### PAPER • OPEN ACCESS

# Analysis of energy consumption in a wireless sensor network using fuzzy memberships function

To cite this article: Wan Isni Sofiah Wan Din et al 2020 IOP Conf. Ser.: Mater. Sci. Eng. 769 012062

View the article online for updates and enhancements.

# Analysis of energy consumption in a wireless sensor network using fuzzy memberships function

Wan Isni Sofiah Wan Din<sup>1</sup>, Mritha Ramalingam<sup>2</sup> and Nuramera Abdullah<sup>3</sup>

Faculty of Computing, Universiti Malaysia Pahang, 26300, Gambang, Pahang, Malavsia<sup>1, 2, 3</sup>

#### sofiah@ump.edu.my1

Abstract. The use of sensors in every application as well as in today's life has become a huge demand. Since that, the public sphere is starting to move to mobile applications that are easy to access. For example, child monitoring at school, health monitoring, tracking system, fire detection and so on. Thus, the Wireless Sensor Network (WSN) is the preferred choice to meet these needs. However, despite the enthusiasm of building various applications using sensors, the WSN itself is still hampered by limited battery usage. Due to some applications that require long battery life, various types of research have been conducted to address this problem. One of these is the involvement of artificial intelligent (AI) in extending the life of a battery on a sensor. Fuzzy Logic (FL) is one of the preferred AI that have been chosen by researchers to be implemented with WSN especially to protract the lifetime of WSN. In Fuzzy Logic, there are three membership functions that need to investigate the capability of it towards the WSN applications. The proposed approach is by combining different types of Fuzzy Logic membership functions which are Triangular with Gaussian, Gaussian with Trapezoidal and Trapezoidal with Triangular to get the best results for analysing the use of sensor batteries. The parameters involved for the cluster head selection are communication cost, centrality and residual energy as Fuzzy inputs. This approach will use an existing algorithm which is Multi-Tier Algorithm (MAP) and this is a part of the MAP enhancement towards the WSN lifetime. The results will compare, discuss and analyst the number of dead nodes and energy usage of the sensor node during data transmission. In conclusion, through this approach, it able to prolong the lifetime for the sensor network since the proposed technique can reduce the energy usage of the sensor nodes.

#### 1. Introduction

In the past few years, the use of sensors has become wider due to the advances in technologies such as in vehicles, manufacturing, wireless communication, agriculture, healthcare and many more. A WSN is known as tiny devices that are used for monitoring and reporting the environmental conditions like pressure, sound, humidity, image and temperature. Typically, a WSN comprises of hundreds to thousands of sensor nodes and there must allocate at least one or more base stations in sensing area for user to access the data. The sensor nodes will collect all the data or information such as videos, audios, humidity, temperature, smoke and any others relevant data. After that, all the data will be placed at base station and user can retrieved the data from that. The base station will act as interference or in other names, an intermediary between users and network where it will observe and analyze the data after gathering them from the sensor nodes. The application of WSN that has been used nowadays in

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under licence by IOP Publishing Ltd 1

our daily life like the water quality detection, intrusion detection, tracking for military application, forest fire detection and recent development is in air pollution monitoring application [1] [2].

Unfortunately, there is a major constraint of WSNs where it has limited energy resource inside the batteries and it will be impossible in replacing the batteries when WSN is implemented far from the power source. Thus, it gave a bad impact on the lifetime of the network since it cannot complete the monitoring task. Due to this issue, the solution must be addressed and one of the effective methods that can be used to optimize the usage of energy and prolong the lifetime of the network is called clustering [3]. Clustering is one of the method to organizes all the distribute sensor nodes in numerous clusters and for every cluster that exist in the network, it must have at least a head of cluster. The cluster head may be assigned by the members of the cluster and the cluster head will become the intermediary between the base station and the other nodes. The cluster head will gather all the information that the other nodes have sent to them and forward the data to the base station [4] [5].

In [6] the researcher proposed a new clustering algorithm that forms a multitier network or often called as "n-tiers" topology to lengthen the lifetime of WSN. The two-tier technique has been designed. The author used Fuzzy Logic technique to choose the cluster head [6]. For the input parameters, they used remaining energy power, the central of the sensor node and also the cost of the communication. From aforementioned parameters, each of that will be combined to satisfy the characteristic of cluster head such as combination of energy with centrality, or energy with cost or cost with energy and centrality. In each tiers, it must locate cluster head as it responsible to collect the data and transmit to the base station. For the data transmission, it involves another two representatives in the network. The sensor node will send the data to their respected cluster head. Cluster head will pass the data to its primary node and then primary node will find the nearest secondary cluster head before all the data safely arrived at base station. Hence, this paper proposed three combinations of fuzzy logic membership function such as Gaussian with Triangular, Triangular with Trapezoidal and Trapezoidal with Gaussian. All these three combination will be run into multi-tier network topology to identify which combinations can enhanced the lifetime of sensor network.

#### 2. Literature review

There are a lot of studies done specifically for energy consumption and optimization issues in Wireless Sensor Network. Many studies had proposed using Artificial Intelligent (AI) techniques in order to solve this main issue that always occurs in WSN. For example, the author in [7] has proposed the Genetic Algorithm (GA) to minimize the communication distance between the sensor nodes and cluster heads at the same time thus enhanced the WSN lifetime. In their work, a cluster-based method is used where there are two phases proposed known as set-up phase and steady-state phase. In the setup phase, the clusters are created and not changed in the whole network. The selection of the new cluster head is done by selecting a node that has the highest residual energy if the current cluster head has residual energy below average after being checked at the end of each round. This cluster-based and GA method was proven helps to expand the lifetime of wireless sensor network after the simulation is done.

According to Gupta et al. [8], a fuzzy logic approach is proposed to depend on three parameters energy, centrality and the concentration to choose a suitable node as the cluster head in every cluster. On that paper, fuzzy logic will help in minimizing the energy used in every node at once increase the lifetime of the network in WSN by using three parameters - energy level in each node, the nodes centrality and the nodes concentration. In order to get a better result, fuzzy logic can blend the different parameters that are proposed together to produce the best result. There is a research where the author proposed an enhancement for clustering in wireless sensor network which also implemented Fuzzy Logic [9]. They used three parameters which are energy of the sensor, the range to the base station and also the concentration of the sensor nodes.

Meanwhile, according to a study by Rostami and Marzieh et al [10], a new approach has been introduced for clustering WSN based on particle swarm optimization (PSO) that aim to extend the network lifetime by using fitness function. This study proposed a new cost function for the selection of

cluster heads that combined different criteria that affects the energy efficiency in cluster heads. There are five parameters that were used in this study and they are residual energy, the distance of cluster head from each other, distance from the base station, density and lastly, the intra-cluster distance. In this study, all the sensor nodes used power control to adjust the power transmission that is dependent on the distance to the receiver. For clustering, the distribution of particles is done randomly and nodes with the best point are appointed as the cluster heads and the other member nodes will form clusters. All of the decisions such as the cluster head selection and clustering were performed in the base station as the proposed algorithm is a centralized algorithm. After the simulation is done, the result showed that the proposed method is way effective in terms of energy consumption and network lifetime compared to other methods that have been used in the previous studies such as LEACH, CHEF and PSO-MV [11].

The Multi-tier Protocol (MAP) is another approach that has been introduced for selecting the optimal cluster head in WSN. It used Fuzzy Logic approach with three suitable parameters to identify the cluster head. From three input with Fuzzy Logic, it will produced one output and will be ranked to identify the potential nodes to become a cluster head [6]. The transmission between the sensor node and cluster head using multi-hop clustering technique. In MAP, each of the network tiers will have a cluster head, for example, the two-tier network clustering, the cluster head in second-tier will responsible to collect and aggregate the packet from other sensor nodes in the same tier, then the information will be forward to the nearest cluster head in a first tier. The process of cluster head in first-tier will be same, the packet will collect from the sensor node in first-tier members and the cluster head in the second tier, and then it will directly send to the base station. It will keep continuing the process until the cluster head loss of energy, then the new cluster head will be select based on K-Optimal formula [12].

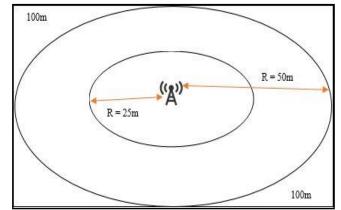
## 3. Methodology

The proposed algorithm is stimulated by using MATLAB tools. Table 1 shows the preliminary design setup for the whole network and values of metrics for the simulation process. Two-tier network design was used in order to compare the result with the previous research [13].

Table 1. Preliminary Design Setup				
Parameter	Value			
Number of Tiers Used	2			
Network Diameter	100M x 100M			
Number of Nodes Used	100			
Base Station Coordination	(50,50)			
Energy Used (Joule)	(1)			
Data (Bits)	2000, 4000, 6000,8000			

Figure 1 shows the proposed two-tier network design. The location of the base station had been set with coordinate (50, 50). The sensor area measured for tier 1 radius from the base station is 25 meters meanwhile for tier 2 is 50 meters.

IOP Conf. Series: Materials Science and Engineering 769 (2020) 012062 doi:10.1088/1757-899X/769/1/012062





For the proposed approach, three parameters have been identified which are energy, centrality and communication cost. All these three parameters will used by 100 sensor nodes and three type of Fuzzy Logic membership function will involve to select the cluster head. The Fuzzy Logic membership function were Triangular, Gaussian and Trapezoidal. From the three input, Fuzzy Logic will only produce one output. The purpose of blending the parameters with different combinations of the membership function is to identify the best combination that can produce the output hence improving the energy efficiency for cluster head selection and prolonging the lifetime of the wireless network.

Figure 2 shows the Fuzzy inference system for three parameters proposed. As mentioned previously, all three parameters will be blended into Fuzzy Logic to produce an output. For each parameter, it will have three fuzzy sets. Centrality was presented by Far (F), Near (N) and Satisfactory (S). Meanwhile, for residual energy and communication cost, Low (L), Medium (M) and High (H) were used.

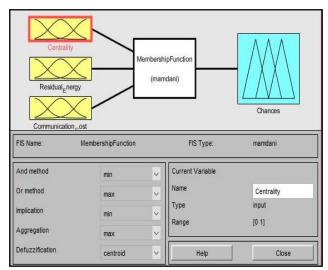


Figure 2. Fuzzy Inference System (FIS) for three parameters

Figure 3 shows the fuzzy sets that represent the parameter of centrality which is Far (F), Near (N) and Satisfactory (S). For Far, Near and Satisfactory, the value is (0-0.5, 0-1 and 0.5-1) respectively. For example to the respective sensor, if the centrality of the sensor node is near to the base station, the residual energy is high and the communication cost low, the chances of the sensor to become a cluster head is high compare to the other sensor nodes. Therefore, for all the sensor nodes, the output chances will be ranked and the sensor with the higher chances will become a cluster head.

Since there are three parameters that are involved in this study, so each input variable will have three linguistic states. Therefore, it specifies the total number of fuzzy inference rules which are a total of 27 rules ( $3x_3x_3 = 27$ ). Table 2 shows the total of 27 rules for the three parameters – residual energy, communication cost and centrality. The output of the variables indicates the chances of each sensor nodes to be the cluster head. The higher the output, the higher the chances and possibility of the sensor nodes to be the cluster head.

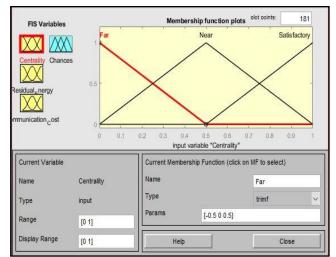


Figure 3. Membership function for parameter centrality

<b>Table 2.</b> Fuzzy rules [6]
---------------------------------

Rule	Centrality	Residual Energy	Communication Cost	Output Chances [6]		
		[6]	[6]	-		
1	Far	High	High	Normal		
2 3	Near	High	High	Normal		
3	Satisfactory	High	High	Strong		
4	Far	High	Medium	Normal		
5	Near	High	Medium	Strong		
6	Satisfactory	High	Medium	Strong		
7	Far	High	Low	Strong		
8	Near	High	Low	Strong		
9	Satisfactory	High	Low	Strong		
10	Far	Medium	High	Weak		
11	Near	Medium	High	Normal		
12	Satisfactory	Medium	High	Normal		
13	Far	Medium	Medium	Normal		
14	Near	Medium	Medium	Normal		
15	Satisfactory	Medium	Medium	Strong		
16	Far	Medium	Low	Normal		
17	Near	Medium	Low	Strong		
18	Satisfactory	Medium	Low	Strong		
19	Far	Low	High	Weak		
20	Near	Low	High	Weak		
21	Satisfactory	Low	High	Normal		
22	Far	Low	Medium	Weak		
23	Near	Low	Medium	Normal		
24	Satisfactory	Low	Medium	Normal		
25	Far	Low	Low	Normal		

26	Near	Low	Low	Normal
27	Satisfactory	Low	Low	Strong

#### 4. Result and discussion

#### 4.1. Analysis of dead node for different data bits

This experiment was run by using MATLAB tool where the combination of the fuzzy membership function is implemented in Fuzzy Logic Controller in order to select the cluster head in the two-tier network. In order to measure the effectiveness of the cluster head, the analysis will be done based on data transmission. This experiment involves 100 sensor nodes and the data used in this experiment were 2000, 4000, 6000 and 8000 bits based on LEACH protocol [14] [15]. The maximum iteration is at maximum 20000 iterations. In order to maintain the consistency of the result for the lifetime of sensor nodes, we have compared with 4000, 6000 and 8000 bits of data. Table 3 shows the iteration of the dead node on different bits of data starting from 2000 to 8000 for the combination of MF that has been proposed in this study.

Based on Table 3, it was observed that the combination of Trapezoidal and Gaussian MF survived longer for every testing that has been done on different data bits compared to the combination of Triangular and Trapezoidal MF. Trapezoidal and Gaussian MF has better performance because the Gaussian MF provided smooth transitions and provided a way to stimulate the maximum number of rules in the rule base thus an accurate representation of input and output relationship is achieved [16]. In addition, with Trapezoidal MF where its width can be adjusted and at once produced a desirable result. When the width of Trapezoidal MF is adjusted to a higher number, the greater the result would be. It was concluded that, the bigger the data bits, the faster the nodes to die which caused the network lifetime to be shorter. It can be seen at 2000 data bits the range of dead node iteration was at 9000th to 10000th and above while, for 8000 data bits, the range of dead node iteration was literally at 3000th and below which was about 25% to 27 % less than 2000 data bits. The combination of Trapezoidal and Gaussian is proven to be efficient and helps to prolong the lifetime of the wireless sensor network.

MF								
Combination	20	000	40	000	60	00	80	00
	First	Last	First	Last	First	Last	First	Last
Triangular +	$7448^{th}$	10121 <sup>st</sup>	3724 <sup>th</sup>	5061 <sup>st</sup>	2482 <sup>nd</sup>	$3374^{th}$	1862 <sup>nd</sup>	2530 <sup>th</sup>
Trapezoidal								
Triangular +	$7448^{th}$	9844 <sup>th</sup>	3724 <sup>th</sup>	4932 <sup>nd</sup>	2482 <sup>nd</sup>	$3288^{th}$	1862 <sup>nd</sup>	2466 <sup>th</sup>
Gaussian								
Trapezoidal	7620 <sup>th</sup>	10507 <sup>th</sup>	3810 <sup>th</sup>	5254 <sup>th</sup>	2540 <sup>th</sup>	3502 <sup>nd</sup>	1905 <sup>th</sup>	2627 <sup>th</sup>
+ Gaussian								

**Table 3.** Dead node for each combination using different data bits

#### 4.2. Analysis of energy used in node 50th using different data bits

Figure 4 shows the graph of energy possessed in 50th node in Joule versus the number of node in 2000 bits and 8000 bits of data for every combination of membership function proposed which were Triangular and Trapezoidal, Triangular and Gaussian and Trapezoidal with Gaussian. There are 3 areas that have been marked in these graphs that represent the highest energy possessed based on the combination of MF proposed.

From Figure 4, it was observed that the energy possessed for 2000 and 8000 bits' data are almost in the same pattern. For the first 2000 bits' data, the total energy possessed recorded to be higher compared to 8000 bits. The total energy possessed for 8000 bits is recorded as the lowest because, at

the beginning of the iteration, the nodes used a lot of energy to send bigger bits of data. This can be seen in Table 4 where the energy was recorded for 3 areas that indicated the combination of MF proposed. This explained why the energy in 50th node is lower because the nodes before have consumed a lot of energy for data transmission in the network.

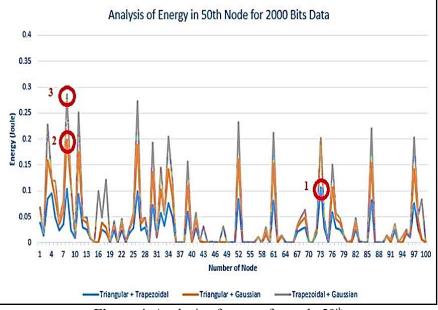


Figure 4. Analysis of energy for node 50<sup>th</sup>

Area	The energy in 2000 bits	The energy in 8000 bits
1	0.0953706599998195	0.0952539200000434
2	0.093994300000205	0.093859200000019
3	0.08860572000202	0.08863592000002

# 4.3. Analysis Of Data Sent By 100 Nodes During Data Transmission

During the data transmission phase, every sensor node will send the different size of data to its appointed cluster head. The cluster head then will gather all the data before sending it directly to the base station. In this study the data from sensor node in tier one is sent to its own cluster head and then, the cluster head will compress, aggregate and analyse before sending it to the PCH appointed in its own cluster. The PCH will send the data that is received from a cluster head in tier 1 directly to the base station. Meanwhile, in tier 2 all node will send their data to the SCH in their own cluster and SCH will compress and aggregate all the data that have been received and transmit them to the nearest PCH in tier one. After PCH received all the data from SCH, it will instantly send the data to the base station. Figure 5 showed the graph of the number of data versus the number of nodes where all data sent for 100 sensor nodes for 4000 bits is recorded and plotted. As we can see, there are several nodes recorded to have data sent above 350000000. The highest data sent for the combination of Triangular Trapezoidal and Triangular Gaussian are at the same which were 3500000000 bits at node 28th. Meanwhile, the highest data sent recorded to be at 48th node which for the combination of Trapezoidal and Gaussian was 361096000 bits.

IOP Conf. Series: Materials Science and Engineering 769 (2020) 012062 doi:10.1088/1757-899X/769/1/012062

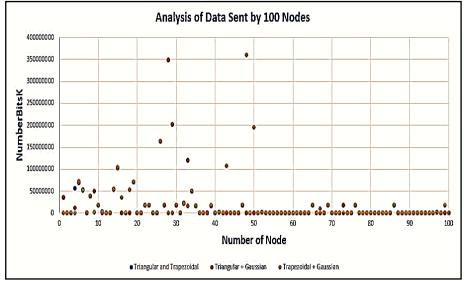


Figure 5. Numbers of Data vs Number of Nodes

Figure 6 shows the total of data sent by all 100 sensor nodes to the selected cluster head in the clusters. For the combination of Triangular Trapezoidal, it showed that they had the lowest total number of data sent which was only 1633692000 bits followed by 1635408000 bits for the combination of Triangular Gaussian function. While, the combination of Trapezoidal and Gaussian function recorded the highest total number of data sent, 1651532000 which was 0.54% higher than the total data of the combination of Triangular and Trapezoidal function. The total number of data sent in this combination recorded the highest because the iteration can sustain up to 5000 iterations. This means the longer the iteration, the more data that can be sent by every sensor node in the network.

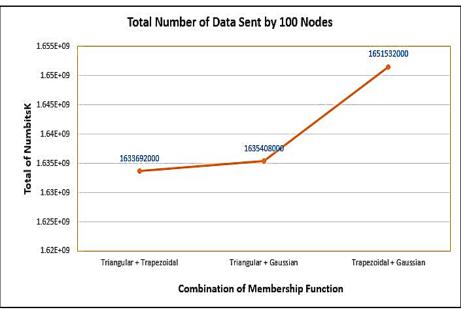


Figure 6. Total Data for each Combination of Membership Function

#### 5. Conclusion

The cluster head selection successfully implemented using Fuzzy Logic approach. Three parameters have been used which are energy, centrality and communication cost. From these three parameters, with the Fuzzy Logic approach it can produce one output and the highest output will be selected as a highest chances for the sensor nodes to become a cluster head. There are also three Fuzzy Logic membership function that involved in this proposed method which are Triangular. Trapezoidal and Gaussian. Based on that, we combined the membership function as; Triangular and Trapezoidal, Triangular and Gaussian and lastly Trapezoidal with Gaussian. The effectiveness of the cluster head gave a big impact on network lifetime [17]. Hence, it is important to make sure which combination of fuzzy membership function helps to prolong the network lifetime [18] [19]. In this study, the effectiveness of cluster head selected is measured based on data transmission between the nodes in the cluster. It was observed that, the smaller the data, the longer the network would survive. By using data up to 6000 bits it is proven to shorten the network lifetime as the energy used by the nodes higher. The nodes can sustain up to 10000 iterations when using 2000 bits' data which is 20 % longer than using 8000 bits' data. In this study, it was observed that the combination of Trapezoidal and Gaussian function is the best to use as it can support up to 1000 of iterations. Besides, this combination also has the highest total number of data sent by all 100 nodes which were 0.54 % higher compared to the combination of Triangular Trapezoidal and Triangular Gaussian. Thus, it concluded that the implementation of Fuzzy Logic with the help of the combination of Trapezoidal and Gaussian membership function has allowed the nodes use their energy effectively and helps in prolonging the network lifetime.

### Acknowledgement

The author would like to thank UMP RDU Grants (RDU-180364) for providing financial support and facilities in this study. Also gratefully, we would like to acknowledge and thank the Faculty of Computing and Systems Network & Security (Sysnet) research group UMP that support this research.

#### References

- J. A. Nayak, "Improving the Network Life Time of Wireless Sensor Network using EEEMR Protocol with Clustering Algorithm," Int. J. Electron. Electr. Comput. Syst. IJEECS, vol. 6, no. 6, pp. 342–346, 2017.
- [2] K. Okokpujie, M. Odusami, E. Noma-Osaghae, O. Modupe, S. John, and O. Oluwatosin, "A Smart Air Pollution Monitoring System," Int. J. Civ. Eng. Technol. (IJCIET, vol. 9, no. 9, pp. 799–809, 2018.
- [3] A. A. Qasem, A. E. Fawzy, M. Shokair, W. Saad, S. El-Halafawy, and A. Elkorany, "Energy Efficient Intra Cluster Transmission in Grid Clustering Protocol for Wireless Sensor Networks," Wirel. Pers. Commun., vol. 97, no. 1, pp. 915–932, 2017.
- [4] J. Anzola, J. Pascual, G. Tarazona, and R. González, "A clustering WSN routing protocol based on k-d tree algorithm," Sensors (Switzerland), vol. 18, no. 9, pp. 1–26, 2018.
- [5] D. R. Das Adhikary and D. K. Mallick, "An Energy-Aware Unequal Clustering Algorithm using Fuzzy Logic for Wireless Sensor Networks," J. ICT Res. Appl., vol. 11, no. 1, pp. 56– 77, 2017.
- [6] W. I. S. W. Din, S. Yahya, R. Jailani, M. N. Taib, A. I. M. Yassin, and R. Razali, "Fuzzy logic for cluster head selection in wireless sensor network," in AIP Conference Proceedings, 2016, vol. 1774, p. 050006.
- [7] M. Iqbal, M. Naeem, A. Anpalagan, A. Ahmed, and M. Azam, "Wireless Sensor Network Optimization: Multi-Objective Paradigm," Sensors, vol. 15, no. 7, pp. 17572–17620, 2015.
- [8] C. P. Gupta and A. Kumar, "Optimal Number of Clusters in Wireless Sensor Networks with Mobile Sink," Int. J. Sci. Eng. Res., vol. 4, no. 8, pp. 1706–1710, 2013.
- [9] T.-T. Nguyen, J.-S. Pan, S.-C. Chu, T.-K. Dao, and V.-C. Do, "Improved Performance of Wireless Sensor Network Based on Fuzzy Logic for Clustering Scheme," Adv. Smart Veh.

IOP Conf. Series: Materials Science and Engineering 769 (2020) 012062 doi:10.1088/1757-899X/769/1/012062

Technol. Transp. Commun. Appl. VTCA 2018. Smart Innov. Syst. Technol. Springer, vol. 128, pp. 104–113, 2019.

- [10] A. S. Rostami, M. Badkoobe, F. Mohanna, H. Keshavarz, A. A. R. Hosseinabadi, and A. K. Sangaiah, "Survey on clustering in heterogeneous and homogeneous wireless sensor networks," J. Supercomput., vol. 74, no. 1, pp. 277–323, 2018.
- [11] A. A. R and O. H. I, "Performance Evaluation of Selected Distance-Based and Distribution-Based Clustering Algorithms," Int. J. Softw. Eng. Comput. Syst., vol. 4, no. 2, pp. 38–48, 2018.
- [12] A. A. Rahman, M. N. M. Kahar, and W. I. S. W. Din, "Distance-based thresholds for 2-tier relay nodes selection in WSN," Comput. Stand. Interfaces, vol. 66, no. April 2019.
- [13] N. K. Jassi and S. S. Wraich, "An enhanced K-Means clustering technique with Hopfield Artificial Neural Network based on reactive clustering protocol," in 2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence), 2014, pp. 821–825.
- [14] S. Vasuhi and V. Vaidehi, "Target tracking using Interactive Multiple Model for Wireless Sensor Network," Inf. Fusion, vol. 27, no. Supplement C, pp. 41–53, 2016.
- [15] M. Mirzaie and S. M. Mazinani, "Adaptive MCFL: An adaptive multi-clustering algorithm using fuzzy logic in wireless sensor network," Comput. Commun., vol. 111, no. Supplement C, pp. 56–67, 2017.
- [16] S. Ramakrishnan and S. P. Shyry, "Fuzzy-based cluster head election mechanisms for wireless sensor networks: A review," in 2017 International Conference on Information Communication and Embedded Systems (ICICES), 2017, pp. 1–6.
- [17] A. A. Allahham and M. A. Rahman, "A Smart Monitoring System for Campus Using Zigbee Wireless Sensor Networks," Int. J. Softw. Eng. Comput. Syst., vol. 4, no. 1, pp. 1–14, 2018.
- [18] C. Wu, T. Yoshinaga, X. Chen, L. Zhang, and Y. Ji, "Cluster-Based Content Distribution Integrating LTE and IEEE 802.11p with Fuzzy Logic and Q-Learning, "IEEE Comput. Intell. Mag., vol.13, no.1, pp 41-50, 2018.
- [19] A. Saxena et al., "A review of clustering techniques and developments," Neurocomputing, vol. 267, pp. 664–681, 2017.