FAFOC: Fog-Based Energy-Efficient Clustering Technique for Wireless Sensor Networks

R. Dayana and G. Maria Kalavathy

Keywords Fog computing · WSN · Clustering · Fuzzy logic · Artificial flora · Energy efficiency · Network lifetime

1 Introduction

In general, the proficient deployment of Internet of Things (IoT) examines the atmosphere and performs specific operations. Hence, the responsibility of a computing method which manages the IoT devices is highly effective. The IoT devices are mainly composed of sensors, and a collection of ubiquitous sensors forms the wireless sensor network (WSN). Initially, cloud computing (CC) depends upon the WSN and is utilized for managing the sensor data [1]. Transmission and computation of massive data which is passed through CC resulted in various complexities like delayed service response time and additional energy requirement. Moreover, the domains of a WSN has been developed extensively and demands for practical computations and data transmission which finally enhances the IoT management task [2]. A novel computing method has been introduced, named as fog computing; in this process, the volume of data produced by a sensor is maximized, and regular data processing becomes a complicated issue.

Here, fog computing is a technique applied for data management and performs communication with the nearby field using a sensor. Since the devices in fog computing applies near-field communication, the response speed becomes robust and rapid when compared to CC. Besides, the sensor count assigned to a device is minimum than CC. Under the assumption of data aggregation, the power consumption can be reduced which has been caused because of the elimination

R. Dayana (✉)
Department of CSE, Jeppiaar Institute of Technology, Chennai, India

G. Maria Kalavathy
Department of CSE, St. Joseph's College of Engineering, Chennai, India

R. Misra et al. (eds.), *Machine Learning and Big Data Analytics*,
Springer Proceedings in Mathematics & Statistics 401,
https://doi.org/10.1007/978-3-031-15175-0_35

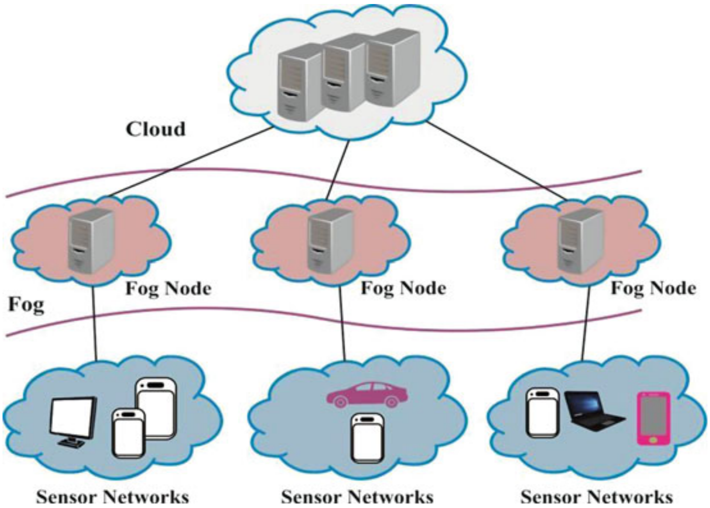

Fig. 1 Architecture of fog-based WSN

of redundant data transmission among nearby devices in a sensor node. On the other hand, the clustering-based hierarchical process offers massive benefits [3–5]. Furthermore, fog computing services are data development, computation, and transmission, as depicted in Fig. 1. Besides, a cloud server is helpful in extracting features from server on the basis of data type when it was actually created. Followed by, a server contains set of rules, pattern analysis, as well as inference for data saving.

Data that is computed is fed into a data owner or authorized server. At this point, data used for a service user has been saved in the form of plain text and transmitted to observing server so that it can be applied in future. Data distribution and applications are the predictable operations. Data which is produced in a sensor might be an ineffective one for the user; hence, a delegation function is to assign a centralized manager for accessing the data, and revocation function is applied to eliminate the authentication for accessing the data as the user required.

Fog computing [6] is defined as a virtual environment which offers computation, storage, and networking service among a device as well as CC data center. Therefore, it has not been placed uniquely at the edge of network. The operations like computation, storage, and networking are considered to be the building blocks of cloud and fog. The cloud layer is considered to be the main theme of fog computing that computes data virtualization, examination, and machine learning (ML) and upgrades the rules and patterns from fog layer's proxies. The proposed fog-based WSN involves a set of two nodes, namely, advanced nodes, normal nodes, and some FNs. The existence of advanced nodes makes the network heterogeneous, which results in maximum energy efficiency of the network. Though the fog based

WSN offers several benefits, there is a need to develop an energy-efficient clustering technique to select the CHs and organize clusters.

This paper presents a new fuzzy with artificial flora optimization-based clustering (FAFOC) algorithm for CHs selection to attain maximum energy efficiency in fog-based WSN. The proposed FAFOC algorithm involves three major stages, namely, node initialization, tentative CH selection, and final CH selection. Initially, the nodes undergo random deployment, and initialization process takes place. Then, the FAFOC selects the CHs in two stages, namely, tentative CH selection and final CH selection. At the initial stage, all the nodes in fog-based WSN undergoes fuzzy logic-based tentative CH selection (FL-TCHS) process which elects a set of tentative nodes based on residual energy (RE), distance to base station (DBS), and node centrality (NC). Then, the tentative CHs execute the artificial flora optimization-based final CH selection (AFA-FCHS) process to select the final CHs. Once the CHs were effectively chosen, the nearby nodes join the CHs and forms clusters. Finally, the cluster members (CM) send the data to CHs, which then forward it to the fog nodes subsequently to cloud. The experimental validation of the proposed FAFOC algorithm is carried out, and the results are examined under several dimensions.

2 Related Works

Energy application management of WSN has gained massive attention from research communities. The developers have examined all other solutions from the application of another power sources, named as solar energy [7], and some other research have focused on the development of power-aware protocol models [8, 9]. Using the objective of enhancing the entire network operation effectively, fog computing is mainly applied for extending the CC model to the network edge [10]. Fog computing is mainly deployed for addressing the issues over mobility, geography, and delay requirements of IoT systems [11]. Because of the demand while installing massive interconnected devices, energy maintenance and self-recovering applications are the two important developing attributes of fog computing environments.

In Al-Fuqaha et al. [12], the major features behind network mechanisms of a fog computing have been examined such as position, distribution, reliability, density of machines, mobility support, real time, and standardization. In addition, the network structure in fog computing devices should be independent and effective with respect to power application as well as network resilience. Due to the network requirements, IEEE 802.15.4 ZigBee standard from ZigBee has inspired the concentration of developers. Diverse researchers are effective and provide the advantages of ZigBee in various Smart City as well as Smart Farming domains. In Lee and Wang [13], the researchers have determined the connection, packet loss rate, and transmission throughput from indoor research region. A hybrid network model, ZigBee/5G, is presented in [14].

The developers have demonstrated efficiency of combining diverse protocols which aims at limiting the energy application. Followed by, the developers from [15] have determined and related the power energy utilization of ZigBee vs. the Advanced and Adaptive Network Technology (ANT). Based on the survey, ANT performs well in ZigBee with respect to power application executed by using sleep and wake technology. Hence, it is stated that, an appropriate management of entire system, radio, and sensors, the energy application addressed by ZigBee systems are reduced gradually. The final outcomes make ZigBee a tremendous environment to learn the advantages of combining energy application management models. Moreover, the extensive application of ZigBee tends to make greater platform for developing Smart City as well as Smart Farming applications. Recently, Piromalis and Arvantis [16] established a reliable hardware structure for WSN and actuator networks which has to be applied in Smart Farming domains. Likewise, the structure of sensor node concentrates on the energy application as well as network resilience methodologies are considered as advanced innovations on fog computing. It refers that the newly developed models takes the structure of sensor node, with main characteristics of fog computing model, especially geographical position, delay, as well as energy provisioning.

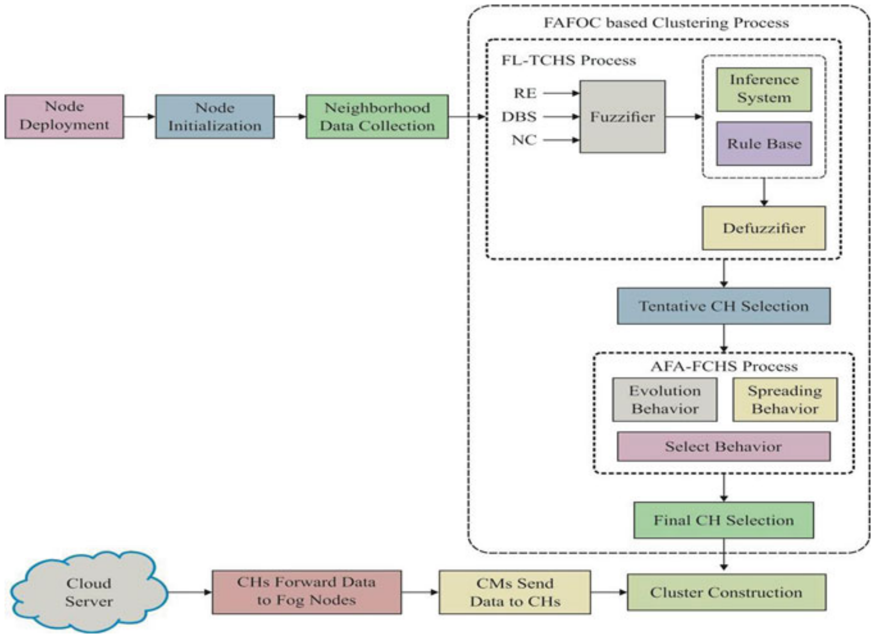

Fig. 2 Block diagram of FAFOC algorithm

3 The Proposed FAFOC Algorithm

The working process involved in the FAFOC algorithm is demonstrated in Fig. 2. The proposed FAFOC algorithm involves three major stages, namely, node initialization, tentative CH selection, and final CH selection. At the beginning stage, the nodes are randomly deployed and initialized, and information collection about nearby nodes takes place. Then, the FL-TCHS process gets executed and selects the tentative CHs. Afterwards, the tentative CHs performs the AFA-FCHS process and elect the final CHs. Once the CHs were chosen, the datatransmission process begins where the data from CMs to cloud takes place via CHs and fog nodes in the network.

3.1 Fuzzy Logic-Based Tentative CH Selection Process

In this section, the tentative CHs were selected using fuzzy logic with the use of three input parameters, namely, RE, DBS, and NC [17].

(i) *Residual Energy*: It plays a significant role for selecting a node as CH which spends maximum energy when compared with CM. A CH node gathers data from members, accumulate the data gathered, and transmits to sink node or BS. Hence, competing energy level is essential for a CH in order to execute the predefined events.

(ii) *Node Centrality (NC)*: Total count of one-hop adjacent nodes inside R_c of a node is named as node degree. NC is a factor that calculates the node's position in the intermediate location over its neighbors. Lower NC measure offers greater opportunity of selecting a node as CH:

$$NC = \frac{\sqrt{\sum_{\mathrm{i=1}}^{\mathrm{ND}} dist_i^2/ND}}{Ntk\_Dimension} \tag{1}$$

In Eq. (3), ND (node degree) corresponds to count of neighbors inside a communication radius R_c of a node, and Ntk-Dimension value is "M" in $M \times M$ field area, as well as $dist_i^2$ implies a distance with *ith* neighbor node which refers that, in 100 m × 100 m, Ntk-Dimension is 100, and, in 200 m × 200 m field area, Ntk-Dimension is 200.

(iii) *Distance to BS*: The power conservation increases while transmitting data and also the distance among a transmitter and receiver nodes. From the point of power conservation, the distance among CH and BS has to be reduced:

$$Distance\,to\;BS = \frac{d_i}{\alpha \cdot Ntk\_Dimensio}, \qquad \alpha = \frac{d_{\max}}{Ntk\_Dimensio} \tag{2}$$

In Eq. (4), d_i refers to the distance among node i and the *BS*, $d_{\max}$ denotes higher distance among a node in a system and *BS*, and A signifies a network dimension constant.

3.2 AFA-Based Final CH Selection Process

Once the tentative CHs are chosen, they compete for final CH selection using AFA. The AFA method is composed of four fundamental units, namely:

- Original plants
- Offspring plants
- Position of plants
- Propagation distance

Initially, the original plant means that the crops are ready for distributing the seeds. Secondly, the offspring plant is considered as the seed of original plant, and it is not suitable for spreading seeds in specific duration. Thirdly, the position of the plant refers the actual localities of the plant. Finally, propagation distance defines the distance of seeds spreading. It is composed of three main behavioral patterns like:

- Evolution
- Spreading
- Select

At the first stage, evolution behavior describes the possibility of time taken by a plant for adapting with the ecological behavior. Secondly, the spreading behavior defines the distribution of the seed. Finally, the select behavior refers that flora might stay alive or becomes expired because of the atmosphere [18].

Evolution Behavior

The original plants distribute seeds over the circle with radius named as propagation distance, which has been developed from propagation distances of the parent as well as grandparent plants.

$$d_j = d_{1j} \times \text{ran}\,(0, 1) \times c_1 + d_{2j} \times \text{rand}\,(0, 1) \times c_2 \tag{3}$$

where $d_{1\text{j}}$ and $d_{2\text{j}}$ denotes the propagation distance of the grandparent and parent plants, c_1 and c_2 defines the learning coefficients, and *rand*(0,1) stands for autonomous and uniformly distributed within (0,1).

A novel grandparent propagation distance is as follows:

$$d'_{1j} = d_{2j} \tag{4}$$

The novel parent propagation distance is meant to be a standard deviation (SD) among the positions of original plant as well as offspring plant:

$$d'_{2j} = \frac{\sqrt{\sum_{i=1}^{N} \left(_{i,j} - P'_i \right)^2},}{N} \tag{5}$$

Spreading Behavior

Initially, the AFA has produced the data randomly from original flora with N solutions, where N plants are existed in flora. The location of the original plant is represented by matrix $P_{i,}$ where i denotes a dimension and j refers to the flora plant count.

$$P_{i.j} = \text{ran}\,(0, 1) \times d \times 2 - d \tag{6}$$

where d indicates the higher limit area and *rand*(0,1) denotes the array of arbitrary values which are distributed uniformly from (0,1) .

The location of the offspring plant can be produced on the basis of a propagation function in the following:

$$P'_{i,j \times m} = D_{i,j \times m} + P_{i,j} \tag{7}$$

where m refers count of seeds which a single plant propagates, $P'_{i,j \times m}$ means a location of offspring plant, $P_{i,j}$ stands for location of original plant, and $D_{i,j \times m}$ signifies a random value with the Gaussian distribution of mean 0 and *varianced*$_j$. When none of the offspring plant existed, then a new plant would be produced on the basis of Eq. (6).

Select Behavior

The active plant can be computed using survival probability as given in the following:

$$p = \left| \frac{\sqrt{F\left(P'_{i,j \times m}\right)}}{F_{\max}} \right| \times Q_x^{(j \times m - 1)} \tag{8}$$

where $Q_x^{(j \times m-1)}$ is Q_x to the power of (j $\times$ m $-$ 1) and Q_x implies the selective probability. It can be viewed that the fitness of offspring plants are far away from

Table 1 Parameter setting

Parameters	Value
Area	$100 \times 100\ m^2$
$^{E}0$	0.5 J
Node count	500
$^{E}elec$	50 nJ/bit
$^{\varepsilon}fs$	10 pJ/bit/m^2
$^{\varepsilon}mp$	0.0013 pJ/bit/m^4
Packet size	4000 bits

actual plant. Q_x estimates the searching ability capability of the method. Q_x has to be maximum for the problem which is simple to enter the local optimal solution. $F_{\max}$ refers to the higher fitness in flora, and $\left(P'_{i,j\times m}\right)$ denotes the fitness of j-th solution.

Roulette wheel selection approach has been applied for determining whether the plant is active or not. The main objective of this model is to "accept according probability"; which means that there may various instances and it is composed of a potential value. Therefore, selection is completely based on value of potential score. The maximum the score, then the higher the accepting probability becomes.

4 Performance Validation

In this section, an extensive experimental analysis is carried out, and the results are examined under diverse aspects. The parameter settings involved in the simulation process is shown in Table 1.

4.1 Network Lifetime Analysis

Figure 3 shows the network lifetime analysis of the FAFOC model in terms of FND, HND, and LND. The figure portrayed that the FAFOC model has delayed the death of the first as well as the last node in the network. It is noted that the FND of the FAFOC model is 391 rounds, whereas the FND of FEAR, P-SEP, QACMOR, and AFA is 75, 198, 88, and 111 rounds, respectively. At the same time, it is observed that the FAFOC model has attained the HND of 1315 rounds, whereas the HND of FEAR, P-SEP, QACMOR, and AFA is 455, 984, 625, and 711 rounds, respectively. At last, it is exhibited that the FAFOC model has delayed the LND to 1989 rounds , whereas the LND occurs at the earlier rounds of 995, 1612, 544, and 1329 by FEAR, P-SEP, QACMOR, and AFA algorithms.

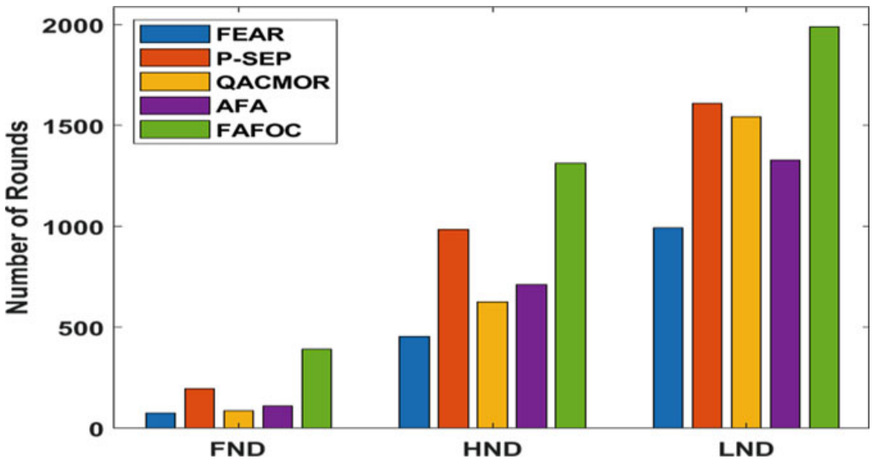

Fig. 3 Network lifetime analysis of FAFOC model

Figure 4 observes the performance of the FAFOC method in terms of PDR under varying node count. The figure exhibited that the FEAR model has depicted PDR over the compared approaches. Concurrently, the QACMOR and AFA models have illustrated certainly higher PDR over the FEAR model. However, the P-SEP model has tried to demonstrate moderate PDR. Yet, the highest PDR has been achieved by the projected FAFOC algorithm. For instance, under the node count of 50, the PDR by the presented FAFOC model is 0.847, whereas the PDR by FEAR, P-SEP, QACMOR, and AFA is 0.777, 0.677, 0.731, and 0.747 correspondingly. Also, under the node count of 500, the PDR by the projected FAFOC model is 0.727, while the PDR by FEAR, P-SEP, QACMOR, and AFA is 0.657, 0.512, 0.527, and 0.567, correspondingly.

Figure 5 showcases the performance of the FAFOC method with respect to average residual energy under varying number of rounds. The figure demonstrated that the FEAR method has shown minimum average residual energy over the compared methods. At the same time, the QACMOR and AFA methods have illustrated certainly superior average residual energy over the FEAR model. However, the P-SEP model has tried to demonstrate moderate average residual energy. However, the highest average residual energy has been attained by the presented FAFOC algorithm. For instance, under the round of 500, the average residual energy by the presented FAFOC model is 0.489J, while the average residual energy by FEAR, P-SEP, QACMOR, and AFA is 0.277J, 0.4875J, 0.3925J, and 0.3425J, respectively. Similarly, under the round number of 1500, the average residual energy by the proposed FAFOC model is 0.35J, whereas the average residual energy by FEAR, P-SEP, QACMOR, and AFA is 0J, 0.225J, 0.1J, and 0.09J, correspondingly.

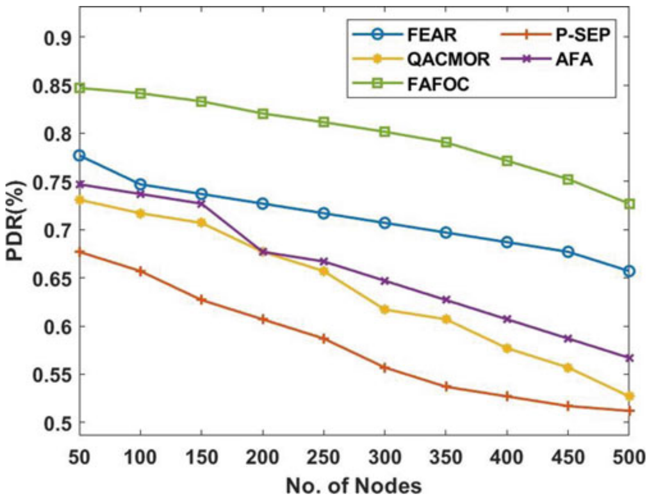

Fig. 4 PDR analysis of FAFOC model

Figure 6 depicts the performance of the FAFOC method with respect to packet loss under varying node count. The figure demonstrated that the FEAR model has shown the highest packet loss over the compared methods. At the same time, the QACMOR and AFA models have shown certainly lower packet loss over the FEAR model. However, the P-SEP model has tried to demonstrate minimum PDR. However, the low packet loss has been achieved by the projected FAFOC method. For instance, under the node count of 50, the packet loss by the projected FAFOC model is 0.02, while the PDR by FEAR, P-SEP, QACMOR, and AFA is 0.2, 0.11, 0.056, and 0.04, correspondingly. Likewise, under the node count of 500, the PDR by the presented FAFOC model is 0.14, but the packet loss by FEAR, P-SEP, QACMOR, and AFA is 0.489, 0.275, 0.26, and 0.22, respectively.

5 Conclusion

This paper has developed a new FAFOC algorithm for CHs selection to attain maximum energy efficiency in fog-based WSN. Initially, the nodes are randomly deployed and initialized, and information collection about nearby nodes takes place. Then, the FL-TCHS process gets executed and selects the tentative CHs. Afterwards, the tentative CHs performs the AFA-FCHS process and elect the final CHs. Once the CHs were chosen, the data transmission process begins where the data from CMs to cloud takes place via CHs and fog nodes in the network. An extensive experimental analysis is carried out, and the results are examined under

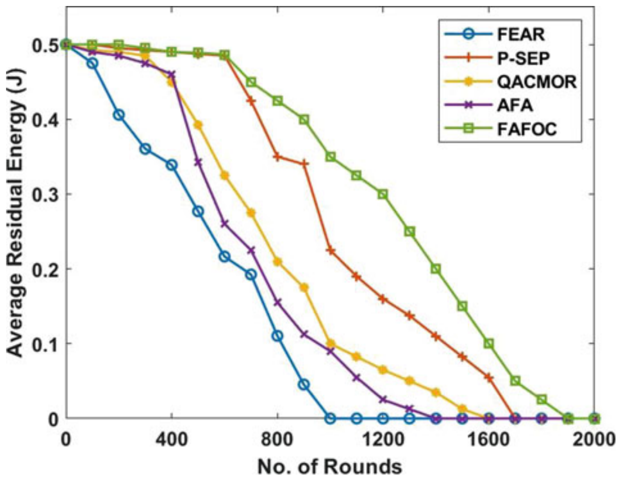

Fig. 5 Average residual energy analysis of FAFOC model

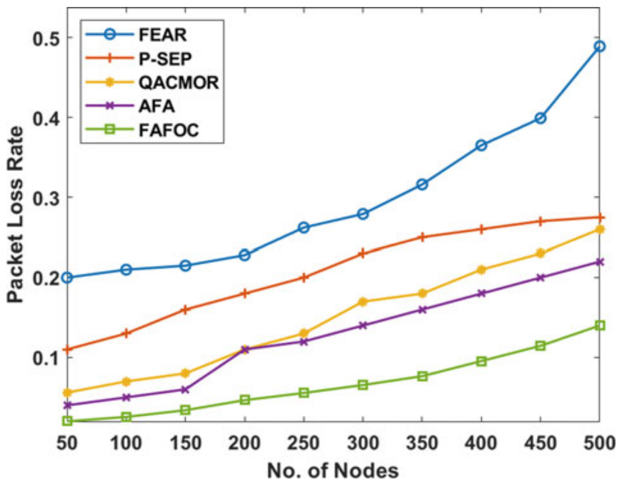

Fig. 6 Packet loss rate analysis of FAFOC model

diverse aspects. The experimentation outcome ensured the effective performance of the FAFPC algorithm in terms of energy efficiency, network lifetime, packet loss, PDR, and delay. As a part of future work, the network lifetime can be lengthened by the use of bio-inspired algorithm-based routing protocol to select the optimal forwarding node set.

References

1. Doukas, C.; Maglogiannis, I. Bringing IoT and cloud computing towards pervasive healthcare. In Proceedings of the 2012 Sixth International Conference Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), Palermo, Italy, 4–6 July 2012; pp. 922–926.
2. Stojmenovic, I.; Wen, S. The Fog computing paradigm: Scenarios and security issues. In Proceedings of the 2014 Federated Conference Computer Scienceand Information Systems (FedCSIS), Warsaw, Poland, 7–10 September 2014; pp. 1–8.
3. Krishnamachari, L.; Estrin, D.; Wicker, S. The impact of data aggregation in wireless sensor networks. In Proceedings of the 22nd International Conference on Distributed Computing Systems Workshops, Vienna, Austria, 2–5 July 2002; pp. 575–578.
4. Lindsay, S.; Raghavendra, C.S.; Sivalingam, K.M. Data gathering in sensor networks using the energy delay metric. In Proceedings of the 15th International Parallel & Distributed Processing, 23–27 April 2001; IEEE Computer Society: Washington, DC, USA, 2001; p. 188.
5. Tan, H.O.; Korpeoglu, I. Power efficient data gathering and aggregation in wireless sensor networks. ACM Sigmod Record 2003, 32, 66–71.
6. Arkin, H.R.; Diyanat, A.; Pourkhalili, A. MIST: Fog-based data analytics scheme with cost-efficient resource provisioning for IoT crowdsensing applications. J. Netw. Comput. Appl. 2017, 82, 152–165.
7. W. B. Pramono, P. Setiawan, and Firdaus, "Solar power supply for ZigBee wireless sensor network," in 2016 International Seminar onApplicationfor Technology of Information and-Communication (ISemantic), pp. 336–340, Semarang, Indonesia, 2016.
8. T. Matsui and H. Nishi, "ECORS: energy consumptionoriented route selection for wireless sensor network," in 2016 IEEE 14th International Conference on Industrial Informatics (INDIN), pp. 1044– 1049, Poitiers, France, 2016.
9. C. Pérez-Garrido, F. González-Castaño, D. Chaves-Díeguez, and P. Rodríguez-Hernández, "Wireless remote monitoring of toxic gases in shipbuilding," Sensors, vol. 14, no. 2, pp. 2981–3000, 2014.
10. F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in MCC '12 Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, pp. 13– 16, New York, NY, USA, 2012.
11. D. Spirjakin, A. Baranov, A. Karelin, and A. Somov, "Wireless multi-sensor gas platform for environmental monitoring," in 2015 IEEE Workshop on Environmental, Energy, and Structural Monitoring Sy
12. A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: a survey on enabling technologies, protocols, and applications," IEEE Communications Surveys & Tutorials, vol. 17, no. 4, pp. 2347–2376, 2015.
13. J.-S. Lee and Y.-M. Wang, "Experimental evaluation of ZigBeebased wireless networks in indoor environments," Journal of Engineering, vol. 2013, Article ID 286367, 9 pages, 2013.
14. J. Mu, "An improved AODV routing for the ZigBee heterogeneous networks in 5G environment," Ad Hoc Networks, vol. 58, pp. 13 –24, 2017.
15. S. Gharghan, R. Nordin, and M. Ismail, "An ultra-low power wireless sensor network for bicycle torque performance measurements," Sensors, vol. 15, no. 5, pp. 11741–11768, 2015.
16. D. Piromalis and K. Arvanitis, "SensoTube: a scalable hardware design architecture for wireless sensors and actuators networks nodes in the agricultural domain, " Sensors, vol. 16, no. 8, p. 1227, 2016.
17. Balakrishnan, B. and Balachandran, S., 2017. FLECH: fuzzy logic based energy efficient clustering hierarchy for nonuniform wireless sensor networks. *Wireless Communications and Mobile Computing*, *2017*.
18. Cheng, L., Wu, X.H. and Wang, Y., 2018. Artificial flora (AF) optimization algorithm. *Applied Sciences*, *8*(3), p.3.