

DDCD Algorithm based Energy Efficient Clustering for WSNs

Y. Chalapathi Rao^{1*} and Ch. Santhi Rani²

¹Department of ECE, Anurag Engineering College, Kodad - 508206, Telangana, India; chalu.8421@gmail.com

²Department of ECE, DMSSVH College of Engineering, Machilipatnam - 521002, Andhra Pradesh, India; santhirani.ece@gmail.com

Abstract

Objectives: In Wireless Sensor Networks (WSNs) the nodes are moving in different geographical conditions and the nodes may be left looked after for a long period of time. So to increase the network life time and to reduce the traffic the energy should be conserved. Energy efficient WSNs will lose for a long period time. **Methods/Analysis:** To increase network lifetime different methods like data aggregation, cluster head, network coding, correlating data set, etc can be used for correlated data environment. These methods are used to calculate to increase the network lifetime. DDCCD (Data Density Correlation Degree) algorithm on wireless sensor network is works as a middleware for aggregating data sustained by a more number of nodes within a network. **Findings:** The problem encountered in the recent past was of the more battery power consumption. Therefore, this paper proposed the efficient and effective mechanism of energy efficient procedures for data aggregation in wireless sensor network and increase the network lifetime. **Application/Improvement:** The results of the simulation are acceptable showing the DDCCD algorithm to have good performance abilities compared to the Weighted-Low Energy Adaptive Clustering Hierarchy and Ant Colony Optimization algorithms. This process is useful where the sensor nodes are densely established.

Keywords: Clustering Head, Data Aggregation, DDCCD, Wireless Sensor Networks (WSNs)

1. Introduction

A Wireless Sensor Network will track the physical environment such as weather conditions, health monitoring, forest fire detection, etc and transmits the data into base station¹. The main aim of WSN will sense the territory conditions and collect the data. The collected data are transmitted through the network to the base station (sink) and prevent the detection². In sensor network there are several number of sensor nodes are presented. Every sensor node is connected to one or more nodes. Every node has number of parts such as radio transceiver with an internal connection to external circuits for interfacing with the sensors and an energy source³. The technologies of wireless sensor network are used in military applications and currently these are used in so many industrial such as control, health care monitoring applications and so on.

There are different set of nodes each nodes collect the data and send directly to the base station so the congestion occur on the network⁴. To avoid that the cluster head is created between each set of nodes. Now the node sends the data to the Cluster Head (CH) and the CH sends the data to the base station^{5,6}. In order to save the energy and to increase the network lifetime the nodes should be in off state⁷. So saving minimum energy dissipation and maximizing the network lifetime. Only the transmission and communication will be in off state, the nodes will be sensing at all time. If the node receives any data while sensing, the transmission and communication part will be on state based on Weighted Low Energy Adaptive Clustering Hierarchy (W-LEACH).

In this paper, we are presented a Data Density Correlation Degree (DDCCD) to estimate the spatial correlation between a sensor's data and neighbouring sensor's

*Author for correspondence

data⁸. Each node receives the data and aggregate the data by using clustering method and the CH send the data to the base station^{9,10}. CH is selected based on which node has the high capacity of transmission and communication part. Thus the CH sends the data to the base station (sink).

In the current work, spatial correlation between a signals spatial direction and average received signal gain. There are different types of nodes. Each node carries different works such as temperature, pressure, military surveillance etc. In WSN the coverage is defined as an estimate of how long and how well for the sensors are able to sense and discover the entire space. Each set of nodes have the CH because if each node send the data directly to the base station the congestion occurs on the network. So we form cluster head by the technique using DDCCD clustering method. If we create cluster head the congestion will be less on the wireless sensor network. CH is selected based on which node has the high capacity of the coverage area and high capacity of transmission and communication between the sensor nodes¹¹.

In present system they are not concentrating in energy because all nodes are active at all time so the wireless sensor network lifetime is decreased. It is the main drawback in existing system, to solve this problem W-LEACH to minimize the energy consumption and to maximize the wireless sensor network lifetime. One of the nodes will be assume as a W-LEACH based on which node has the low capacity of sensing, transmission, communication and coverage area in the WSN¹². Thus the W-LEACH will allot a time slot for each node, which node should be in ON state and which node should be in OFF state. Off state means only the transmission and communication part goes to OFF state. Sensing the environment mode will be on at all time. While sensing if the node receives any data the transmission and communication part will be ON state based on the W-LEACH.

The energy consumption of DDCCD is similar to W-LEACH, the energy of nodes is used to amalgamation, transmit and receive the data. The node transmit a l-bit data to its receiver over d distance, the energy consumption is composed of the transmitting circuit loss and power amplifier loss. The model as

$$E_{Tx}(L,d) = \begin{cases} L.E_{elec} + L.\epsilon_{fs}.d^2; & \text{if } d < d_0 \\ L.E_{elec} + L.\epsilon_{mp}.d^4; & \text{if } d \geq d_0 \end{cases} \quad (1)$$

For receiving this data, narrate the required energy is

$$E_{Rx}(L,d) = L.E_{elec} \quad (2)$$

Where E_{elec} is the consumed energy per bit, d_0 is the threshold, ϵ_{mp} is the multipath fading mode, ϵ_{fs} is the free space mode, and d is the distance¹³. From Equation (1), by equating the two expressions at $d = d_0$, we get

$$d_0 = \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \quad (3)$$

Data fusion is consume certain energy, the energy consumption of aggregating the L-bit data is:

$$E_{DF}(L) = L.E_{DA} \quad (4)$$

Where, E_{DA} is the energy dissipation of data aggregating per bit.

Ant Colony Optimization (ACO) is a random technique that exhibits the social behaviour of ant colonies, which organize to build the shortest path to feeding sources. Ants forage for food lay down quantities of pheromone marking the route that they follow. ACO tap the search experiences and domain grasping to accelerate the search process. ACO is performed by a team of intelligent agents which shows the ants behaviour, walking along the graph to solve the issue by using procedures of cooperation and adaptation.

ACO techniques have been successful applied to many areas like scientific and industries. This routing based algorithm is used for improving the energy efficiency in unicasting¹⁴, transmitting and data gathering. The routing issues in WSNs has focused the ACO algorithms, for extending the lifetime of sensor network by detection the number of connected covers are maximum.

2. Materials and Methods

This paper addresses the cluster based routing methods: Sensor nodes (V) are grouped to form a set of mutually exclusive clusters. The Cluster Head (CH) receive the data from other clusters randomly. The received data is aggregated by the CH and directly transmits to the sink. The topology can be modelled as an undirected graph $G(V, E)$ with V as the group of vertices and an edge (u, v) is included in the edges set E if two nodes u and v can share with each other. This topology is a widely accepted solution for organizing sensor nodes into clusters for energy

conservation. The gathered observations from the nodes can be suppressed by the CHs that are almost correlated in the densely deployed network. The higher energy nodes act as CHs, after processing the received data it is sent to sink directly. The low energy nodes represent the cluster members which perform the sensing of physical phenomena.

2.1 Spatial Data Correlation Degree (SDCD)

A wireless sensor network composing of a set of N sensor nodes, in this paper, we model the spatial correlation between surveillances of each node. The data gathered by N number of sensors in an event area. In practice the Gaussian random distribution is encountered¹⁵. Thus, to steady the data at nodes by using a *jointly Gaussian model*, with an N -dimensional multivariate Gaussian distribution $F_N(\mu, C_Y)$: The N -dimensional Gaussian density is

$$f(Y) = \frac{1}{\sqrt{2\pi}|C_Y|^{1/2}} \exp\left\{\frac{1}{2}(Y - \mu)^T C_Y^{-1}(Y - \mu)\right\} \quad (5)$$

where μ is the mean, C_Y is the co-variance matrix (+ve), C_Y^{-1} is the inverse matrix. The diagonal elements of C_Y are the variances $C_{ii} = \sigma^2$ and $C_{ij} = \sigma^2 \exp(-\alpha d_{i,j}^2)$. Then combination of index $I = \{i_1, \dots, i_k\} \in \{1, \dots, N\}$, $k \leq N$, $Y = (Y_{i_1}, \dots, Y_{i_k})$ is k -dimensional Gaussian distributed. Its co-variance matrix is $C[I]$ selected from C_Y , The entropy of a k -dimensional multivariate Gaussian distribution $F_k(\mu, C_Y)$ is:

$$h(F_k(\mu, C_Y)) = \frac{1}{2} \log(2\pi e)^k |C_Y| \quad (6)$$

From Equation (6), compute the SDCD of a group, which consists of a node and its close members. The clustering techniques can segregate the CHs and combine the nodes into separate clusters. Two different correlated random data sources Y_i and Y_j , the entropies of Y_i and Y_j are $P(Y_i)$ and $P(Y_j)$, then Y_i and Y_j both can communicate, jointly code data coding rate $P(Y_i Y_j)$. It is essential that their conditional entropies $P(Y_i/Y_j)$ and $P(Y_j/Y_i)$ are equal to their coding rates $P(Y_i Y_j)$.

It can be extended to multi-dimensional conditions. For n -correlated random data sources $Y = (Y_0, Y_1, Y_2, \dots, Y_n)$, then the sources are jointly coding rate $P(Y_0, Y_1, Y_2, \dots, Y_n)$ to code do not communicate their data. Let the set Y of sources are mapped with respect to the distances to Y_0 , and then assigned the coding rates (W) are as follows:

$$\begin{aligned} W_0^* &= P(Y_0) \\ W_1^* &= P(Y_1 / Y_0) \\ W_2^* &= P(Y_2 / Y_1, Y_0) \\ &\vdots \\ W_n^* &= P(Y_n / Y_{n-1}, Y_{n-2}, \dots, Y_1, Y_0) \end{aligned} \quad (7)$$

Where Y_0 is a cluster head and (Y_1, Y_2, \dots, Y_n) are the cluster members.

2.2. Problem Definition

After build out the clusters, the total cost of communication sustained throughout the collecting data of two clusters, $E_{total} = E_{clusters} + E_{to\ sink}$, where $E_{clusters}$ is the reduced cost for sending the data to the CH by each cluster member, and $E_{to\ sink}$ is the cost acquired by the CH at the time of sending data to the sink. For grouping the small number of clusters, it not only processes the higher data rate compression, and the cost of communication process $E_{to\ sink}$ is lower.

Suppose a group of sensor nodes are selected as a set of m -cluster heads, $C = \{c_1, c_2, \dots, c_m\}$. The following conditions are holds to select the cluster set C :

- Reduce the total cost of communication $E_{total} = E_{clusters} + E_{to\ sink}$.
- For every node $u \in V$, either $u \in CH$ or $u \in V - CH$ holds. For every node $u \in V - CH$, exists a $v \in CH$ and $(u, v) \in E$.
- The inspired sub graph is connected to CH .
- The sink is a part of the CH set which is a node.
- The consumed energy of a CH is below a threshold.

To reduce the data gathering energy cost, we need to reduce $E_{clusters}$ and $E_{to\ sink}$. To meet the above purpose, the average cost of energy for aggregating the data by one node in one period is acquired. A set of CHs denoted by $C = \{c_0, c_1, c_2, \dots, c_m\}$ with c_0 distributes as the CH. Assume the energy dissipated by node c_i is $T(c_i)$ for sending the data process in one period by c_i to c_0 . $T(c_i)$ can be evaluated by

$$T(c_i) = 2E_{elec} + W_i^* \{d(c_i, c_0)^\gamma \cdot E_t + E_r\} \quad (8)$$

where E_{elec} is the energy cost, W_i^* is the coding rate, $d(c_i, c_0)$ is the distance between c_i and c_0 , γ is the path

loss exponent, E_t is the transmitting energy and E_r is the dissipated energy by c_0 for receiving a unit data.

Let $P(c_0)$ is the cost of energy by c_0 for transmitting a data to the BS, and then the energy dissipated by cluster C for gathering the data process is:

$$Cost = \sum_{i=1}^m T(c_i) + \frac{\theta}{\lambda} \sum_{i=1}^m W_i^* . P(c_0) \quad (9)$$

Where θ is the data gathering frequency, λ is the data size, $\sum_{i=1}^m T(c_i)$ is the $E_{clusters}$, and $\frac{\theta}{\lambda} \sum_{i=1}^m W_i^* . P(c_0)$ is the E_{tosink} .

The average cost of energy in one period is for gathering the data by one node:

$$Cost = \frac{1}{m} \sum_{i=1}^m T(c_i) + \frac{\theta}{m\lambda} \sum_{i=1}^m W_i^* . P(c_0) \quad (10)$$

To reduce E_{tosink} , $\sum_{i=1}^m W_i^*$ is also need to reduce and maximize the DDCCD of the cluster. Evaluate the $P(c_0)$ as given below:

$$P(c_0) = \frac{d(c_0, BS)}{d^*} \left(E_{elec} + \lambda (d^*)^\gamma . E_t \right) + \left(\frac{d(c_0, BS)}{d^*} - 1 \right) (E_{elec} + E_r) \quad (11)$$

where d^* is the optimum hop between c_0 and BS. Assume x is the average hop from c_0 to BS, and then the energy dissipated by c_0 for transmitting data to the BS is

$$f(x) = \frac{d(c_0, BS)}{x} (E_{elec} + \lambda x^\gamma E_t) + \left(\frac{d(c_0, BS)}{x} - 1 \right) (E_{elec} + \lambda E_r) \quad (12)$$

Differentiating Equation (12),

$$f'(x) = -\frac{d(c_0, BS)}{x^2} (2E_{elec} + E_r) + \frac{d(c_0, BS) . \lambda x^{(\gamma-2)} . E_t}{\gamma - 1} \quad (13)$$

According to the condition of extremism, Equation (13) is equal to 0, then $x=d^*$ can be computed.

$$d^* = \sqrt[\gamma]{\frac{(\gamma - 1)(2E_{elec} + \lambda E_r)}{\lambda E_t}} \quad (14)$$

This clustering problem can be modelled as a weighted connected controlling. A controlling graph set $G(V, E)$

is a subset of $S \subseteq V$, and for every sensor node $u \in V$ is either in adjacent to a node of S or set S . A node of S is control the all adjacent vertices. G is called as weighted graph, if the different weights of the nodes are allocated in the graph G . This clustering issue can be modelled for evaluating a controlling set of weights. The total weight of the CHs summation is maximum¹⁶.

3. Clustering Method Description

In wireless sensor network the clustering method is done by DDCCD clustering method¹⁷. The clustering method is to aggregate the data which receives from the sensor node and then it send to the BS. By choosing the cluster head the traffic will be reduced in the WSN.

There is n neighbouring sensor nodes $V_i (i = 1, \dots, n)$ in the communication radius of V is the core sensor node, so the definition of DDCCD for V as follows:

$$Sim(V) = \begin{cases} 0, & N < \min Pts \\ a_1 (1 - 1 / \exp(N - \min Pts)) + a_2 (1 - d\Delta / \varepsilon) \\ + a_3 (1 - d / \varepsilon), & N \geq \min Pts \end{cases} \quad (15)$$

Where ε is the threshold data, $\min Pts$ is the threshold amount, d is the average distance between D and N , $d\Delta$ is the distance between the data centre of the data objects and D , ε is the neighbourhood of D , a_1, a_2, a_3 are weights, D is the data object of V and $a_1 + a_2 + a_3 = 1$. The sensor node V of DDCCD is $Sim(V)$ defined by Equation (15).

The cluster is selected among the set of nodes in the WSN Figure 1. Therefore the CH is selected based on which node has the high capacity of the coverage area, transmission and communication between the set of nodes. The DDCCD algorithm has two techniques:

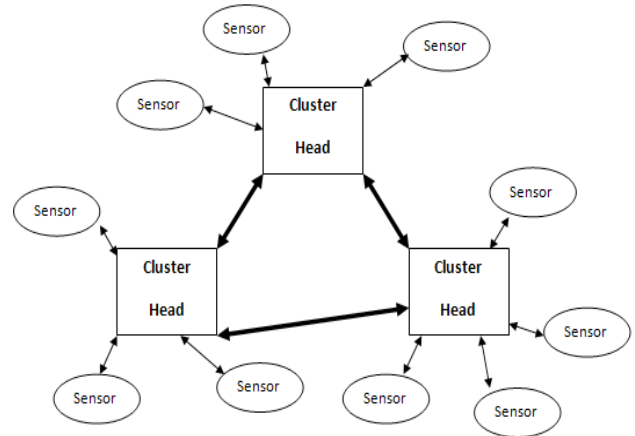


Figure 1. Cluster formation.

- Sensor type Calculation Techniques (STC)
- Local Cluster Construction techniques (LCC).
The following algorithm is used for clustering and data aggregation.

3.1 STC Techniques

Input :

Threshold Data – α ;

Threshold Amount – min Pts;

Weights – a_1, a_2, a_3 ;

Set of neighbour sensor nodes – $N(i)$.

Output :

Types of sensor nodes – core or non – core;

Set of sensor nodes – $SNSetin(i)$ and $SNSetout(i)$;

Data Density Correlation Degree – $Sim(i)$

Step1:

for each $i: i \in V$ *par do*

{

a non – core sensor node is I ,

Near Num = 0;

$SNSetin(i) = I, SNSetout(i) = I$

for each $j: j \in N(i)$

{

if (*Near Num* > min Pt)

{

a core sensor node is I ;

$Sim(i) = x_1(1 - 1 / \exp(n - \text{min Pt})) + x_2(1 - d / \alpha)$

$+ x_3(1 - d / \alpha)$;

else

$Sim(i) = 0$;

}

Step2:

for each $i: i \in V$ *par do*

{

if (*a core sensor node is* I)

{

for each $j: j \in N(i)$

{

if ($\|d(i) - d(j)\| \leq \alpha$)

$SNSetin(i) = \{j\} \text{ union } SNSetin(i)$;

else

$SNSetout(i) = \{j\} \text{ union } SNSetout(i)$;

}

}

}

3.2 LCC Techniques

Input :

Types of sensor nodes – core or non – core;

Set of sensor nodes – $SNSetin(i)$ and $SNSetout(i)$;

Data Density Correlation Degree – $Sim(i)$

Output :

$ClusterSet = \{ClusterSet(i) | i \in V\}$

$DDCD \text{ Set} = \{DDCD_{\max}(i) | i \in V\}$

Step1:

for each $i: i \in V$ *par do*

{

If

{

for each $j: j \in SNSetin(i)$

{

sends the data from I *sensor node* – $(I, -1, Sim(i))$ *to* j

}

or each $j: j \in SNSetout(i)$

{

sends the data from I *sensor node* – $(I, -1, Sim(i))$ *to* j

}

if (*non – core sensor node is* I)

{

or each $j: j \in N(i)$

{

sends the data from I *sensor node* – $(I, -1, Sim(i))$ *to* j

}

}

}

Step2:

for each $i: i \in V$ *par do*

{

$DDCD_{\max}(i) = (I, Sim(i))$

for each $j: j \in SNSetin(i)$

{

if $Sim(j) > Sim(i)$

$DDCD_{\max}(i) = (j, Sim(j))$,

}

if (*a core sensor node is* I)

$ClusterSet(i) = SNSetin(i)$;

if $ReceiveDataRelationship == 1$

$ClusetrSet(i) = \{ReceiveData.ID\} \text{ union } ClusterSet(i)$;

}

4. Results and Discussion

In this paper the DDCCD algorithm exhibits a significant improvement in the performance metrics like Energy

consumption, Throughput, Nodes alive, Delivery Ratio, etc.

4.1 Network Energy Consumption

Energy consumption is crucial important for network lifetime because sensor nodes operated in an irreplaceable battery. When number rounds increase in the network energy dissipation of the nodes should decrease so that we need efficient transmission to avoid nodes to get quickly drain out of their energy. Figure 2 represents the network energy consumption greatly achieved when compared to W-LEACH and ACO strategy.

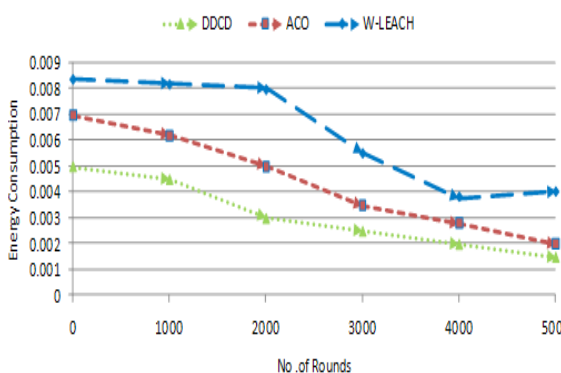


Figure 2. Number of rounds versus energy consumption.

4.2 Network Throughput

It is usually computed in bps, i.e., bits per second. The successful data packets delivered over a communication channel to the receiver is defined as throughput. Figure 3 represents the throughput is better for multi sink moving when compared with W-LEACH and ACO.

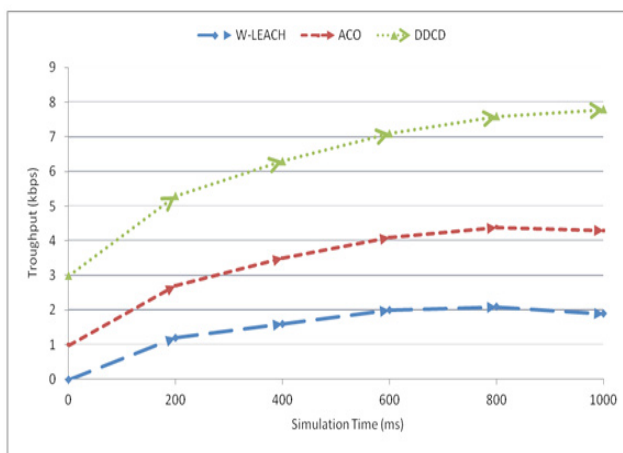


Figure 3. Simulation time versus throughput.

4.3 Nodes Alive

In order to compute the number of alive nodes in each round for maximize the network lifetime. The round denotes the each transmission may need to be rebuilt the each sensor node which generates DATA_PAK that has to be sent to the base station. Then we collect the data about the death rates of rounds versus nodes. Figure 4 represents that DDCCD method contains maximum number of alive nodes when compared to W-LEACH and ACO algorithms.

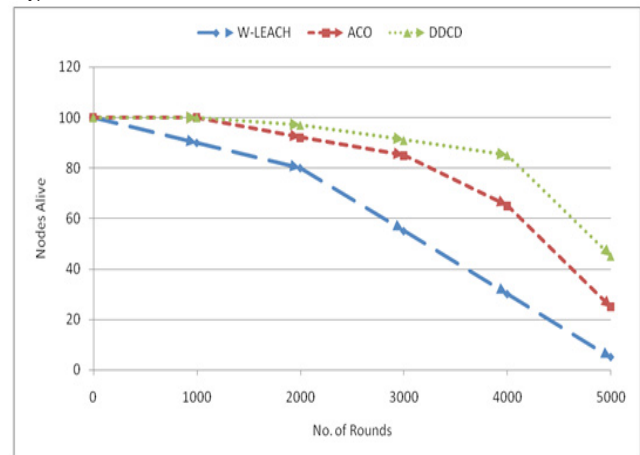


Figure 4. Number of rounds versus nodes alive.

4.4. Packet Delivery Ratio

The percentage of data packets gathered by the nodes that are successfully delivered is packet delivery ratio. Figure 5 represents the packet delivery ratio when compared with density of node, the proposed scheme can achieve better. When the density of node is 50 the packet delivery ratio of DDCCD algorithm has 95% but W-LEACH method has 66% and ACO method has 85%, further increase the density of nodes DDCCD algorithm has still 90% but W-LEACH and ACO algorithms are sharp decrease below 80%.

5. Conclusion

The main aim of this paper is the introduction of Data Density Correlation Degree (DDCD) clustering approach and energy efficient manner in WSN. This paper highlights the experiment of reducing the traffic between the nodes in WSN and clustering and aggregating the data in efficient manner by using STC and LCC techniques. Simulation experiment results show that DDCCD significantly reduces the amount of data transmission, especially

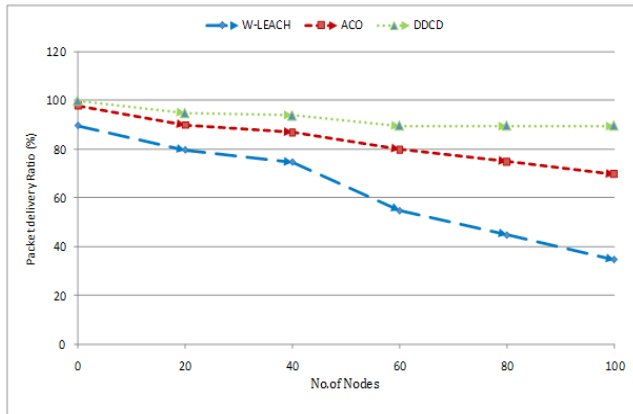


Figure 5. No. of nodes versus packet delivery ratio.

for the cluster heads, and increases the lifetime of sensor nodes and the network. In this algorithm, the sensor nodes in different clusters have a low correlation degree, when compare to the sensor nodes of the same clusters have a high correlation degree. The execution of this clustering procedure is more energy efficient and obtained better data description than the other clustering procedures like Weighted-Low Energy Adaptive Clustering Hierarchy and Ant Colony Optimization. As a result of this, DDCD technique is used where the sensor nodes are densely established.

6. References

1. Yick J, Mukherjee B, Ghosal D. Wireless sensor network survey. *Comput Netw.* 2008; 52(12):2292–330.
2. Zhu C, Zheng C, Shu L, Han G. A survey on coverage and connectivity issues in wireless sensor networks. *J Netw Comput Appl.* 2012; 35(2):619–32.
3. Yuan J, Chen H. The optimized clustering technique based on spatial-correlation in wireless sensor networks. *Proc IEEE Youth Conf Inf Comput Telecommun, YC-ICT; 2009 Sep.* p. 411–4.
4. Vuran MC, Akan OB, Akyildiz IF. Spatiotemporal correlation: Theory and applications for wireless sensor networks. *Comput Netw.* 2004; 45(3):245–59.
5. Ma Y, Guo Y, Tian X, Ghanem M. Distributed clustering-based aggregation algorithm for spatial correlated sensor networks. *IEEE Sensors J.* 2011 Mar; 11(3):641–8.
6. Vuran MC, Akyildiz IF. Spatial correlation-based collaborative medium access control in wireless sensor networks. *IEEE/ACM Trans Netw.* 2006 Apr; 14(2):316–29.
7. Rajeswari A, Kalaivaani P. Energy efficient routing protocol for wireless sensor networks using spatial correlation based medium access control protocol compared with IEEE 802.11. *Proc Int Conf PACC; 2011 Jul.* p. 1–6.
8. Sivakumar S, Diwakaran S. An energy efficient routing technique to improve the performance of wireless sensor network through adaptive tree based sink relocation. *2014 IEEE International Conference on Computational Intelligence and Computing Research; 2014.*
9. Al-Karaki JN, Ul-Mustafa R, Kamal AE. Data aggregation and routing in wireless sensor networks: Optimal and heuristic algorithms, *Comput Netw.* 2009; 53(7):945–60.
10. Hua C, Yum TS. Optimal routing and data aggregation for maximizing lifetime of wireless sensor networks. *IEEE/ACM Trans Netw.* 2008 Aug; 16(4):892–903.
11. Fan G, Jin S. Coverage problem in wireless sensor network: A survey. *J Netw.* 2010; 5(9):1033–40.
12. Kaur J, Gaba GS, Miglani R, Pasricha R. Energy efficient and reliable WSN based on improved leach-R clustering techniques. *Indian Journal of Science and Technology.* 2015 Jul; 8(16):1–6. DOI: 10.17485/ijst/2015/v8i16/70896.
13. Rao Y, Chalapathi Ch, Rani S, Basha SS. Efficient clustering and energy optimization algorithm for WSNs: An approach in environment monitoring system. *2015 International Conference on Innovations in Information Embedded and Communication Systems (ICIIECS); 2015.* p. 927–35.
14. Okdem S, Karaboga D. Routing in wireless sensor networks using ant colony optimization. *Proc 1st NASA/ESA Conf Adapt Hardware Syst; Istanbul.* 2006 Jun 15-18. p. 401–4.
15. Zhang C. Cluster-based routing algorithms using spatial data correlation for wireless sensor networks. *Journal of Communications.* 2010 Mar; 5(3):232–8.
16. Garey ML, Johnson DS. *Computers and intractability: A guide to the theory of NP completeness.* San Francisco: W. H. Freeman; 1979.
17. Fei Y, Zhan Y, Wang Y. Data density correlation degree clustering method for data aggregation in WSN. *IEEE Sensors Journal.* 2014; 14(4):1089–98.