



Fuzzy based novel clustering technique by exploiting spatial correlation in wireless sensor network

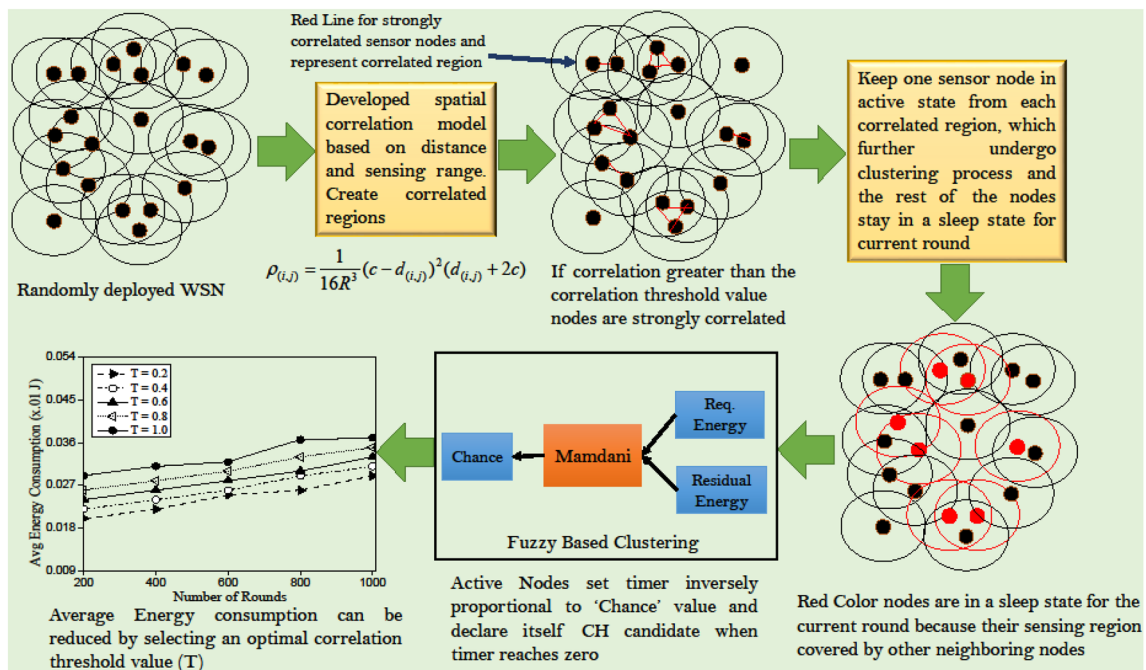
Manjeet Singh¹  · Surender Kumar Soni¹

Received: 20 November 2017 / Accepted: 5 June 2018 / Published online: 12 June 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In wireless sensor networks, the event is detected by multiple closely placed sensor nodes. The spatial relationship can be utilized productively in order to conserve the power banks by halting some sensors to transmit the same information. This paper deals with the segregation of network into the correlated clusters based on correlation value. On the one hand, unlike existing clustering techniques relying on residual energy and distance to select cluster heads, this paper defines more realistic three-dimensional correlation model where cluster heads are elected on the basis of the correlation value, residual energy, and required energy. On the other hand, other than developing theoretical three-dimensional correlation model, a fuzzy-based clustering technique is also proposed to further implement the developed correlation model, where the nodes with similar information are gathered in such a way that data from a solitary node suffices the fidelity constraint to the sink. The effects of node density, sensing range, and the threshold value is studied in detail. Also, the correlation model is clubbed with clustering technique to further take the advantages of exploiting spatial correlation at the network layer. The results have revealed that proposed approach extend network lifetime by 30, 35 and 78% as compared to the FBUC, CHEF, and LEACH respectively. The results of clustering using correlation model show that the number of participating nodes get reduced by 33% when correlation threshold value is decreased from 0.8 to 0.6. Also, it is found that network lifetime gets improved by decreasing the correlation threshold value.

Graphical abstract



Keywords Spatial correlation · Wireless sensor network · Correlated clusters · Fuzzy clustering

1 Introduction

The wireless sensor networks (WSNs) are intended to work in a self-organized way where each sensor is capable of scanning and sensing the surrounding phenomenon. Sensors in the network are event-driven and work in a collaborative manner. On the occurrence of any event, the sensor nodes in the vicinity of the event get activated. After event detection, they all try to send detected information to the sink at the earliest opportunity (Vuran et al. 2004). This kind of situation can also be related to flooding which further increases the likelihood of collision and contention for channel access. In general, it is not energy efficient that all activated sensors send sensed information to the sink because communication bandwidth and residual energy are the two main constraints of the sensor network (Vuran and Akyildiz 2006). In order to conserve battery, if only a few selected sensor nodes among all the activated sensors send information then overall energy consumption can be reduced. This objective of reducing energy consumption can be achieved, if redundant sensor nodes switch to sleep mode periodically in a legitimate manner. An example of the redundant sensor node is depicted in Fig. 1. As shown the node C is redundant node because the region covered by C is also covered by

neighbor nodes A, B, D, and E. Also, it can say that information sensed by node C is strongly correlated with other four sensor nodes. Therefore, while A, B, D, and E are in the active state, node C can switch to the sleep state in order to conserve battery power. A particular region in sensor field where sensor nodes have a high correlation of sensed information is called as Spatial Correlation region and it is sufficient to send information of a single node to represent the spatial correlation region (Zeng and Tang 2011). The readings of a sensor node may be predicted from other strongly correlated sensor nodes in the vicinity having a high-value correlation between them (Vuran and Akyildiz 2006; Zheng and Tang 2011). The main aim of exploiting the spatial correlation between sensors is to reduce the number of reporting nodes because of highest energy expenditure of transceiver system. In previous techniques, according to (Heinzelman et al. 2000, 2002; Yuan et al. 2011), the number of bits transmitted can be reduced by partitioning the whole network into clusters and send only processed information of each cluster to the sink. Although, in clustered hierarchical architecture the information collected from all the activated sensor nodes by their respective cluster heads.

In these approaches, the number of reporting nodes is not reduced and they mainly focused on data aggregation methods which further possess many limitations. In a clustered

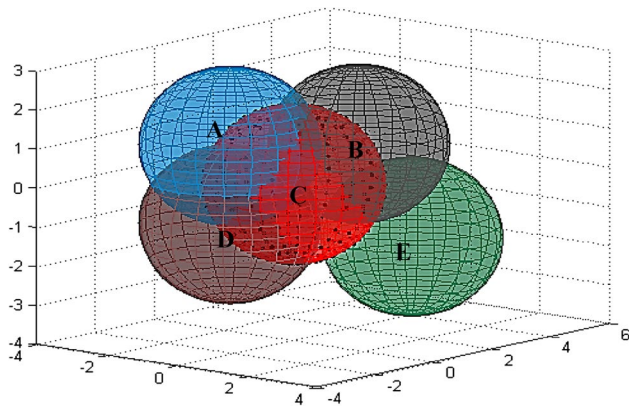


Fig. 1 Coverage area of nodes, indicating that node C is redundant

architecture, the biggest challenge is to select most suitable sensor node as a cluster head by considering various selecting parameters such as energy, distance, node density, the rate of retransmission, locality etc. The cluster heads should be chosen carefully because they directly affect the network lifetime. In Gupta et al. (2005), Kim et al. (2008), Pires et al. (2011), Jin et al. (2011), Lee and Cheng (2012), Mhemed et al. (2012), Izadi et al. (2013, 2015), Nayak and Devulapalli (2016) and Singh et al. (2016), authors have highlighted the limitations of clustering techniques. They further have presented the methods to improve the performance of clustering techniques by utilizing fuzzy logic for reducing the uncertainties in decision making regarding the cluster head selection. Despite the fact that improvement in network lifetime, when these approaches are contrasted with probabilistic clustering techniques, the improvement is not very high because they just partitioned the entire network according to the fuzzy system output which depends upon the residual energy in most of the cases. The probabilistic clustering techniques and fuzzy based clustering techniques have not explored the concept of correlated information. The performance of these existing clustering techniques can be improved further by exploiting the spatial correlation between sensor nodes. Shakya et al. (2013) have presented a two-dimensional analytical model to exploit the spatial correlation between the sensor nodes and further presented the concept of correlated clusters according to the correlation value. By utilizing this 2D analytical model the whole network can be divided into correlated clusters but it does not further present any clustering technique that can utilize this correlation model. Also, this model is two-dimensional but in real scenarios, WSNs are generally deployed in three-dimensional regions. Therefore, to achieve the objective of reducing energy consumption in the network, in this paper, a more realistic three-dimensional analytical model is developed to exploit the spatial correlation between sensor nodes and further a fuzzy based clustering technique is proposed.

This paper is organized as follow. The previous work in this field is reviewed in Sect. 2. The problem is formulated in Sect. 3. The theoretical correlation model is developed in Sect. 4. In Sect. 5, a clustering technique is proposed for both the cases, first without using correlation model and second with using correlation characteristics. The results and analysis are presented in Sect. 6. The paper is concluded in Sect. 7 with future direction.

2 Related work

In this section, firstly the work related to spatial correlation is reviewed and then work related to the fuzzy-based clustering techniques is reviewed. In a dense network, the information detected by sensors is highly correlated in the space domain. This correlation can be exploited to enhance the energy efficiency. Vuran et al. (2004) have presented a theoretical model to estimate the distortion achieved during approximating the information about the event at the sink, when more than one information copies are sent by different nodes from the proximity to the event. In this developed model, the exponential power relation is utilized to approximate the correlation in both the scenarios firstly, between two sensors and secondly between sensor and event. The value of correlation depends upon the distance between two entities. Therefore, according to this correlation model two sensor node correlate to each other even they are out of sensing range to each other.

The same theoretical framework is carry forwarded by Vuran and Akyildiz (2006), where authors have further exploited the spatial correlation at MAC layer and presented the spatial correlation based collaborative MAC (CC-MAC) protocol, which reduces the number of nodes contending for the channel access by allowing only a few nodes to contend for channel access from the event area. Here again, the correlation value between sensor nodes depends upon the distance between them and sensors out of range each other still possess some correlation value which may not be possible in an actual scenario. This work is further extended by Zheng and Tang (2011), where authors presented the circular spatial correlation model for WSNs. The basic difference between previously developed models and in this model is that firstly in the theoretical model, the nodes are placed in a series of concentric circles wherein Vuran et al. (2004) and Vuran and Akyildiz (2006), the nodes deployment was random, secondly the expectation of the source information has some finite value as compared to former models where the expectation of source information is considered as zero. Also, the correlation between the event and sensor is exploited as compared to previous models where the correlation was exploited between sensors as well as between sensors and event. Further, by using correlation model authors

have presented spatial correlation based MAC (SC-MAC) protocol. In SC-MAC, the selection of sensor nodes from the event area relying upon the residual energy and correlation value wherein CC-MAC the selection of sensor nodes only depends upon the correlation values.

To exploit the correlation between sensor nodes according to their sensing range, a theoretical model is presented by Shakya et al. (2013). According to the presented model if two sensor nodes do not have overlapped sensing region then the correlation value between them is zero and if they have overlapped region then can say they correlate to each other up to certain extent. The model presented by Shakya et al. (2013) is more realistic as compared to the previously proposed models. However, this model is developed in 2D planes and according to the actual scenario, 3D model is more close to the reality (Tam and Hai 2016). The most of the literature in the field of spatial correlation mainly focused on exploiting the correlation at MAC layer. The same can be exploited at network layer where correlation model can be clubbed with latest clustering techniques to enhance the network lifetime.

Gupta et al. (2005) have investigated and highlighted the limitations of the LEACH (Heinzelman et al. 2000, 2002) technique and further implemented the fuzzy logic to reduce the uncertainty in selecting the more suitable sensor nodes and have presented novel fuzzy-based clustering technique. This was a centralized clustering technique and has limited applications. Therefore, to reduce these limitations of Gupta et al. (2005), in Kim et al. (2008) authors have presented cluster head election mechanism using fuzzy logic (CHEF), where the decisions regarding cluster head selection and cluster formation are taken at the node level according to the fuzzy system output. The proposed techniques in Pires et al. (2011) and Jin et al. (2011) are inspired from on novel fuzzy-based clustering (Gupta et al. 2005) technique, the only difference exists in the number of input parameters used in the fuzzy logic system. Lee and Cheng (2012) have presented LEACH-ERE to overcome the limitations of the LEACH where the neural network was implemented to estimate the expected residual energy (ERE) consumption. Further, by using the fuzzy logic network is partitioned into clusters. The use of neural network may not be efficient because training data is required and due to the event-oriented behavior of WSN, the training data may not be available. Mhemed et al. (2012) have extended the work of Gupta et al. (2005) approach.

Most of the technique not investigated the consequences of cluster head failure, thus to solve cluster head failure issue the self-configured cluster head election (SCCH) technique has been presented in Izadi et al. (2013, 2015). In SCCH, the network is partitioned into clusters in the first round and after that only, the role of cluster head is rotated among the cluster members rather than forming new clusters in

each round. In SCCH, the cluster heads with a large number of cluster members died earlier as compared to the cluster heads with lower cluster members because the size of the cluster is static after the first round. Thus the uniform distribution of cluster head and cluster members were the issues with this technique. Nayak and Devulapalli (2016) have implemented fuzzy logic two-time to select cluster head and then to select super cluster head from the selected cluster heads. The main feature of this technique is the mobile sink and due to that, the energy consumption of cluster heads is lower as compared to earlier presented techniques. Singh et al. (2016) have implemented fuzzy logic for segregation of the network but did not explore spatial correlation. The 2D correlation model has been presented by Shakya et al. (2013). Sivanandam et al. (2007) have briefly explained the use of fuzzy logic in a sensor network. The possibilities of spatial and temporal correlation have been explored by Akyildiz and Vuran (2010). The distance can calculate using received signal strength (Ju et al. 2010). The swarm optimization algorithm has been implemented by Tam and Hai (2016) for 3D scenarios for lifetime improvement. To tackle the hotspot problem in WSN a fuzzy based energy aware unequal clustering algorithm (EAUCF) has been introduced by Bagci and Yazici (2013). Where the sensor node first calculates the cluster competition radius based on the fuzzy system output. This same work was extended by Logambigai and Kannan (2016) and they introduced fuzzy based unequal clustering (FBUC), where node degree was added in the fuzzy system to find completion radius. In addition to this, cluster joining was based on the fuzzy output value. In Selvi et al. (2016), temporal rules are implemented in fuzzy logic to select the cluster heads. In Selvi et al. (2017a, b), depth-first search (DPS) algorithm has been implemented to find the all possible routes between source and sink, which further efficiently utilized to reduce delay and to improve the network lifetime. In Selvi et al. (2017a, b), fuzzy swarm optimization (FSO) was implemented for dynamic clustering and routing the information and explored the possibilities to improve network lifetime using fuzzy logic and optimization algorithms. Singh and Soni (2017) have presented a detailed review of fuzzy based clustering approaches.

Most of the clustering techniques in literature have not utilized spatial correlation to segregate the network into clusters. The nodes with overlapped sensing region are not taken into account and mostly focused on the residual energy to select the cluster heads. In addition, presented spatial correlation models in literature are modeled for the 2D plane. Therefore, to develop a more realistic model in this paper we proposed 3D spatial correlation model and further on the basis of correlation model fuzzy based clustering technique is presented.

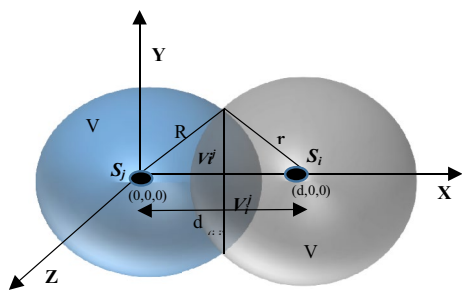


Fig. 2 Two sensors with the overlapped sensing region

Table 1 Notations and symbols used in correlation model

Symbol	Description
R	Sensing range
S_i	Sensing region of node n_i
S_j	Sensing region of node n_j
V	Volume of sensing region
$d_{(i,j)}$	Distance between node n_i and n_j
V_j^i	Volume of region of n_i over n_j
V_i^j	Volume of region of n_j over n_i
$K_c(II,II)$	Correlation function
c	Control parameter
$\rho_{(i,j)}$	Correlation value between n_i and n_j

3 Problem formulation

Let n number of sensor nodes $\{S_1, \dots, S_n\}$ are randomly deployed in the area of interest to detect the event. The all deployed sensors are equipped with same hardware and software. In an Omni-directional sensing, the sensors can detect the events within the sphere of radius r (sensing range) in ideal conditions. In general, extra sensor nodes are deployed to provide full coverage and the event is detected by a large number of sensor nodes in the vicinity of the event. The sink only interested in collective information from the sensors compare to the information from individual nodes (Vuran et al. 2004; Vuran and Akyildiz 2006). If N number of nodes $\{S_1, \dots, S_N\}$ detects the event then due to the overlapped sensing region many of sensor nodes possess a same copy of the information. Thus our objective to reduce the number of copies of the detected information sent to the sink by exploiting the spatial correlation between sensor nodes and reduce the number of nodes participated in clustering process by filtering the redundant nodes. Further, the whole network can be partitioned into the correlated clusters based on the correlation value between sensor nodes and then by selecting only one node with highest residual energy from the each of correlated cluster for clustering process and by keeping the other nodes in sleep mode the number of nodes participating in clustering process reduces automatically (Zheng and

Tang 2011). The number of correlated clusters in the network depends upon the correlation value between them.

4 Proposed spatial correlation model

4.1 Mathematical model design

In this section, a three dimensional (3D) correlation model has been presented with Omni-directional sensing in three-dimensional field flow (e.g. underwater surveillance networks, structural health monitoring of multi-story). Let two sensor nodes S_i and S_j are assumed to be located at $(0, 0, 0)$ and $(d, 0, 0)$ with sensing radius ' R ' and ' r ' respectively as shown in Fig. 2. The notations and symbols are used in the figure are listed in Table 1. If the distance between the i th and j th sensor node is less than the twice of sensing radius ($d_{(i,j)} < 2R$) then S_i and S_j will have an overlapped sensing region and the coefficient of correlation is based on the volume of overlapped sensing region defined as follow.

$$\rho_{i,j} = K_c(d_{i,j}) = \frac{V_i^j + V_j^i}{V} \tag{1}$$

where V is the sensing region of the sensor node. The V_i^j is the overlapped sensing region of an i th sensor node with the j th node sensing region and similarly, V_j^i is the sensing region of a j th sensor node in i th node region. Here $K_c(d_{i,j})$ is the correlation between i th and j th sensor nodes which are a decreasing function of distance $d_{(i,j)}$, its value is equal to one at $d=0$ and of zero for $d \geq 2R$, it means when distance between sensor nodes is zero then the correlation value between them is maximum and if the distance between sensor nodes is greater than twice the value of sensing range the correlation value is minimum.

In Cartesian co-ordinate the equations for sensing region as shown in Fig. 2 are given as follow.

$$x^2 + y^2 + z^2 = R^2 \tag{2}$$

$$(x - d)^2 + y^2 + z^2 = r^2 \tag{3}$$

where S_i and S_j are placed at coordinates $(0, 0, 0)$ and $(d, 0, 0)$ with sensing radius ' R ' and ' r ' respectively as shown in Fig. 2. After solving Eqs. (2) and (3) we get

$$x = \frac{d^2 - r^2 + R^2}{2d} \tag{4}$$

On putting the value of x in Eq. (2), we obtained

$$y^2 + z^2 = \frac{4R^2d^2 - (d^2 - r^2 + R^2)^2}{4d^2} \tag{5}$$

The volume of the three-dimensional lens common to the two spheres can be found by adding the two spherical

caps. The heights of the intersected spherical lens are given as follow.

$$h_1 = \frac{(R - r + d)(R + r - d)}{2d} \tag{6}$$

and

$$h_2 = \frac{(r - R + d)(R + r - d)}{2d} \tag{7}$$

The total overlapped volume between S_i and S_j is given by V as follow

$$V = V(R_1, h_1) + V(R_2, h_2) \tag{8}$$

where ‘ h_1 ’ and ‘ h_2 ’ are the heights of spherical caps. Thus the volume of the spherical cap is given by

$$V(R, h) = \frac{\pi h^2}{3}(3R - h) \tag{9}$$

Using Eqs. (6) and (7) in Eq. (9), we obtained

$$V = \frac{\pi}{12d}(R + r - d)^2(d^2 + 2dr - 3r^2 + 2dR + 6rR - 3R^2) \tag{10}$$

An analytical expression is obtained for intersected volume between two spheres then, ‘ r ’ is substituted by ‘ R ’ abiding by the assumption of the same sensing range of sensor nodes with an omnidirectional antenna. Thus, Eq. (10) further simplified for equal sensing range as follows.

$$V = \frac{\pi}{12}(2R - d)^2(d + 4R) \tag{11}$$

where,

$$V_i^j = V_j^i = \frac{\pi}{12}(2R - d)^2(d + 4R) \tag{12}$$

Hence according to the Eq. (1), the correlation coefficient is given by $\rho_{(i,j)}$ as follow.

$$\rho_{(i,j)} = \frac{1}{16R^3}(2R - d_{(i,j)})^2(d_{(i,j)} + 4R) \tag{13}$$

Let ‘ c ’ be the control parameter, where $c = 2R$.

$$\rho_{(i,j)} = \frac{1}{16R^3}(c - d_{(i,j)})^2(d_{(i,j)} + 2c) \tag{14}$$

According to Eq. (14) it is found that if $d_{(i,j)} = 2R$ the correlation coefficient between two sensor nodes equals to zero and when $d_{(i,j)} < 2R$ the $K_c(d_{i,j})$ possess some positive value. In order to give generalize form to this correlation model a control parameter $c = 2R$ is introduced as a variable to control the coefficient of correlation between sensor nodes. The final correlation model is as follow:

$$\rho_{(i,j)} = K_c(d_{(i,j)}) = \begin{cases} \frac{1}{16R^3}A \times B, & \text{if } : 0 \leq d_{(i,j)} < c \\ 0, & \text{if } : d_{(i,j)} \geq c \end{cases} \tag{15}$$

where $A = (c - d_{(i,j)})^2$ and $B = (d_{(i,j)} + 2c)$. Equation (15) represent the generic model to find out the correlation coefficient between two sensor nodes.

4.2 Formation of correlated clusters

To segregate the WSN into correlated clusters, firstly we find the correlation value $\rho_{(i,j)}$ matrix for every sensor node with respect to the other sensor nodes using Eq. (15). In outset, choose a correlation threshold value between [0, 1] according to the application and set the state of each sensor nodes as a uncluster node by enabling a flag. In first round randomly select one node in the network and set its flag to the clustered node after that check the correlation value from the correlation matrix $\rho_{(i,j)}$ and if the correlation value is greater than the correlation threshold add all the nodes to the cluster $C_k = \{n_i, \dots, n_n\}$ where the $k = \{1, \dots, N\}$. Now repeat this process until the state of each sensor node get changed to the clustered node. The algorithm to form a correlated cluster is presented in Fig. 3.

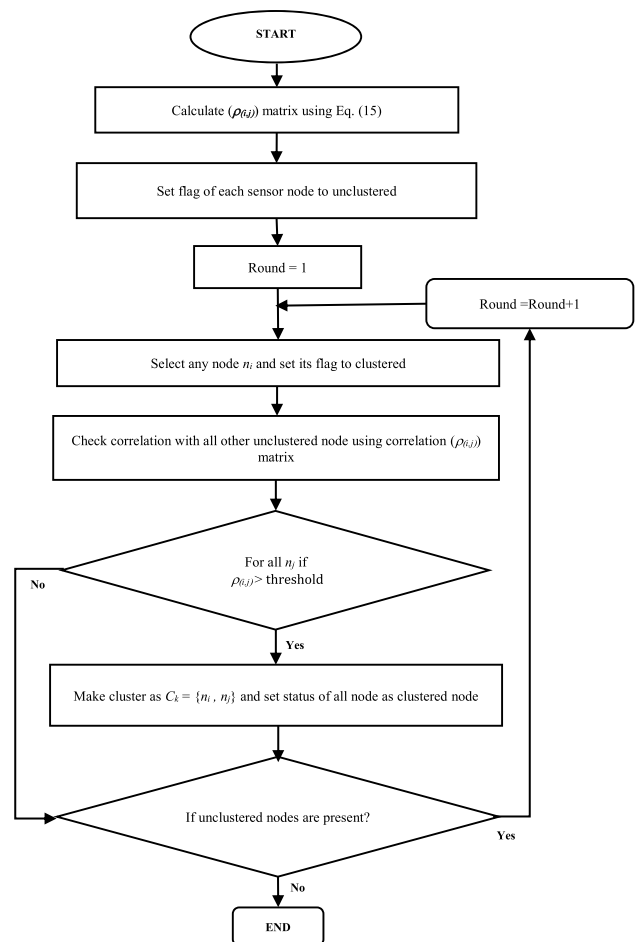


Fig. 3 Flow chart of an algorithm to form correlated clusters

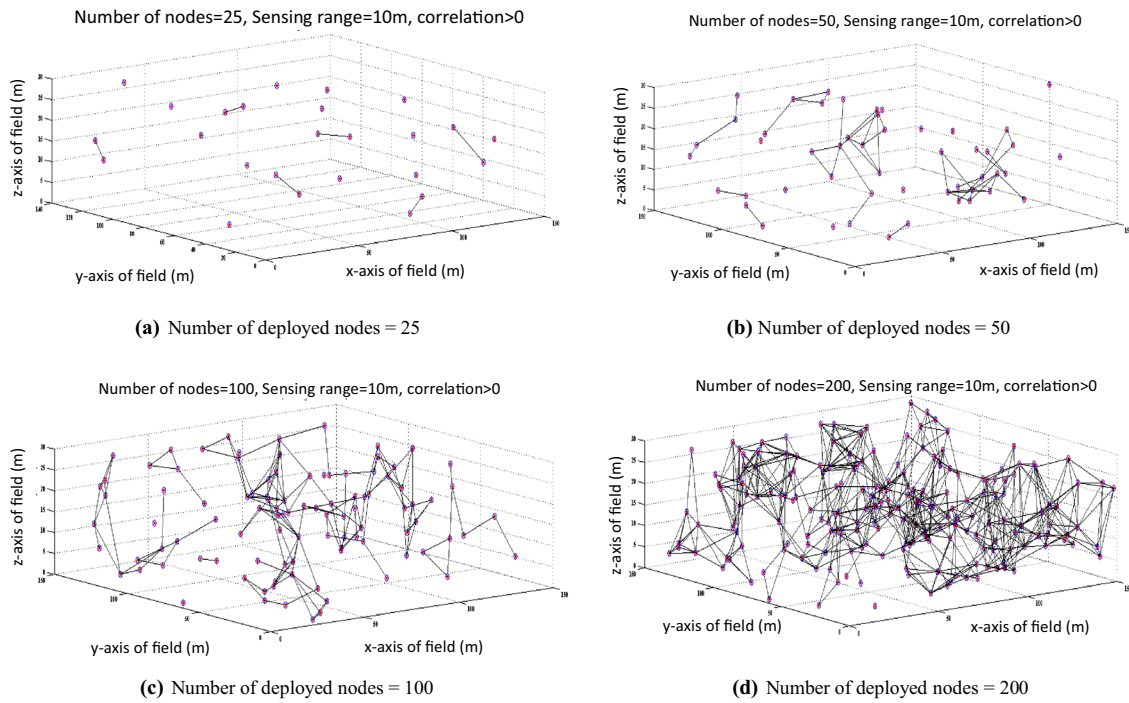


Fig. 4 Random distribution for varying node density, where sensing radius 10 m and correlation threshold is zero. **a** $n=25$, **b** $n=50$, **c** $n=100$, **d** $n=200$

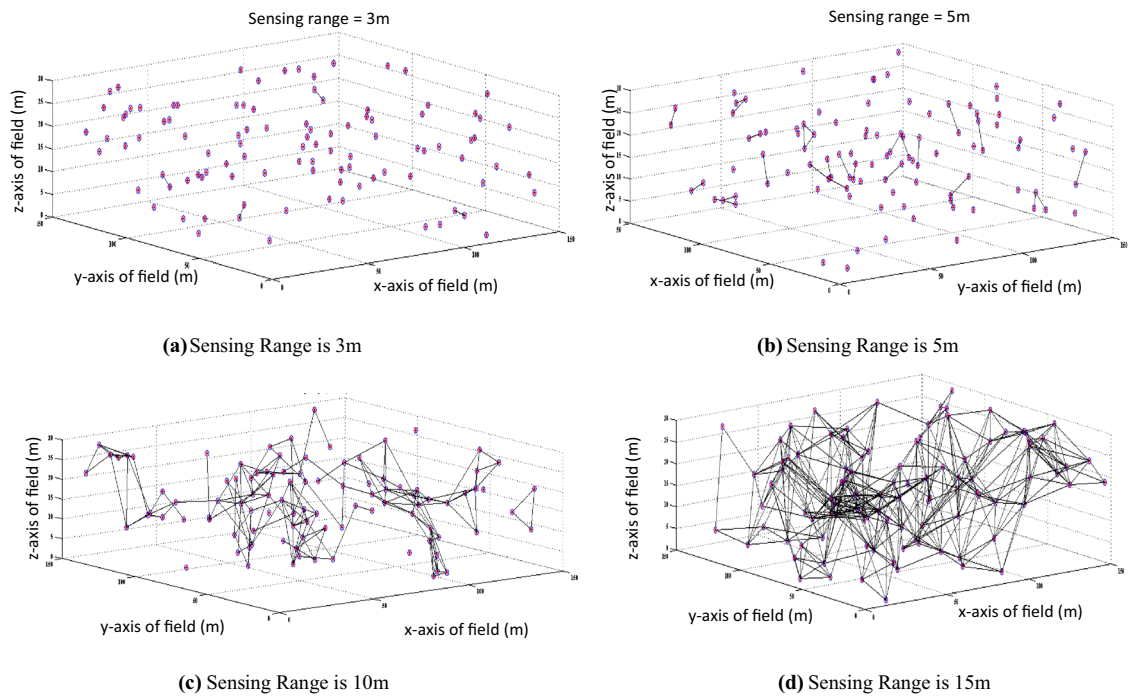


Fig. 5 Random distribution of 100 nodes, where sensing radius (R) is variable and correlation threshold is zero. **a** $R=3$ m, **b** $R=5$ m, **c** $R=10$ m, **d** $R=15$ m

4.3 Discussion

The proposed 3D correlation model and correlated cluster algorithm are implemented in MATLAB. The field of dimensions $150\text{ m} \times 150\text{ m} \times 30\text{ m}$ is created with randomly deployed sensor nodes. The sensor nodes with correlation value greater than the threshold are connected by solid line to represent the correlated clusters. The sensor nodes without connected line are individually considered as correlated clusters having only one node in a cluster. First, we examine the impact of node density on correlation characteristics by varying number of nodes $n=25, 50, 100, 150$ respectively in the field. Figure 4 illustrate that with an increase in a number of nodes the connected lines also get increases. It signifies that higher the node density more closely the sensors in the field and corresponding higher is the correlation between them.

In Fig. 5 as shown the 100 sensor nodes are deployed in the field and the correlation is evaluated between sensors using Eq. (15) with variation in sensing range ($R=3, 5, 10, 15\text{ m}$) to examine the connectivity between sensors. It is inferred from the results that in the same area with the same number of deployed nodes, if sensing range is increased more connected lines appears. This shows that with an increase in sensing range the correlation is increasing.

Algorithm 1 Pseudocode of the proposed clustering scheme.

```

BS_Broadcast_('Hi')
for i = 1:N
    Calculate  $d_{(i,s)}$  using Eq. (16)
    Match their time with the BS
end
for i=1:N
    calculate  $E_{ts(k,d)}$  using Eq. (18)
end
round  $\leftarrow$  1
for i = 1:  $R_{\max}$ 
    for i = 1:N
        fis = readfis('Chance');
        s(i).chance = evalfis([s(i).Er s(i).Ereq]);
        s(i).timer = 1/s(i).chance;
    end
    BS_broadcast_Probe_Signal
    start countdown;
    when (s(i).timer == 0)
        s(i).type == 'cluster head'
        if (s(i) == 0 && receive CH msgs)
            find  $d_{(i,j)}$ ;
            if  $d_{(i,j)} < R_{opt}$ 
                stop countdown;
                send('CH_Join_msg');
            else
                s(i).type = cluster head;
            end
        end
    end
    evaluate energy expenditure;
    round  $\leftarrow$  round+1;
    repeat until alive node > 0;
end

```

In Fig. 6 the correlation is evaluated by varying the correlation threshold value and it is observed that with an increase in threshold value the connected lines between sensors are decreasing and the corresponding number of correlated clusters increasing with decrease in cluster members. After examined the impact of above three parameters, it is observed that a number of correlated clusters and the number of cluster members rely upon the node density, sensing range and the value of correlation threshold value.

The size of the correlated clusters can be adjusted according to the information fidelity requirement using the controlling parameters c and correlation threshold value. It is also inferred from this model that the correlation coefficient is the only function of internode distance and sensing range. Hence, the correlation between any two sensor nodes can only be calculated if sensing range and internode distance are known. Therefore, this model cannot use when sensing range and location of nodes are unknown.

5 Proposed clustering scheme

In this section, firstly we present the presumptions in Sect. 5.1. The objective of the proposed scheme in Sect. 5.2. In Sect. 5.3, cluster formation without using correlation characteristics is presented. The spatial correlation based clustering technique is presented in Sect. 5.4.

5.1 Presumption

The WSNs comprised of homogeneous sensors and location of each sensor nodes is fixed. The base station is static and located outside the field. The sensor nodes are capable of varying its transmission range but not sensing range. It is assumed that geographical locations of nodes are known. The worst condition is considered where each sensor node in the field has data to send in each round. Only cluster heads are authorized to transmit collaborative information to the base station directly or indirectly.

5.2 Objective

The main objective of the proposed scheme is to enhance network lifetime by taking the advantage of spatial correlation characteristics. To enhance network lifetime, we uniformly distribute the cluster heads and reduced the number of nodes sending data to the base station by exploiting spatial characteristics. This is achieved by reducing the number of candidates participated in cluster head process by keeping some nodes in sleep mode which have high correlation value with neighboring nodes. Firstly, we propose the generic clustering technique in Sect. 5.3, where cluster heads are selected on the basis of the output value of the fuzzy

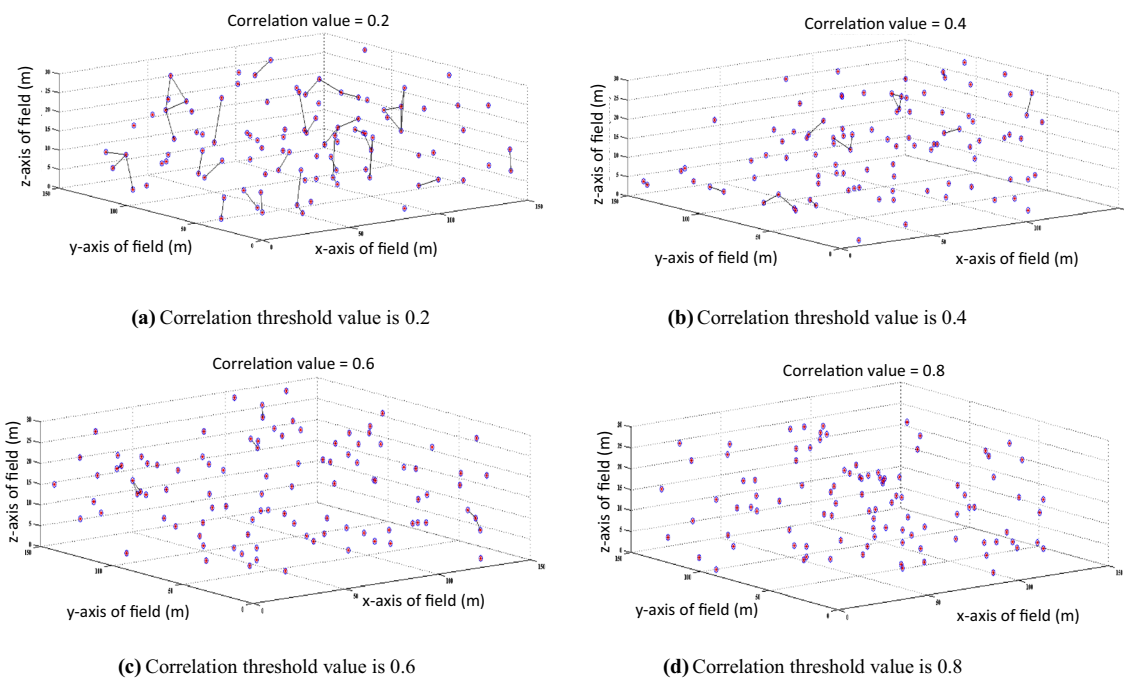


Fig. 6 Random distribution of 100 nodes, where correlation threshold is variable and sensing range is 10 m. **a** Correlation threshold=0.2, **b** correlation threshold=0.4, **c** correlation threshold=0.6, **d** correlation threshold=0.8

system. The detailed information of proposed technique can be seen from Algorithm 1 and from the flowchart in Fig. 7. In Sect. 5.4, we embedded the proposed correlation model with proposed clustering technique to reduce the number of nodes participated in the clustering process.

5.3 Proposed clustering technique without utilizing correlation characteristics

In this subsection, a clustering scheme using fuzzy logic is proposed where the decision regarding cluster head selection and cluster formation is taken at the node level. Here clusters are configured in two phases like LEACH (Heinzelman et al. 2002), CHEF (Kim et al. 2008) and FBUC (Logambigai and Kannan 2016). The first phase of the proposed technique is different from the LEACH, CHEF, and FBUC. The fuzzy logic with two input variables: residual energy and expected required energy is used to find the chance of sensor node to become a cluster head. In the first phase base station broadcast Hi message with timing information, each sensor node calculates required energy to transmit one bit to the base station according to the received signal strength (Xu et al. 2010) and set their clock based on the received timing information. The distance is calculated by sensor node on the basis of received signal power is as follow.

$$d = 10^{[(P_0 - F_m - P_r - 10 \times L \times \log_{10}(f)) + 30 \times L - 32.44] / 10 \times L} \quad (16)$$

where P_0 is the power (dBm) at d_0 distance. The d_0 is depended (m) upon amplification radio factor for free space and multipath respectively as shown in Eq. (17).

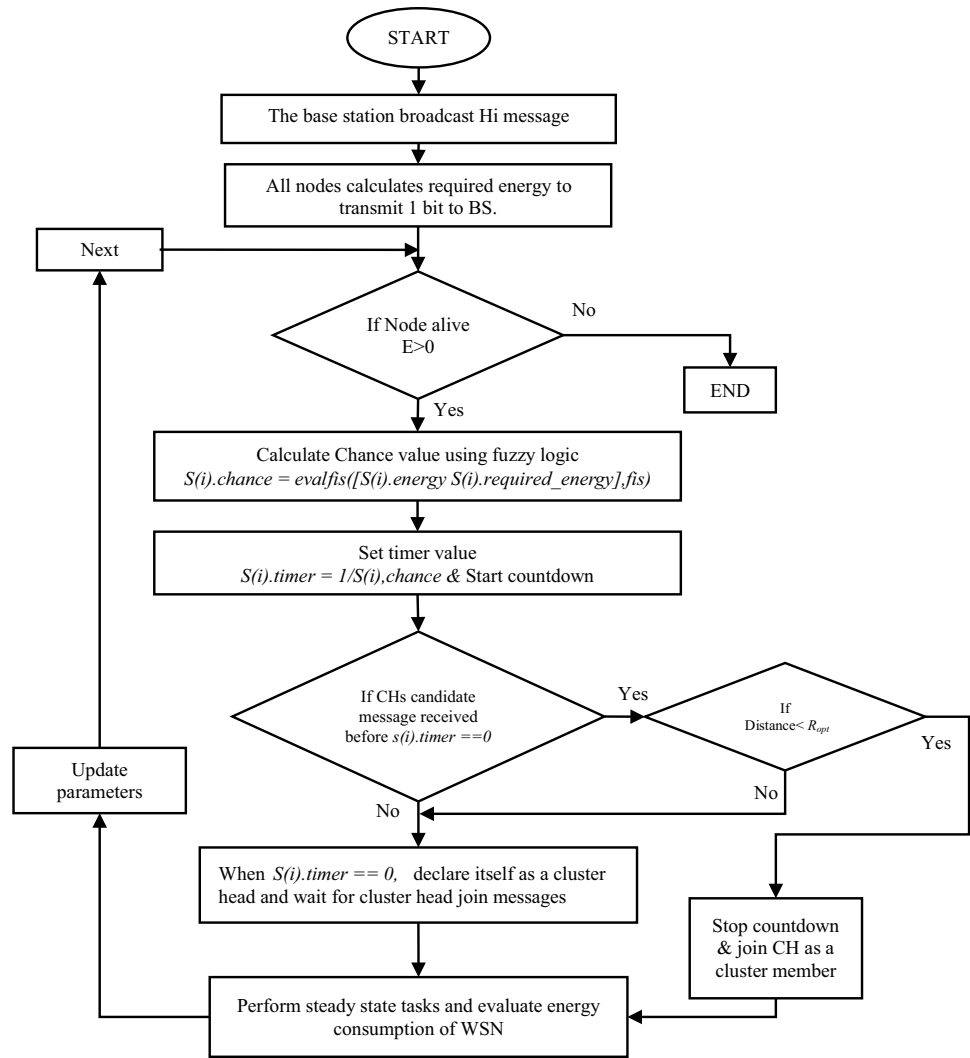
$$d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \quad (17)$$

Now sensor node finds out the energy required to transmit according to the radio energy model (Heinzelman et al. 2002) as given by

$$E_{Tx}(k, d) = \begin{cases} (E_{elec}^{Tx} + \epsilon_{fs} \times d^2) \times k, & d < d_0 \\ (E_{elec}^{Tx} + \epsilon_{fs} \times d^4) \times k, & d \geq d_0 \end{cases} \quad (18)$$

Symbols and notations meanings used in Eqs. (16), (17) and (18) are given in Table 3. After evaluating the required input parameters each sensor node runs fuzzy system shown in Fig. 8 where each input is divided into three triangular membership functions as shown respectively. The fuzzy if-then rules are given in Table 2. Based on rules fuzzy system gives the crisp value that is considered as *chance* value to become a cluster head. As shown in Algorithm 1 and Fig. 7, each node set the timer value which is inversely proportional to the *chance* value. Therefore, for higher chance values, the timer value is lower and vice versa. According to the fuzzy rules sensor nodes with higher residual energy and low required energy have higher chance value, thus more eligible node have lower timer value as compared to the less eligible nodes. After

Fig. 7 Flow chart of proposed clustering technique



this, the base station will broadcast probe signal in the network which contains the timing synchronization information and on receiving this probe message each sensor node will activate its timer and start its countdown. When the countdown reaches zero it declares itself as a cluster head candidate by

broadcasting cluster head candidate message with a chance value within the optimal cluster radius. The optimal cluster radius (R_{opt}) is evaluated as follow.

$$R_{opt} = \sqrt[3]{\frac{volume}{4/3 \times \pi \times C_{opt}}} \tag{19}$$

Table 2 Fuzzy if-then rules for proposed technique

Rule No.	Required Energy	Residual energy	Chance
1	Low	Low	Low
2	Low	Med	High
3	Low	High	Very high
4	Med	Low	Very low
5	Med	Med	Med
6	Med	High	Very high
7	High	Low	Very low
8	High	Med	Low
9	High	High	Med

Notations meaning is given in Table 3. The sensor node after receiving cluster head candidate message find out the distance using RSSI and if it is less than R_{opt} sensor stop countdown and send cluster head join message. In another scenario, when sensor receives more than one cluster head candidate message then in this situation it compares the chance value of candidates and sends the cluster head join message to the candidate having highest chance value within that region. This process it continues and the whole network gets divided into clusters.

In the second phase like existing techniques, LEACH and CHEF cluster heads design channel access schedule and

Table 3 Notations and symbols

Notations/symbols	Meaning
$S(.)$	Sensor node
fis	Fuzzy inference system
E	Residual energy
R_{opt}	Optimal cluster radius
w	Constant for timer adjustment
d	Actual distance between node and sink
d_0	Threshold distance
P_0	Power received at d_0
L	Path loss exponent
f	Frequency
ϵ_{fs}	Free space amplification factor
ϵ_{mp}	Multipath amplification factor
k	Number of bits in a packet
$E_{Tx}(k,d)$	Required energy to transmit k bits up to distance d
E_{elec}	Energy consumed by electrical circuit
$volume$	3D field volume
C_{opt}	Optimal number of cluster heads
B_0	Number of cluster members
E_{CH}	Energy expended by cluster head per round
E_{nonCH}	Energy expended by cluster members per round
n	Total number of nodes deployed in field
$E_{initial}$	Initial energy of sensor
E_{total}	Total energy consumed in one round
E_{avg}	Energy left after performing round
SM	Number of nodes in sleep mode per round

perform its duties. The energy consumption of cluster head is as follow (Heinzelman et al. 2002).

$$E_{CH} = kE_{elec}B_0 + kE_{DA}(B_0 + 1) + E_{Tx}(k, d_{toBS}) \tag{20}$$

The meaning of symbols is listed in Table 3. The energy consumption of non-cluster sensor nodes is evaluated as follow

$$E_{nonCH} = kE_{elec} + k\epsilon_{fs}d_{toCh}^2 \tag{21}$$

The meaning of symbols is listed in Table 3. The total energy consumption in each round and average energy consumption are approximated as follow.

$$E_{total} = C_{opt} \times E_{CH} + (n - C_{opt}) \times E_{nonCH} \tag{22}$$

and

$$E_{avg} = E_{initial} - E_{total} \tag{23}$$

The meaning of symbols is listed in Table 3. If two sensor nodes countdown reaches at zero at the same time then to break the tie, the nodes will generate random number and node with a higher value of random number will become a cluster head for the current round. In this technique all deployed sensor nodes take part in the clustering process. In

order to reduce number of participating nodes, the proposed clustering technique presented in this section is clubbed with correlation model and further presented in the next section.

5.4 Proposed clustering technique using correlation characteristics

Here we present the extended version of proposed clustering technique in presented Sect. 5.3. The main objective of this proposed technique is to reduce the number of reporting nodes using the correlation characteristics between sensor nodes. As shown in Fig. 7, the whole network can be segregated into correlated clusters.

When the correlation threshold is high for example 0.9 which implies that approximately 90% of sensing region overlap to each other. Now if any event occurs nearby these correlated sensors then the sensed information is highly correlated to each other or can assume have same information if the quality of service is not a top priority. Then in such cases, if keep only one node in an active state for each round and others in sleep mode periodically then the energy consumption gets reduced due to less number of reporting nodes.

The most of the energy consumed by transceiver system (Heinzelman et al. 2002), if the number of reporting nodes reduced the load on cluster head and a number of bits communicated also get reduced which directly reduce the energy consumption. The flow chart of the proposed scheme is shown in Fig. 9. The performance of this proposed scheme depends upon the correlation threshold value. Thus according to the required fidelity constraints of the application, the WSNs lifetime can be adjusted using this technique. The energy consumption in each round is approximated as follow.

$$E_{total} = C_{opt} \times E_{CH} + (n - C_{opt} - SM) \times E_{nonCH} \tag{24}$$

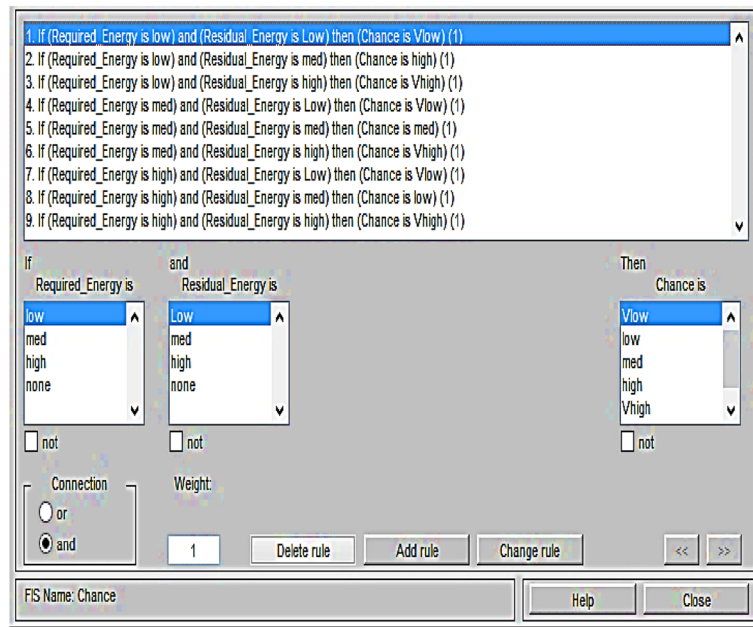
Notations are given in Table 3. A number of nodes in sleep mode depends upon correlation threshold value.

It is clearly seen from Eq. (24) that the approximated energy consumption is less as compared to the Eq. (22). Therefore, the overall energy consumption can reduce by using proposed correlation-based clustering technique. In next section, the simulation results are presented to analyze the performance.

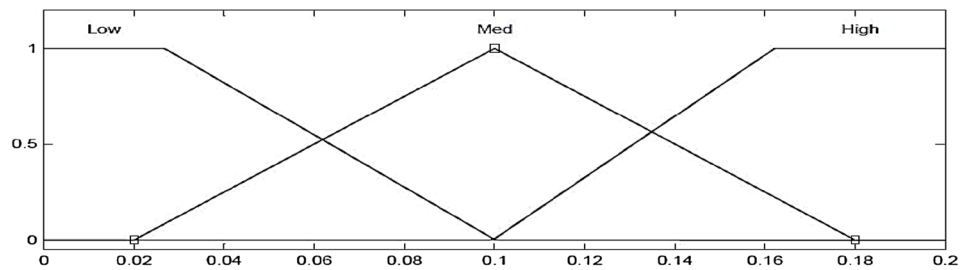
6 Experimental setup and analysis

We have implemented the proposed techniques in MATLAB and compared their performances with LEACH, CHEF and FBUC. The three-dimensional region is considered with field dimensions $150\ m \times 150\ m \times 30\ m$. The number of

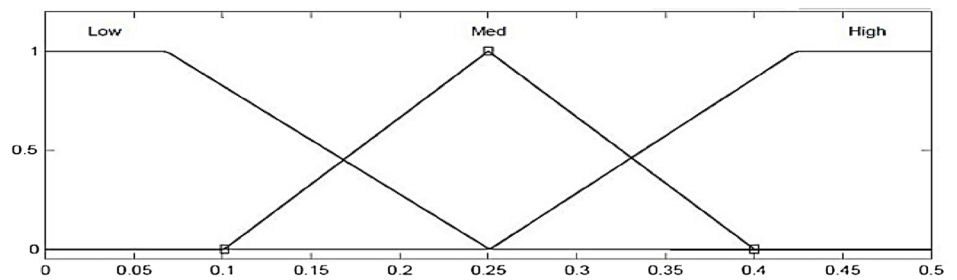
Fig. 8 Fuzzy system with membership functions for inputs and output



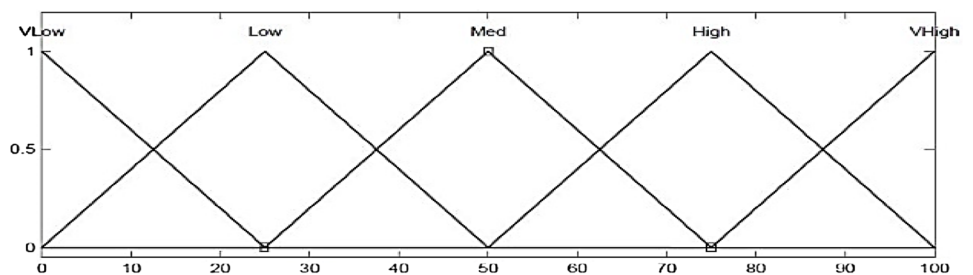
(a) Fuzzy rules implemented in MatLab.



(b) Membership functions for 'Required Energy'.



(c) Membership functions for 'Residual Energy'.



(d) Membership functions for output 'Chance'.

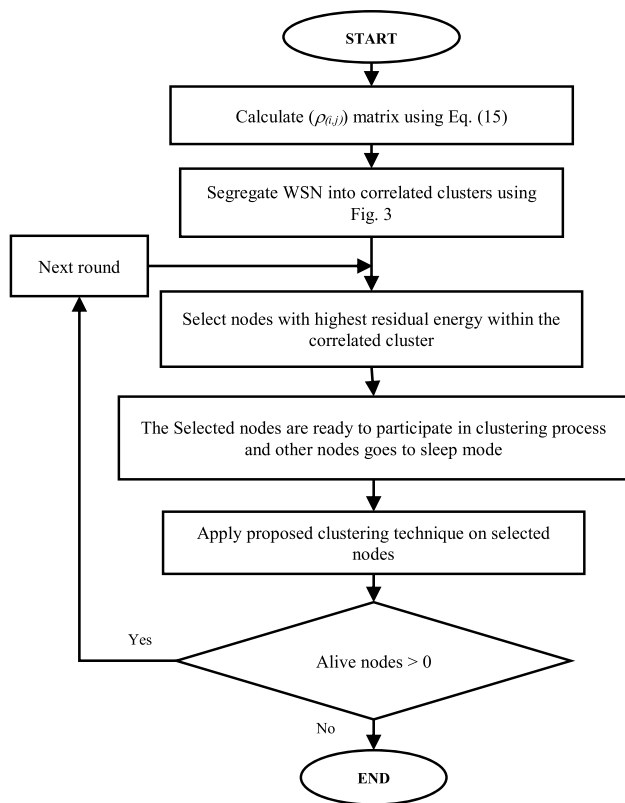


Fig. 9 Flow chart of proposed correlation based technique

Table 4 Simulation parameters

Parameters	Value
Sink	(50,175,0)
Initial energy (E_0)	0.5J
ϵ_{fs}	10×10^{-12}
ϵ_{mp}	0.0013×10^{-12}
k	4000 bits
n	50, 100, 150, 200
R_{opt}	$R_{opt} = \sqrt[3]{\frac{volume}{4/3 \times \pi \times C_{opt}}}$
E_{elec}	50nJ
E_{DA}	5nJ
Fuzzy rules	9
Header	200 bytes
Node distribution	Random
C_{opt}	Dynamic
Packet	500 bytes

nodes deployed in the field is varied $\{n = 50, 100, 150, 200\}$ respectively in order to evaluate the performance in different node densities.

Firstly, the performance of proposed clustering technique without using correlation characteristic is examined. Where we examine the performance on the basis of the number of alive nodes, the first node dies (FND) round, the last node dies (LND) round and a number of cluster head w.r.t number of rounds. Secondly, the performance of proposed clustering technique using correlation characteristics is examined where in order to perform unbiased comparison LEACH, CHEF, and FBUC are modified using correlation model where on the basis of correlation value number of nodes participated in the clustering process. The simulation parameters are listed in Table 4.

6.1 Results and discussion

The comparison for a number of alive nodes in every round is depicted in Fig. 10. It can be clearly seen from the results that the number of alive nodes decreases with increase in the number of rounds. This decrement occurs due to the battery depletion of the sensor nodes. The deployed sensor network can sustain for a longer duration when most of the sensor nodes stay alive for a maximum number of rounds.

In the proposed scheme, the number of alive nodes is higher as compared to other three techniques. In the proposed scheme, on average around 25, 20, 15% more number of alive nodes as compared to the LEACH, CHEF, and FBUC respectively. However, after 1000th round, this percentage is decreasing. But according to Logambigai and Kannan (2016), sensor network can perform satisfactory up to the half of the node dies in the network. Based on this analysis, it can say that proposed scheme performed better than other three approaches.

It is also observed from the comparisons that with an increase in node density the node death is more uniform with respect to increasing in a number of rounds. The first node dies round (FND) and the last node dies round (LND) comparisons are illustrated in Fig. 11. For $n = 50, 100, 150,$ and 200 the first node dies nearby 280th, 400th, 500th, and 590th round. Based on the FND the proposed scheme improved network lifetime by 78%, 35%, and 30% as compared to LEACH, CHEF, and FBUC respectively.

Generally, FND define the network lifetime (Heinzelman et al. 2002; Kim et al. 2008; Selvi et al. 2016) because the node affects the overall performance. In LND comparison, the proposed scheme shows better results as compared to other three techniques. Therefore, on basis of this analysis, it can say that proposed scheme perform better and capable to extend the network lifetime. Figure 12 illustrates the distribution of the alive nodes with respect to the number of rounds for proposed clustering technique using 3D correlation model.

As shown, with an increase in the correlation threshold value the network lifetime is decreasing because higher the

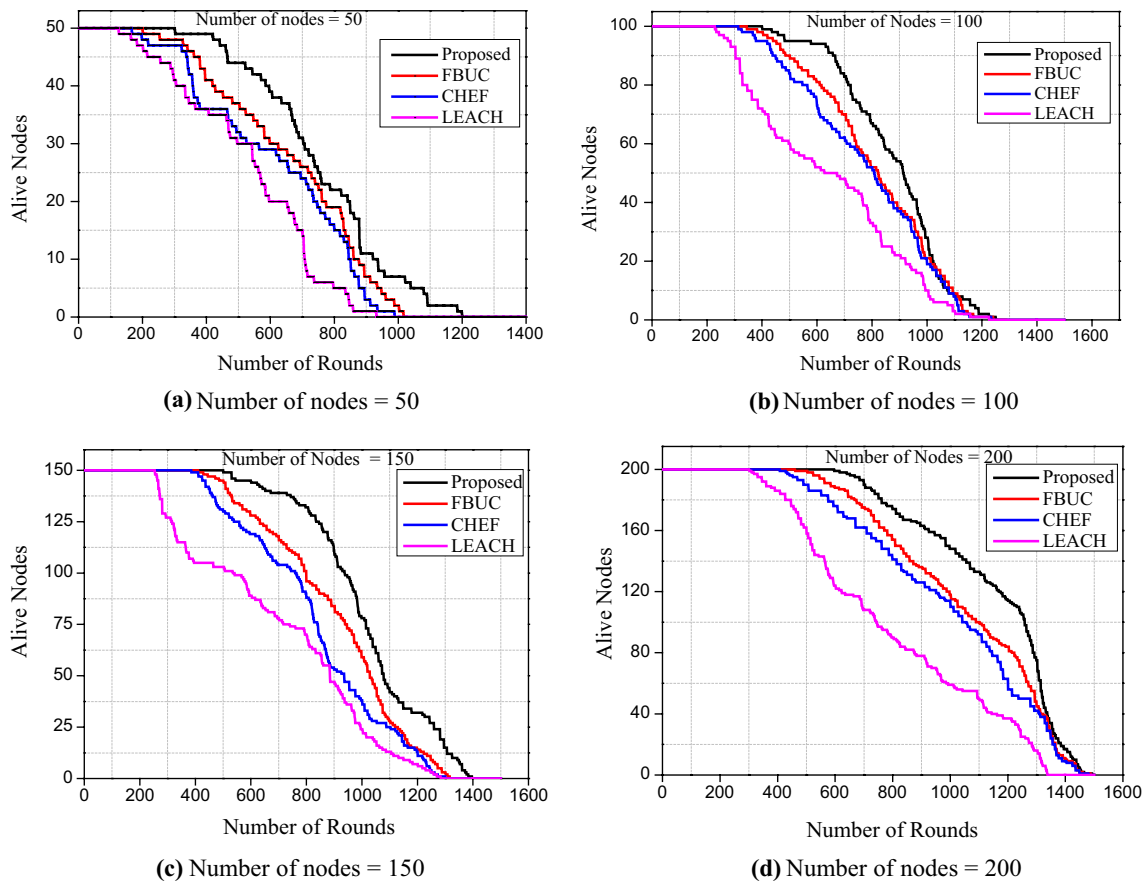
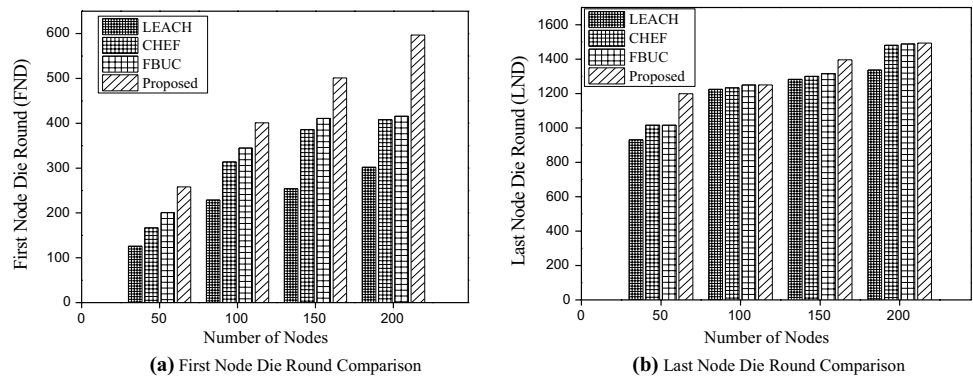


Fig. 10 Comparison of number of nodes (n) alive w.r.t number of rounds with variation in number nodes deployed in the field **a** $n=50$, **b** $n=100$, **c** $n=150$, **d** $n=200$ respectively

Fig. 11 Comparison of FND and LND rounds respectively



correlation threshold value higher is the number of reporting nodes. For example, when the threshold value is 0.2, it means that the nodes having overlapped sensing region greater than or equal to 20%, considered as correlated clusters. In this situation, keep only one node in active state and others in sleep state from each cluster, automatically reduce the number of reporting nodes. In another scenario, when the threshold value is 0.9 the nodes having 90% overlap between sensing region are considered as correlated. The

probability of 20% overlapped sensing region is higher as compared to the 90% overlap. Therefore, a more number of reporting nodes for higher correlation threshold value and lesser the lifetime as compared to the lower threshold value. From Fig. 12, it can be seen that for lower correlation threshold values first node death occur earlier as compared to the higher threshold value. But overall percentage of a number of alive nodes is higher for lower threshold value.

Fig. 12 Network lifetime comparison using correlation characteristics in proposed clustering technique w.r.t change in correlation threshold value

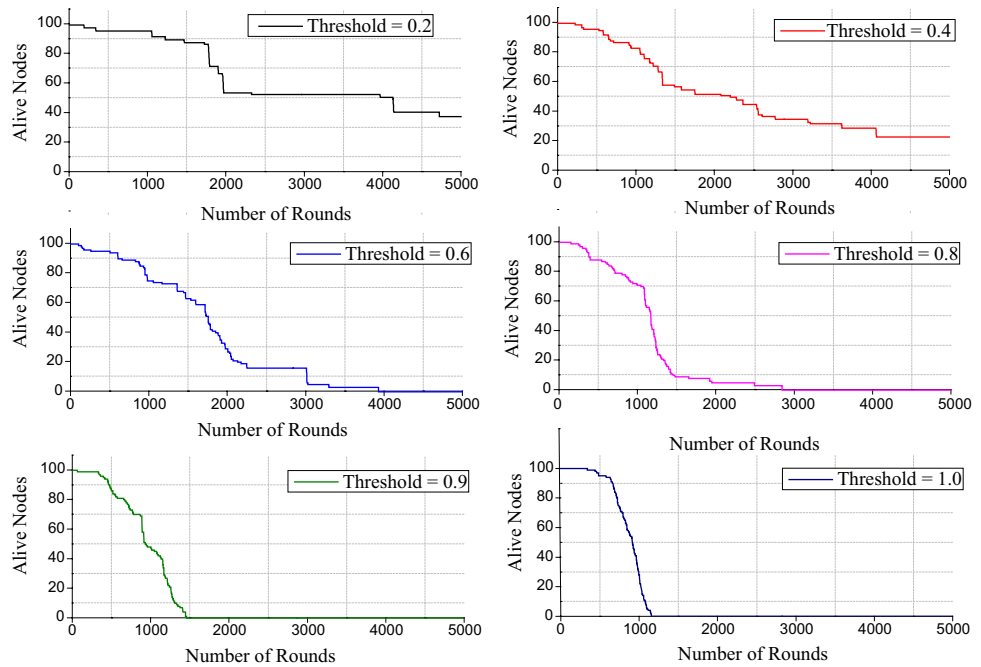


Fig. 13 Comparison of number of alive nodes vs number of rounds using correlation characteristics. **a** Correlation threshold = 0.6, **b** Correlation threshold = 0.8

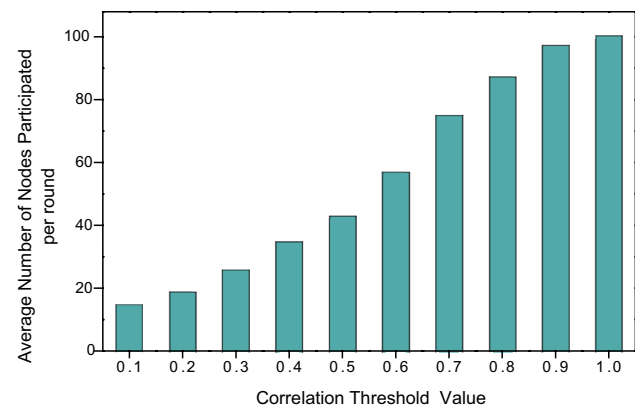
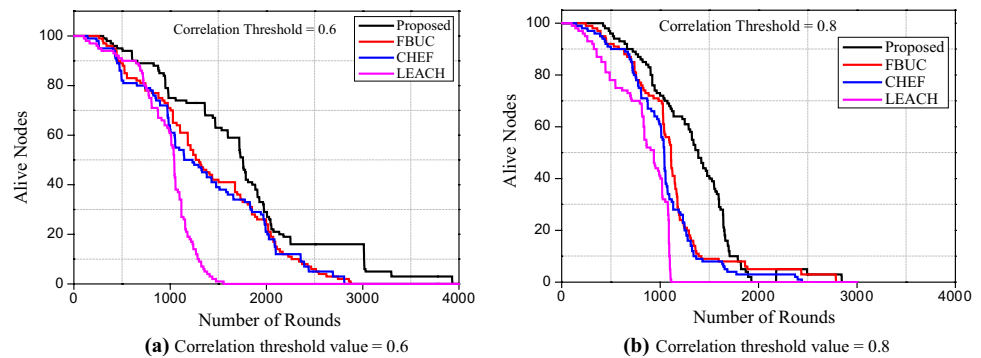


Fig. 14 Number of nodes participated in clustering process vs correlation threshold value

Therefore based on the required application correlation threshold value can be adjusted to prolong the network operational lifetime. On the basis of half node dies round it can say that for lower correlation threshold values, proposed clustering approach perform better as compared to the higher values. The comparisons between proposed, CHEF, LEACH, and FBUC using correlation model are depicted in Fig. 13. It is inferred from results that with an increase in correlation value the network lifetime is reducing. It is observed that when the threshold value is decreased from 0.8 to 0.6, there is 34, 9, 7, and 5% improvement in the network lifetime based on the FND round. The reason for the improvement in network lifetime is a reduction in the number of participating nodes in

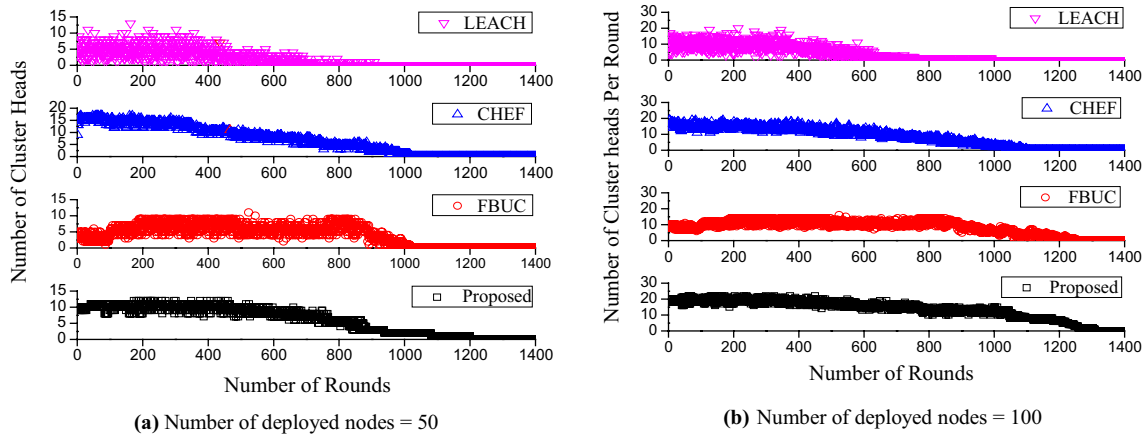


Fig. 15 Number of cluster heads selected per round for Proposed, CHEF and LEACH techniques with a number of nodes (n) deployed in the field. **a** $n=50$, **b** $n=100$

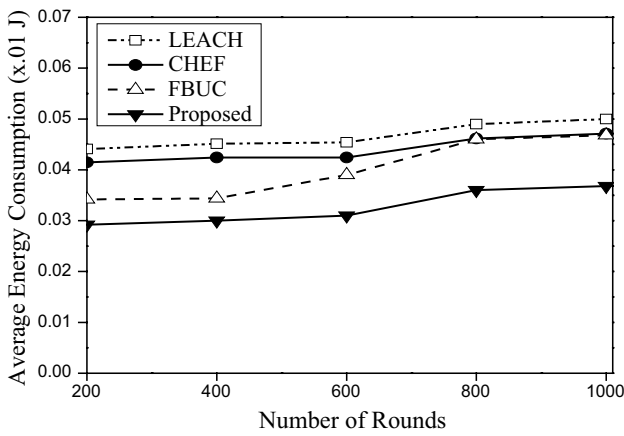


Fig. 16 Average energy consumption per round

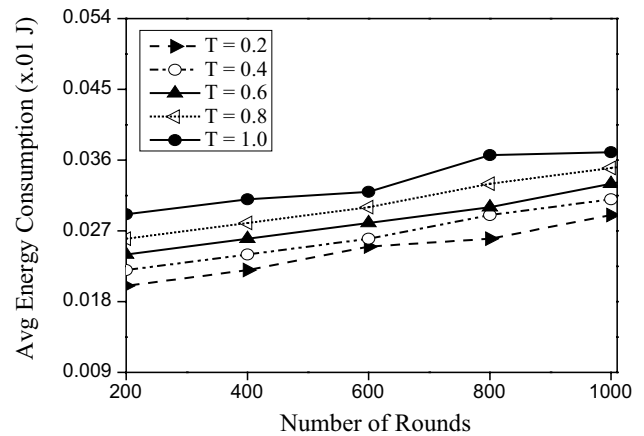


Fig. 17 Effect of correlation threshold (T) on average energy consumption

clustering process in each round by varying the correlation threshold value.

The average number of nodes participated in clustering process for different correlation threshold value are shown in Fig. 14. It is observed that with an increase in correlation value the number of nodes participating in clustering process also get increases. Therefore, based on this analysis it can say that the network lifetime of any clustering technique can be further improved by using the proposed correlation model as shown in Fig. 13. The cluster heads distribution with respect to the number of rounds for proposed, CHEF, LEACH and FBUC techniques without using correlation characteristics are illustrated in Fig. 15. The results show that the number of cluster heads selected in each round for proposed, CHEF, and FBUC are somewhat constant as compared to the LEACH or can say variation is less as compared to the LEACH.

In proposed and CHEF, the cluster radius is fixed. Therefore, due to the cluster radius restriction, the variation in the selection of cluster heads in each round is very less for proposed and CHEF. However, in FBUC, there is an increase in a number of cluster heads with an increase in the number of rounds because of unequal cluster size in the network. Based on these results, it cannot say that proposed system is better than CHEF but it can say that the proposed approach performed well like CHEF. Also, the idea of cluster radius was initially used in CHEF and we carry forward this concept with slight modification.

The comparison for average energy consumption is depicted in Fig. 16. After 200 round it can be seen that for LEACH, CHEF, and FBUC the average energy consumption is 45, 30, and 10% higher than the proposed approach. Also, result reveals that the variation in average energy consumption is lower as compared to other three approaches. The

reason behind this lower energy consumption is the uniform distribution of clusters and the selection of most suitable cluster heads due to fuzzy logic chance value. Therefore, based on average energy consumption it can say that proposed scheme perform better than other three approaches.

Figure 17 illustrates the effect of correlation threshold value (T) on average energy consumption per round in the network for first 1200 rounds. It is observed that energy consumption is increasing with increase in the value of T. This is because a number of participating nodes depend upon the threshold value and as the threshold value increasing, participating nodes also increasing. As shown, when the value of T varies from 0.2 to 1.0, average energy consumption is increasing by 55% and when varying from 0.2 to 0.8 it is increasing by 40%. These results show that the average energy consumption depends upon the correlation threshold value. It is the reason behind the increase in the network lifetime with a change in the value of T. The network lifetime depends upon the average energy consumption in the network. The network can sustain for longer time duration if average energy consumption is low. Therefore, it can say that network lifetime can be extending by reducing the average energy consumption by optimizing the correlation threshold value (T). The above results and analysis reveal that the 3D correlation model presented in Sect. 4 can be utilized in a more effective way to extend the network lifetime. It is also observed from the results that by using correlation model the network lifetime is extended for proposed, CHEF, LEACH and FBUC techniques.

7 Conclusion

This paper presents a 3D spatial correlation model for exploiting the correlation between sensors based on their geographical location and sensing range. In order to implement correlation model a new fuzzy logic based clustering technique is proposed. The performance of correlation model and clustering technique is validated through MATLAB based simulation by creating randomly deployed sensor network scenarios with varying node density, sensing range and correlation threshold value. The effects of node density and sensing range on spatial correlation characteristics are studied and results are presented. It is inferred from the results that by exploiting the correlation between sensors the network lifetime can be extended by reducing the number of reporting nodes. The proposed clustering technique is fuzzy based where a decision regarding the cluster head selection and cluster formation are taken at the node level. The results show that the proposed clustering technique is more energy efficient as compared to the FBUC, CHEF and LEACH techniques. Further, the performance of proposed, FBUC, CHEF

and LEACH clustering techniques are improved by embedding them with proposed correlation model. The correlation model performance depends upon the choice of correlation threshold value. The effect of varying threshold value on network lifetime in term of a number of alive nodes with respect to a number of rounds are studied and simulation results are presented in this paper. The results show that by using correlation model the performance of clustering techniques can also be further improved. As a future work, the correlation model can be implemented in designing of MAC layer protocols for reducing contention for channel access.

References

- Akyildiz IF, Vuran MC (2010) *Wireless sensor networks*, vol 4. Wiley, Hoboken. <https://doi.org/10.1002/9780470515181>
- Bagci H, Yazici A (2013) An energy aware fuzzy approach to unequal clustering in wireless sensor networks. *Appl Soft Comput* 13(4):1741–1749. <https://doi.org/10.1016/j.asoc.2012.12.029>
- Gupta I, Riordan D, Sampalli S (2005) Cluster-head election using fuzzy logic for wireless sensor networks. In: 3rd annual conference on communication networks and services research, May 16–18, 2005, Halifax, NS, Canada, pp 255–260. <https://doi.org/10.1109/CNSR.2005.27>
- Heinzelman W, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: 33rd annual Hawaii international conference on system sciences, Jan 4–7, 2000, Maui, Hawaii, USA, pp 1–10. <https://doi.org/10.1109/HICSS.2000.926982>
- Heinzelman W, Chandrakasan A, Balakrishnan H (2002) An application-specific protocol architecture for wireless microsensor networks. *IEEE Trans Wirel Commun* 1(4):660–670. <https://doi.org/10.1109/twc.2002.804190>
- Izadi D, Abawajy J, Ghanavati S (2013) A new energy efficient cluster-head and backup selection scheme in WSN. In: IEEE 14th international conference on Information reuse and integration (IRI), Aug 14–16, 2013, San Francisco, CA, USA, pp 408–415. <https://doi.org/10.1109/IRI.2013.6642500>
- Izadi D, Abawajy J, Ghanavati S (2015) An alternative clustering scheme in WSN. *IEEE Sens J* 15:4148–4155. <https://doi.org/10.1109/JSEN.2015.2411598>
- Jin R, Wei N, Shi X, Gao T, Zou J (2011) Clustering routing protocol based on fuzzy inference for WSNs. In: 7th international conference on wireless communications, networking and mobile computing (WiCOM), Sept 23–25, 2011, Wuhan, China, pp 1–4. <https://doi.org/10.1109/wicom.2011.6040346>
- Kim JM, Park SH, Han YJ, Chung TM (2008) CHEF: cluster head election mechanism using fuzzy logic in wireless sensor networks. In: 10th international conference on advanced communication technology (ICACT), Feb 17–20, 2008, Gangwon-Do, South Korea, pp 654–659. <https://doi.org/10.1109/ICACT.2008.4493846>
- Lee JS, Cheng WL (2012) A fuzzy-logic-based clustering approach for wireless sensor networks using energy predication. *IEEE Sens J* 12:2891–2897. <https://doi.org/10.1109/JSEN.2012.2204737>
- Logambigai R, Kannan A (2016) Fuzzy logic based unequal clustering for wireless sensor networks. *Wirel Netw* 22(3):945–957. <https://doi.org/10.1007/s11276-015-1013-1>
- Mhemed R, Aslam N, Phillips W, Comeau F (2012) An energy efficient fuzzy logic cluster formation protocol in wireless sensor networks. *Procedia Comput Sci* 10:255–262. <https://doi.org/10.1016/j.procs.2012.06.035>

- Nayak P, Devulapalli A (2016) A fuzzy logic-based clustering algorithm for wsn to extend the network lifetime. *IEEE Sens J* 16:137–144. <https://doi.org/10.1109/JSEN.2015.2472970>
- Pires A, Silva C, Cerqueira E, Monteiro D, Viegas R (2011) CHEATS: a cluster-head election algorithm for WSN using a Takagi-Sugeno fuzzy system. In: *IEEE Latin-American conference on communications (LATINCOM)*, Oct 24–26, 2011, Belem do Para, Brazil, pp 1–6. <https://doi.org/10.1109/LatinCOM.2011.6107427>
- Selvi M, Logambigai R, Ganapathy S, Ramesh LS, Nehemiah HK, Arputharaj K (2016) Fuzzy temporal approach for energy efficient routing in WSN. In: *Proceedings of the international conference on informatics and analytics*, August 25–26, 2016, Pondicherry, India, pp 1–5. <https://doi.org/10.1145/2980258.2982109>
- Selvi M, Logambigai R, Ganapathy S, Nehemiah HK, Arputharaj K (2017a) An intelligent agent and FSO based efficient routing algorithm for wireless sensor network. In: *Second international conference on recent trends and challenges in computational models (ICRTCCM)*, Feb 3–4, 2017, Tindivanam, India, pp 100–105. <https://doi.org/10.1109/ICRTCCM.2017.43>
- Selvi M, Velvizhy P, Ganapathy S, Nehemiah HK, Kannan A (2017b) A rule based delay constrained energy efficient routing technique for wireless sensor networks. *Cluster Comput*. <https://doi.org/10.1007/s10586-017-1191-y>
- Shakya RK, Singh YN, Verma NK (2013) Generic correlation model for wireless sensor network applications. *IET Wirel Sens Syst* 3:266–276. <https://doi.org/10.1049/iet-wss.2012.0094>
- Singh M, Soni SK (2017) A comprehensive review of fuzzy-based clustering techniques in wireless sensor networks. *Sens Rev* 37(3):289–304. <https://doi.org/10.1108/SR-11-2016-0254>
- Singh M, Soni S, Kumar V (2016) Clustering using fuzzy logic in wireless sensor networks. In: *3rd international conference on computing for sustainable global development (INDIACom)*, March 16–18, 2016, New Delhi, India, pp 1669–1674
- Sivanandam SN, Sumathi S, Deepa SN (2007) *Introduction to fuzzy logic using MATLAB*. Springer, Berlin. <https://doi.org/10.1007/978-3-540-35781-0>
- Tam NT, Hai DT (2016) Improving lifetime and network connections of 3D wireless sensor networks based on fuzzy clustering and particle swarm optimization. *Wirel Netw*. <https://doi.org/10.1007/s11276-016-1412-y>
- Vuran M, Akyildiz I (2006) Spatial correlation-based collaborative medium access control in wireless sensor networks. *IEEE/ACM Trans Netw* 14(2):316–329. <https://doi.org/10.1109/TNET.2006.872544>
- Vuran M, Akan Ö, Akyildiz I (2004) Spatio-temporal correlation: theory and applications for wireless sensor networks. *Comput Netw* 45(3):245–259. <https://doi.org/10.1016/j.comnet.2004.03.007>
- Xu J, Liu W, Lang F, Zhang Y, Wang C (2010) Distance measurement model based on RSSI in WSN. *Wirel Sens Netw* 2(8):606–611. <https://doi.org/10.4236/wsn.2010.28072>
- Yuan HY, Yang SQ, Yi YQ (2011) An Energy-efficient unequal clustering method for wireless sensor networks. In: *International conference on computer and management (CAMAN)*, May 19–21, 2011, Wuhan, China, pp 1–4. <https://doi.org/10.1109/CAMAN.2011.5778810>
- Zheng G, Tang S (2011) Spatial correlation-based MAC protocol for event-driven wireless sensor networks. *J Netw* 6(1):121–128. <https://doi.org/10.4304/jnw.6.1.121-128>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Manjeet Singh¹  · Surender Kumar Soni¹

✉ Manjeet Singh
manjeet.nith@gmail.com

Surender Kumar Soni
surender.soni@gmail.com

¹ Department of Electronics and Communication Engineering, National Institute of Technology, Hamirpur, Himachal Pradesh, India