



Clustered Relay Node Placement for WSN through Metaheuristics Algorithm

Nageswar Rao Amgoth^{1*}

Rajendra Naik Bhukya¹

Nirmala Devi Lavadya¹

¹*Department of Electronics and Communication Engineering,
University College of Engineering, Osmania University, India*

* Corresponding author's Email: nagiitkgp@gmail.com

Abstract: In the field of Wireless Sensor Networks (WSNs), after the deployment of sensor nodes, transmission of data with an effective energy is becoming a critical issue due to the limited battery of sensor nodes. To solve this issue, most of the past developed approaches focused on the implementation of clustering in which the sensor nodes forward their data through Cluster Heads (CHs) to Base Station (BS). However, the main issue is to select an optimal CH or CHs such that every sensor node is ensured with at least one path to base station, directly, through CHs or through relay nodes. Towards this objective, we have proposed a new method called Multi-Objective Assisted Clustered Relay Node Placement (MOCRNP) to achieve improved energy efficiency along with an enhanced network lifespan. Firstly, to select the CH, the Multi-Objective Genetic Algorithm is introduced based on energy, distance and node density. Secondly, in the process of relay node placement, Multi-objective firefly algorithm is introduced and for this purpose, two constraints such as Coverage and Connectivity are considered. We have conducted several experimentations using several setups of the WSN and the performance is evaluated in terms of Network lifetime, average energy consumption, average residual energy, and throughput. On an average, the proposed approach has gained an improvement of 150 sec network lifetime, 104 Kbps Throughput when compared with state-of-art methods.

Keywords: Wireless sensor networks, Network lifetime, Clustering, Genetic algorithm, Relay nodes, Throughput, Energy.

1. Introduction

In recent years, WSNs have gained a greater importance due to various advanced properties like Communication technology, embedded technology, distributed data processing, and integrated sensor technology etc. These advanced features of WSNs make it to employ in various fields, including, Disaster prevention, manufacturing, health care, traffic management, environmental monitoring, national defense and military [1]. In WSNs, the Sensor Nodes (SNs) gather and process the physical and environmental data from area monitored by the network and send it to the BS. The SNs accomplish these tasks in a collaborative fashion [2]. Generally, to monitor an area, there is a need of more number of sensor nodes which have limited resources like storage, computational and communication capabilities. For example, the Micaz [3] sensor

developed by Crossbow is a representative sensor node type and equipped with 4KB RAM, 128 KB Flash memory, CC2420 chip and Atmega128L microprocessor. Moreover, the sensor nodes are of having limited battery capacity and if it is not utilized in an optimal manner, the network lifetime will be affected. Thus network will get failed after a particular percentage of node's die. Hence there is a need of reducing energy consumption of SNs such that the network lifetime will get prolonged, is major hurdle in WSNs.

To overcome this hurdle, many researchers have developed several types of solutions [4]. The first and basic method is to transfer the entire data sensed at SNs to the BS. However, this approach has more energy consumption, especially for sensor nodes which are far way from base station. The next strategy is based on the clusters which is more suitable for energy limited WSNs which have continue data flow. A clustered WSN is generally

consists of BS and some clusters. Every cluster is formulated through one CH and some Cluster Nodes (CNs). The major responsibilities of CH are to collect the information from all CNs, aggregate the information and then forward the entire information to BS either directly, or through other CHs or via one or multiple Relay Nodes (RNs). However, the main problem in cluster assisted network is the selection of CH. A random CH selection has serious effects on the network lifetime. Moreover, the CH is also a SN and it may be profoundly loaded since it needs to aggregate and transmit the information of the entire cluster, consequences to quick energy depletion. This may condense the CH's lifespan. Hence, lowering the energy consumption of CHs generally plays an important role in cluster assisted WSNs.

There are several clustering protocols that have been developed in the past [5]. LEACH [6], PEGASIS [7] and LEFC [8] are some of the clustering protocols for WSNs. In the LEACH protocol, the selection of CH is accomplished in a random fashion and performs the data aggregation process in each round. Compared to the LEACH, PEGASIS has less energy consumption. Next, the LEFC clustering protocol clusters the network based on the location information of SNs. Even though these methods had achieved an efficient performance in the reduction of energy consumption, the methodology followed for CH selection is limited to several constraints. For example, if the CH selection is accomplished based on the location, the connectivity and coverage problems may arise. In the case of connectivity, the all the CHs may not be nearer to the BS by which the SNs has to suffer an excessive delay for data transmission. In the case of coverage, few clusters may not cover all the SNs in the network. In such condition the SNs will transmit the information directly to BS and depletes the energy quickly.

To overcome the above mentioned problems, in his paper, we have focused to develop an efficient clustering strategy in which the CH selection strategy follows multiple objectives such as energy cost, distance and node density. In this mechanism, these three objectives are considered as Fitness Parameters and a composite Fitness Function is formulated by combining them. And it is solved through Multi-Objective Genetic Algorithm. The energy cost solves the network lifetime problem, distance solves the connectivity problem and node density solves the coverage problem. To further maintain the connectivity as well as coverage, a novel relay node placement (RNP) is proposed in which the CH nodes will communicate to the BS through RNs. After the CHs aggregate the data, the Multi-Objective Firefly

(MOFF) algorithm is accomplished to choose the RNs between CH and BS such that the energy of CH is also minimized.

Remaining paper is prepared as follows: the details of related work are explored in section II. Next the details of MOCRNP are explored in section III. Section IV explores the details of simulation experiments and performance evaluation. Finally, the conclusions are explored in section V.

2. Related work

Many authors have propounded their research strategies and reported a broad review of the various routing protocols based on clustering mechanism [9]. The clustering idea was first exposed by LEACH protocol and further onwards a paramount importance has given for clustering as well as for CH selection in WSNs. Similar to LEACH, some more clustering protocols are developed in earlier. They include EECS [10], HEED [11], MOECS [12], AEEC [13], and pLEACH [14]. Among these algorithms, MOECS have found a longer lifetime compared to HEED and EECS. AEEC and pLEACH are the extension of basic LEACH protocol and performs better than LEACH protocol. However, there is no comparison between them and MOECS has been done. All these approaches strictly formulated the clusters with energy factor but not focused on other parameters such as node density and the nodes which are not covered by any cluster. Moreover, they didn't focus on the energy depletion rate of CH.

Considering the dynamic process of CH selection in which the responsibility of CH is alternated at some pre-defined conditions, the optimization plays a significant role. The optimization technique opted for the CH selection aims to reduce its energy consumption. But the selection of an optimal CH assisted routing in energy efficient manner is a "Non-deterministic Polynomial-time hard (NP-Hard)" problem. To solve this issue, several metaheuristics are employed that incorporates some important factors for the selection of CH in their Fitness Function (FF) establishment [15]. During this process, the metaheuristic algorithm is applied over the basic characteristics of network such that the FF gets converged and gives a better solution. "Particle Swarm Optimization (PSO)" [16, 22, 23], "Genetic Algorithm" [17-19], and "Artificial Bee Colony (ABC)" [20, 21] are some of the noticeable metaheuristic algorithms considered in earlier cluster based routing protocols in WSNs. Among these algorithms, GA has more efficiency due to its robust nature and for every round of optimization, it ensures

an improved optimization. Moreover, the GA generally focuses on the SN's energy saving.

In earlier, several cluster based routing approaches have been applied GA scrupulously to optimize the selection of CH. Gupta S. K and Jana P. K [17] considered the GA for clustering in WSNs. The clustering process considered the Residual Energy (RE) of CHs and also the distances from CHs to their corresponding CNs. Next, in [18], a cluster based fault tolerance method is developed using GA. Here the clusters are formulated according to the energy efficient distance between CHs and SNs. For each CH, a set of backup nodes are nominated through GA based on the RE and sponsored coverage.

T. Bhatia et al, [19] proposed a new scheme, coined as GADA-LEACH, makes use of evolutionary GA for improving CH selection in legacy LEACH routing protocol in sensor networks. The concept of relay node is introduced which acts as an intermediary between CH and base station (BS) to ease the communication between the CH and BS. The fitness function is modeled with respect to three parameters such as Residual Energy of Nodes, Distance between CH and nodes, and distance from all CHs to BS. However, this approach not focused on the Residual energy at cluster node and node density.

Next, a novel hierarchical clustering approach proposed by Karaboga D et al, [21] employed ABC algorithm based on the energy cost. But the single energy factor is not sufficient for an effective clustering. S. Bayrakali and Erdogan S. Z [24] proposed a "Genetic Algorithm Based Energy Efficient Clustering (GABEEC)". This method is an enhancement to the convention LEACH. In GABEEC, GA is accomplished in two phases; they are set-up phase and steady state phase. In the setup phase, the clusters are formed and they are kept constant for entire network. However, this work not focused on the scalability of network. As the network size increases there is a need of more number of clusters to maintain the coverage.

Next, considering the entropy, a new algorithm called as "Genetic Dual Fuzzy Entropy Clustering (GDFEC)" is proposed by S. S Kumar and S. Vishwas [25] to obtain total dissipated energy. In [26] S. P Singh et al. proposed "Genetic Algorithm Based Energy efficient clustering (GAEEC)" in which the GA is accomplished two times with dissimilar operators and parameters to achieve optimal and static clustering. Further the CH selection is accomplished by considering the RE and Average Transmission Cost (ATC) to enhance the lifetime of network. The node density is also an

important parameter which is not focused in this approach.

Pritee P. and Sireesha R, [27] proposed the optimization of clustering through GA which will help to optimize the communication distance. In this approach, the FF is designed based on the active and sleep modes of SNs and cluster overhead. Next, Vipin pal et al, [28] considered GA to optimize the selection of CH based on the optimization of FF. In this approach, the FF is formulated with respect to Communication distance between BS and CH, Cluster Distance and Energy. However, the energy consumption at sleep modes is very negligible. Moreover, the nodes left after clustering need a direct link to BS which is typically complex.

X. Yuan et al, [29] proposed "Genetic Algorithm based Self-organizing network clustering (GASONEC)" method to formulate dynamic and optimal clusters in WSNs. The FF is modeled with respect to RE, number of nodes in the neighborhood, and distance to the BS and it searches for an optimal and dynamic structure of WSN. M. Elhoseny et al, [30] introduced a new GA assisted CH selection technique for both single-hop transmission and multi-hop transmission networks. In this approach, the FF is modeled based on six metrics such as degree of mobility, node vulnerability, energy aware distance, number of close neighbors, consumed energy and remaining energy. However, the additional mobility assisted parameters constitutes an extra computational burden over the optimization process.

From these studies, we have noticed that almost all the approaches have supposed that the sink or BS is located at the midpoint of network which is not feasible in real time applications, for example, if sensor network is deployed in forest or fire detection area, the BS won't be lie at center of network. In such scenarios, the BS may lie at the outside of network in which any CH of CN needs multi-hop communication to transmit their data. This introduces a problem of long-haul transmission and to solve this problem, this paper developed a multi-objectives oriented clustering strategy such that every CH will transmit the information to the BS through RNs.

3. Proposed approach

3.1 Overview

In this paper, we have proposed a new CH Selection followed by RNP framework based on the two metaheuristic algorithms such as GA and Firefly Algorithm (FF). The proposed framework is named as Multi-Objective Clustered Relay Node Placement (MOCRNP). In this framework, the Genetic

Table 1. Notations

Notation	Meaning
E_e	unit energy consumption
E_T	Energy for transmission
E_R	Energy for receiving
E_D	Energy for data aggregation
E_a	extra energy due to amplifier circuit
E_C	Energy Cost
E_{Re}	Residual Energy
L_T	Lifetime
$d(s, CH)$	Distance between source and CH
d_{IC}	Intra_Cluster Distance
d_{Far}	Farthest distance
d_A	average distance
d_{ij}	distance between node i and node j
NC_i	Node cover of i^{th} cluster
FP	Fitness parameter

Algorithm is adopted for CH selection and Firefly algorithm is adopted for relay node selection. For both algorithms, the fitness function is modeled as Multi-objective Function and hence they are called as “Multi-objective Genetic Algorithm (MOGA)” and “Multi-objective Firefly Algorithm (MOFF)”. Under the Multi-objective constraint for GA, this paper formulated three objectives such Energy, Distance and Node Density. Then by combining these three objectives, the fitness function is formulated and it is optimized through GA. Table 1 defines several notations used in this paper.

Under the first objective, energy constraint considers both the energy Cost at CH level and also the energy cost at SN level. Next, under the second objective, the distance constraint considers the distance from every SN to BS. Finally, at third objective, the node density considers the total possible number of SNs covered by one candidate CH. Based on these three objectives, the GA finds optimal CHs. Once the CHs are discovered and the entire network is clustered into several clusters, then the proposed framework applies MOFF to find out the optimal location for relay node placement. Here the MOFF considers two objectives such as Connectivity and Coverage constraints to obtain an optimal number of relay nodes.

3.2 Fitness parameter 1 (Energy Cost)

Since the energy is a main factor in WSNs through which the lifetime followed by Quality of Service (QoS) is decided, we have considered it as a first and major parameter. The basic aim of any routing approach in WSNs is to lessen the total amount of energy consumption occurring in a network such that the nodes will live for more time and hence the network lifetime will be increased.

Under this section, the energy cost is considered as first Fitness Parameter (FP) and formulated with respect to energy consumption at both SN level and cluster level. At cluster level, the FP is modeled by assuming that a node is chosen as a CH and at node level, it is modeled according to the normal operations of node. Under both levels, the possible energy costs are studied and based on the individual fitness parameters, one final fitness parameter is formulated.

3.2.1. Energy cost (EC) at cluster head

The major responsibility of a CH is to receive the information packets from SNs and transmit them to the BS. Additionally, the CH is also in charge for the data aggregation which was received from its CNs. As the CN's consumes, the energy consumption of a CH also rises proportionately. Hence if the total number of cluster member nodes is minimized, then the energy cost at CH also reduces. Thus the lifetime of CHs will get increased. Let E_T , E_R and E_D are the amount of energies consumed at CH due to data transmission, data reception and data aggregation respectively. For a packet of size k bits, these three ECs are computed as;

$$E_R = E_e \times k \quad (1)$$

$$E_D = E_e \times k \times N \quad (2)$$

$$E_T = E_e \times k + E_a \times k \times d^2 \quad (3)$$

Where E_e is the unit energy consumption, k is the number of bits present in a packet, N is total number of cluster member nodes, E_a is the extra energy consumed by an amplifier circuit to maintain particular level of Signal to Noise Ratio (SNR), d is the Euclidean distance between CH and BS or RN.

Here the total energy cost (E_C) at a CH per round is formulated as the summation of Transmission energy, Reception energy and Data Aggregation Energy, as

$$E_C = E_R + E_D + E_T \quad (4)$$

Then the lifetime (L_T) of a CH with the remaining energy left, called as Residual Energy (E_{Re}), is calculated as;

$$L_T = E_{Re} / E_C \quad (5)$$

Based on the Eq. (5), we can state that the maximum lifetime is proportional to the maximum

residual energy or minimum energy cost. Now our main objective is to decrease the energy cost of the CHs so that the lifetime of network will get maximized. Based on this analysis, the fitness parameter 1 at CH is formulated as;

$$f_{p_1} \propto L_T \quad (6)$$

Where f_{p_1} is the sub-fitness parameter of Fitness Parameter 1 (FP1), which is directly proportional to the lifetime of CH. Based on this, the node which has less energy consumption is chose as CH.

3.2.2. Energy cost at cluster node

Unlike the Energy cost computation at CH which is composed of three forms of ECs such as transmission cost, aggregation cost and reception cost, the CN have only one energy consumption, i.e, due to the data packets transmission to the CH. However, the EC of a CN is in direct proportion with the distance between CN and its respective CH. To minimize the EC at cluster node, the distance between CN and its respective CH should be lessened. Here the distance between CNs and their corresponding CHs is defined as Intra_Cluster Distance. To minimize this Intra_Cluster Distance between CN and its corresponding CH, there is only one possible solution, i.e, assigning a sensor node to its nearest CH. Let (s_x, s_y) and (c_x, c_y) be the coordinates of a cluster node and CH, then the distance between them, $d(s, CH)$ is computed as;

$$d(s, CH) = \sqrt{(s_x - c_x)^2 + (s_y - c_y)^2} \quad (7)$$

Lets d_{min} and d_{max} be the minimum and maximum Intra_Cluster Distances, then the average Intra_Cluster Distance (d_{IC}) is computed as;

$$d_{IC} = \frac{1}{N} \sum_{i=1}^N \left(\frac{d(s_i, CH) - d_{min}}{d_{max} - d_{min}} \right) \quad (8)$$

Here the main objective is to lessen the average Intra_Cluster Distance such that the EC of CN will be reduced. Hence the FF has an inverse relation with the average Intra_Cluster Distance, mathematically represented as

$$f_{p_2} \propto \frac{1}{d_{IC}} \quad (9)$$

Now, the final fitness parameter 1 (FP1) is obtained by combining the fitness parameter at CH (f_{p_1}) and the fitness parameter at Cluster Node (f_{p_2}), as

$$FP_1 \propto \frac{L_T}{d_{IC}} \quad (10)$$

i.e,

$$FP_1 = \beta \times \frac{L_T}{d_{IC}} \quad (11)$$

Where β is a proportionality constant. Now our main aim is to maximize the FP1 so that the lifetime of network will be more with less EC.

3.3 Fitness parameter 2 (distance)

Generally, the energy cost has linear relation with distance, i.e, as the distance between node increases, the amount of energy required to communicate also increases. Hence we have considered the Euclidean distance between SNs and BS as a second FP in the selection of CH. Let B be the base station, and $S(s_1, s_2, \dots, s_N)$ be the set of SNs deployed randomly in a network, the distance between SN, s_i and BS, B, is denoted as $d(s_i, B)$. In the network, all the SNs are not having equal distances with BS. Some nodes lies very nearer, some nodes lies nearer and other nodes lies far from base station. Under the FP2 computation, initially the Euclidean distances between all SN and BS are computed with the help of Eq. (7) and it is denoted as $d(s_i, B), \forall i = 1, \dots, N$. Based on these distances, the farthest node is discovered as the node which has larger distance from base station and it is formulated as;

$$d_{Far} = \max_{i \in N} (d(s_i, B), \forall i = 1, \dots, N) \quad (12)$$

And the average distance is computed as

$$d_A = \sum_{i=1}^N d(s_i, B) / N \quad (13)$$

The fitness parameter 2 (FP2) is formulated by combining the above Eq. (12) and Eq. (13) as;

$$FP_2 = \sum_{i=1}^N \left(\left(\frac{d(s_i, B)}{d_{Far}} \right) + \left(\frac{1}{d_A} \right) \right) \quad (14)$$

Here the FP2 is measured by summing up the distance cost earned for every SN s_i , where $i = 1$ to N . The involvement of average distance in FP2 makes the CH to be selected which was close to the

BS and the CNs are selected which have distances less than or equal to the d_A . From the above expression shown in Eq. (14), we can state that larger the FP2 value, chances are more for a node to get chosen as CH since it guarantees the minimum probable distance to communicate with BS. **Note:** According to FP2, the SN which is nearer to the BS is chosen as CH, hence the EC will get minimized. However, a CH closer to the BS may not cover maximum SNs in the network. Hence there is a need of additional cluster and corresponding CHs. In such case, they are definitely far away from the BS and need the help of RNs to forward the data to base station.

3.4 Fitness parameter 3 (node density)

Node density is an important factor through which we can reduce total number of RNs need to be positioned. As the CH node covers more SNs, the number of clusters will get reduced followed by minimum number of RNs. Hence we have considered the node density as a FP3 through which one SN is chosen as CH if it covers larger number of SNs. For a given node, the number of SNs surrounded by it is evaluated initially. For this purpose the surrounded nodes must be within the communication range of respective node and they are discovered based on their distances to the respective node.

For this purpose, initially, for all nodes in the network, a common communication range is defined (in this paper we have defined the communication range as 25% of network area). Next, the Euclidean distances are evaluated between all nodes. Next, for a given node the Euclidean distances are compared with its communication range and the nodes which fall below the range are considered as neighbor nodes or surrounding nodes. Hence a SN is selected as CH which has more neighbor nodes, means, having larger density of nodes. Towards this formulation, initially we construct a distance matrix, for a given set of sensor nodes $S(s_1, s_2, \dots, s_N)$, as

$$D = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1N} \\ d_{21} & d_{22} & \dots & d_{2N} \\ \vdots & \vdots & \dots & \vdots \\ d_{N1} & d_{N2} & \dots & d_{NN} \end{bmatrix} \quad (15)$$

Where d_{ij} is the distance between node i and node j .

In the above distance matrix, the first row denotes the Euclidean distances between first sensor node and the remaining N-1 sensor nodes. For example, the distance term d_{12} is the Euclidean distance between SN-1 and SN-2. Similarly, the values in the second

row denote the distance between SN-2 and the remaining N-1 SNs. In this way, the i^{th} row denotes the distances of an i^{th} node from the remaining node in the network.

Note: in the above distance matrix, the diagonal values are zero, denotes the self-distance.

From the distance matrix, to find the node cover (SNs count those comes under the communication range of a specific SN), we have to scan the distance matrix row-by-row. For a given row (R_i) every element (R_{ij}) is compared with the communication range of i^{th} sensor node and if it is found that it is less than or equal to the communication range, then that particular j^{th} sensor node is buffered into the respective node cover of i^{th} node. Mathematically it is defined as follows;

$$NC_{i=1,\dots,N} = \text{find} \left(d_{ij} \leq C_R(i), j = 1, \dots, N \right) \quad (16)$$

Where NC_i is the Node Cover of the i^{th} node, denotes the node count those are within the boundary of specific node and $C_R(i)$ denotes the communication range of i^{th} node. Finally the fitness parameter 3 (FP3) is formulated as the one node which has maximum cover size, i.e, the node cover which has maximum size is selected as CH. Mathematically it is formulated as

$$FP_3 = \max_{i \in N} (NC_i), i = 1, 2, \dots, N \quad (17)$$

The obtained FP3 denotes the maximum node density of i^{th} node that node is chosen as CH. Fig. 1 shows the simple illustration about the node density assisted CH selection.

3.5 Fitness function

Based on the above mentioned individual fitness parameters such as FP1, FP2, and FP3, a final fitness function are designed by combining all these parameters. For every FP, an individual weight factor is assigned such that it signifies the importance of that particular FP. Based on these three FPs, the final fitness function is modeled as

$$FF = \frac{1}{(w_1 \times FP_1) + (w_2 \times FP_2) + (w_3 \times FP_3)} \quad (18)$$

Where w_1 , w_2 , and w_3 denotes the weight factors, EC is weighted by w_1 , distance is weighted by w_2 and node density is weighted by w_3 . In this paper, all the weights are assigned with equal priority, i.e, $w_1 = w_2 = w_3 = 0.3333$, subjected to the constraint $w_1 + w_2 + w_3 = 1$.

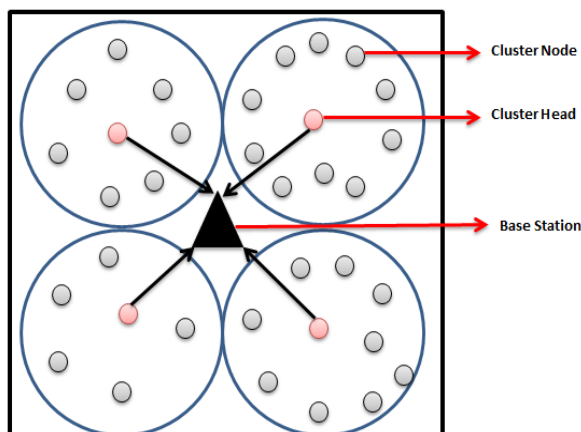


Figure. 1 Node density assisted CH selection

3.6 Algorithm

In this paper, to optimize the fitness function, we have accomplished multi-objective GA which considers three FPs as three objectives. GA is a type of metaheuristic algorithm which helps to find an optimal solution through the evaluation of natural evolution process. According to the standard GA, initially some simulations are initiated to obtain an Initial Population, termed as IP and comprises of one or more chromosomes. In a chromosome, a single unit is termed as gene. After the generation of IP, a FF is accomplished to choose the individual entities which had shown an optimal performance.

Next two random individuals were chosen as parents and new children are produced through Cross Over (CO) operation. A Mutation Operation (MO) is accomplished to obtain children. Next the FF is accomplished to compute the fitness of parents as well as the generated new children. An important note is that mutation process generates the new children and eh selection process generates the parents. Lastly the algorithm choses half of the individuals which have shown better performance and those were added to IP and creates New Population (NP). If termination is obtained at NP, then the algorithm ends, otherwise the algorithm starts with new children generated based on CO and MO. The general process of GA is shown in the following Fig. 2.

The step-by-step process of proposed MOGA is illustrated as;

Step 1: Initialize the parameter of GA, such as Number of nodes or chromosomes, Cross over rate (c_r), mutation rate (m_r), $gen = 0$, Initial energy, Communication range = 25% of network area, weights factors w_1 , w_2 and w_3 .

Step 2: choose some chromosomes randomly and for these chromosomes, the fitness parameters are

computed through Eq. (11), (14) and (17). Further the final fitness is computed through Eq. (18).

Step 3: Repeat the process for every chromosome and every time check the stopping criterion is reached. If stopping criterion is reached, then the chromosome node at which the stopping criterion is obtained is chosen as best fit chromosome. And after the network operates in a steady state manner.

Step 4: if stopping criterion is not reached, then increment the generation number, i.e., $gen = gen + 1$ and apply rank selection method to select two individual chromosomes as parents to generate children.

Step 5: Apply cross over to generate two new individuals, i.e., combining of parents with particular cross over rate c_r .

Step 6: Generate new chromosomes after CO process over the MO process with particular m_r to produce new children.

Step 7: Place the newly generated chromosomes in the new population and formulate new set of individuals.

Step 8: repeat from step – 2 to step – 8.

Even though the cluster assisted data transmission in WSNs has an optimal performance, in a distant communication, the clusters also need the help of relay nodes and hence we have extended the proposed work to rely node placement. Once the clusters are formed, they will check the distance from BS and if it is found that the BS is not in the communication range of any cluster, that cluster opts for RNP. Since there exists more than one such type of clusters, all of them need the assistance of RNs and hence we have extended the proposed work. Non-planned relay node deployment consequences to worst problems like too much cost and computation burden over the nodes. Hence to select an optimal location at which the RN can be placed, we have accomplished MOFF according to the methodology described in [31]. The MOFF find outs optimal locations as well as optimal number of RNs by considering the connectivity and coverage in mind. In this approach, the connectivity and coverage are considered as main constraints and a new FF is formulated based on these two constraints. The FF is solved through MOFF.

4. Simulation results

To alleviate the performance enrichment of proposed MOCRNP, here we have conducted extensive simulations over it through MATLAB software. Under this section initially the details of simulation set up is explored and then the obtained performance metrics are explained.

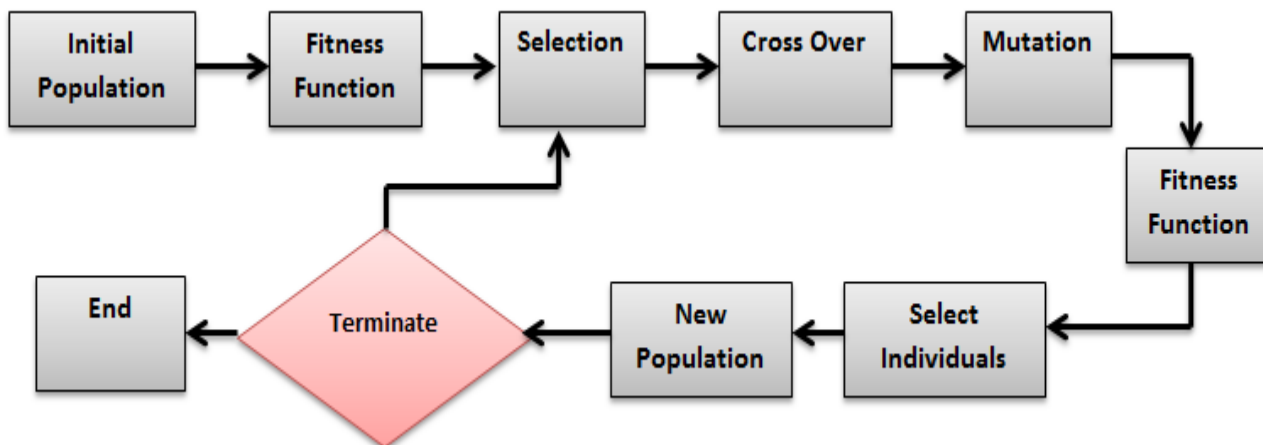


Figure. 2 General process of GA

4.1 Simulation set up

The performance of proposed MOCRNP is evaluated using MATLAB software. Under the simulation setup, we have a considered randomly varying network with varying node count. The varying node count is varied from 20 to 60 and the area of network is varied as Area = [300 × 300, 500 × 500, 700 × 700, 900 × 900, 1100 × 1100] and the BS is located at a random location. Unlike the conventional approaches which were generally assumed that the BS is located at the center of network, in this work, we have located the BS at random locations. For every simulation the location of BS is changed. Next, the transmission range of a sensor nodes is considered as one fourth of the network area, for example, if Area is 500 × 500 m², then the communication range is kept as 125 m. The communication range of both sensor node and relay nodes is assumed as same. The details of simulation parameters are presented in Table 2 Next, a sample

Table 2. Simulation set up

Parameter	Value
Number of nodes	[20, 30, 40, 50, and 60]
Network Area	[300 × 300, 500 × 500, 700 × 700, 900 × 900, 1100 × 1100]
Communication range	1/4 (Network Area)
Base station location	Random
Data packet size	2000 bits
Data type	Constant Bit Rate (CBR)
Energy threshold	10% of initial energy
Packet size	512 bits
w ₁ , w ₂ and w ₃	0-1
Population size	30
Simulation time	100 sec
Mutation rate	0.005
Cross over rate	0.5

random network and its clustered version are shown in Fig. 3 and Fig. 4 respectively.

4.2 Performance metrics

The performance metrics we have considered are Throughput, Total number of RNs, Network lifetime, and Average Energy Consumption. The definitions of performance metrics are done as;

Network lifetime: This metric is defined as the time taken by the process from the starting of a network to the time until any source sensor node is failed to send the data to the sink node. In this paper, the network is assumed to be disconnected when any of the source node is not connected to even one RN thereby it can't send the sensed data to sink. Higher the network lifetime, better the performance.

Average Residual Energy: This metric is defined as the total amount of energy left at the sensor nodes or CHs after every round of data transmission. The energy of network decreases progressively with the proceedings of information transmission. The efficiency of proposed approach is revealed when the average residual energy left after particular round is more.

Average Energy Consumption: This metric is defined as the total amount of energy consumed at the SNs or CHs after every round of data transmission. With an increment in the time proceeding for information transmission, the total energy consumption of network also increases gradually. The efficiency of proposed approach is revealed when the average energy consumed for every round is less.

Throughput: this metric is defined as the total number of packets received at BS successfully with in the specific simulation time. Higher the throughput, better the quality of service (QoS).

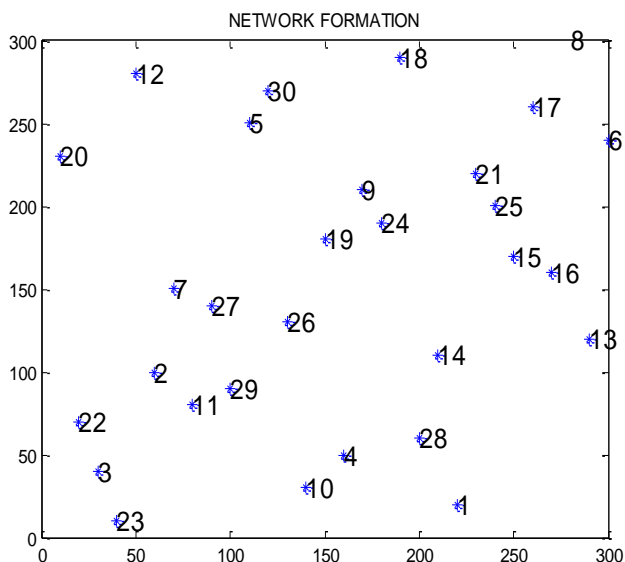


Figure. 3 randomly deployed network

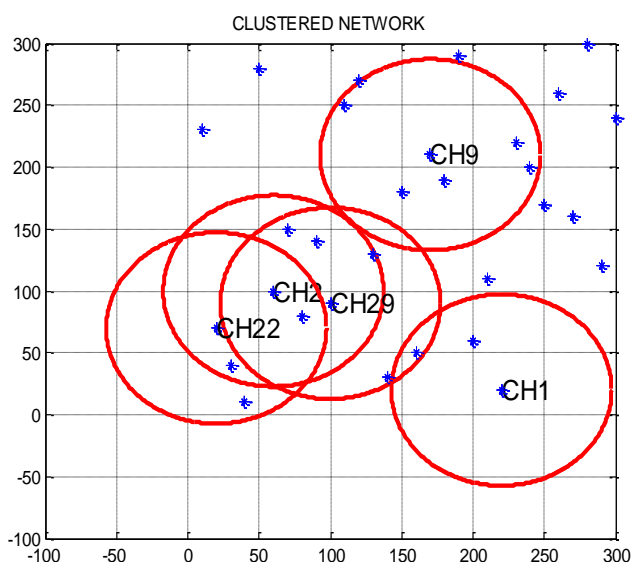


Figure. 4 Clustered network with five clusters

4.3 Results

In this section, we discuss the particulars of performance metrics evaluated after the simulation of proposed model over varying network characteristics. Simultaneously, a detailed comparison between the proposed MOCRNP and existing methods is also discussed here. Under the existing methods, we have compared the proposed MOCRNP with LEACH [6], LEACH-GA [28], GAEEC [26], GDEFC [25], and GABEEC [24]. Under the first simulation study, we have varied the node count and measured the network lifetime and the total number of rounds taken by MOGA until the first node dies. Under the second simulation study, we have varied the network size and measured the network lifetime. Under third simulation study we have varied the number of

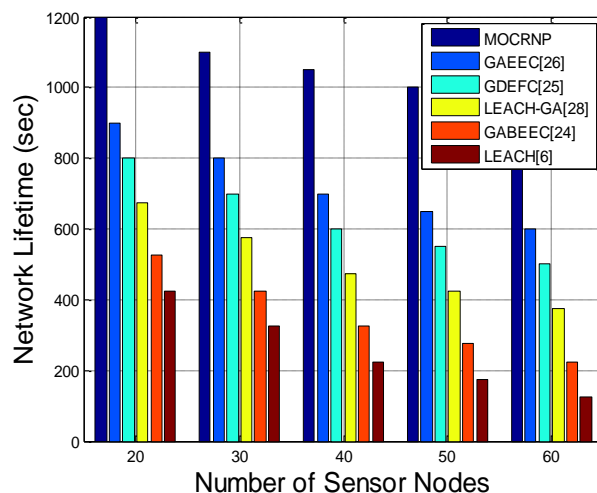


Figure. 5 Network lifetime (sec) vs. SNs count

rounds and measured the Average Energy consumption, and average RE. At last the simulation is conducted through varying number of clusters and the performance is measured through throughput (kbps).

Fig. 5 shows the varying lifetime of network for varying SNs count. It is observed that the proposed MOCRNP has better network lifetime compared to the existing methods. The reason behind that is, we have considered the energy consumption at both Cluster level and at node level. This consideration makes our method more relaxable because any node in the network can choose an optimal CH through which the energy consumption will be less. A sensor node seeks the CH which is more nearer to BS and also has higher residual energy. Hence the proposed MOCRNP has gained more network lifetime. Whereas the conventional approaches didn't provide a balanced energy consumption at CH level. In the conventional approaches, once a CH is selected, then the entire responsibility is accomplished through that CH only by which its energy will get deleted very quickly either by reducing the network lifetime. In comparison between conventional approaches, the LEACH has very much less network lifetime due to the random CH selection. Even through the GAEEC, GDEFC and GABEEC have accomplished GA for optimization, however, they didn't focus over the multi-objective assisted CH selection. On an average, the proposed MOCRNP gained a network lifetime of 1060 sec while the GAEEC, GDEFC, LEACH-GA, GABEEC and LEACH have gained 730 sec, 632 sec, 505 sec, 355 sec and 310 sec respectively.

In general, as the network size increases, the number of clusters required to cover entire nodes will also increase. At this phase, the intra-cluster distance between cluster node and CH will become high and for a cluster node, to transmit data to CH, it

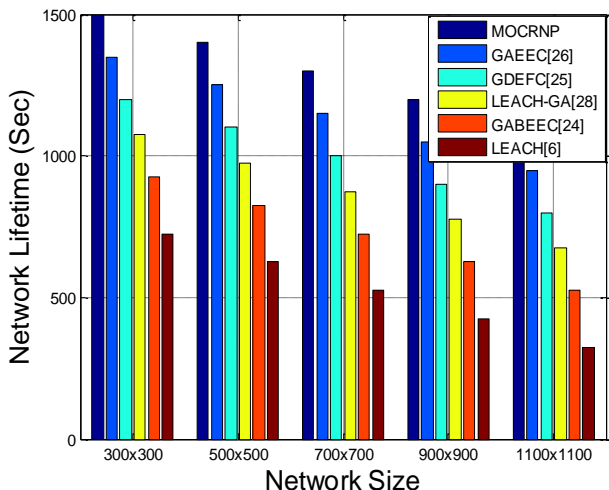


Figure. 6 Network lifetime (sec) vs. network size

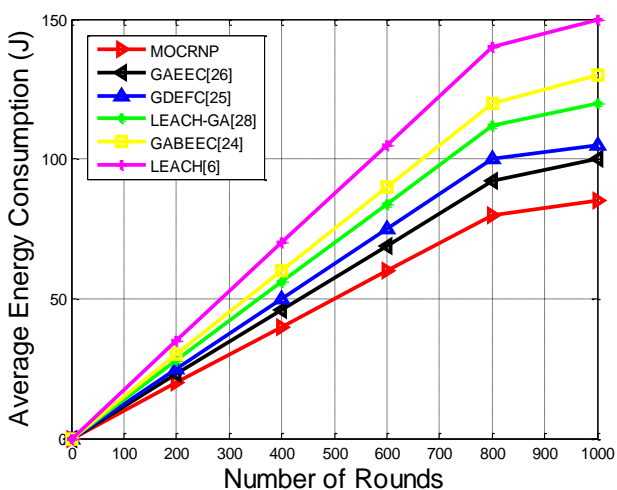


Figure. 7 Average energy consumption vs. number of rounds

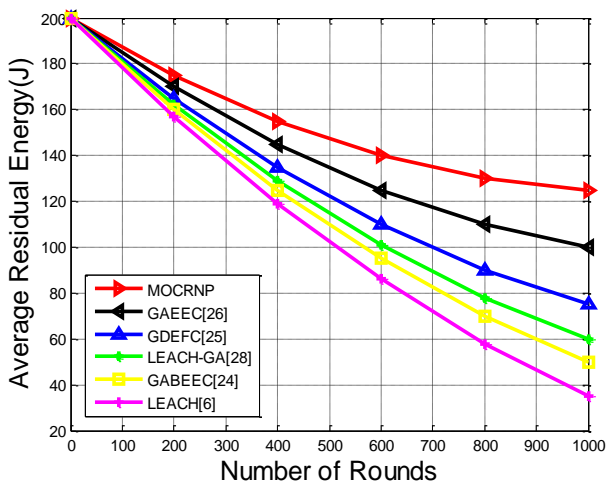


Figure. 8 Average residual energy vs. number of rounds

needs more energy, consequences to quick energy depletion. This is shown in figure.6, for varying network size from 300 × 300 to 1100 × 1100. For this simulation, the number of sensor nodes is kept

constant and it is taken as 40. For a given set of sensor nodes, if the network area is considered as 300 × 300, then the communication range is 75m and for 1100 × 1100, it is 275m. From this we can observe that the total number of sensor nodes possible to cover by a CH is more in second case when compared with the first case. But due to the fixed node set, the more intra-cluster distance resulting in a less network lifetime. Though the network lifetime has followed a decreasing characteristic, it is observed as high when compared to the conventional approaches. Since the proposed MOCRNP considered the node density also as one of the constraint during the selection of CH, it has obtained a less Intra-cluster distances between every CH and its respective cluster nodes, compared to the conventional approaches. Moreover, the proposed approach also placed relay nodes between CH and Base station for a distant communication. On an average, the proposed MOCRNP gained a network lifetime of 1300 sec while the GAEEC, GDFEC, LEACH-GA, GABEEC and LEACH have gained 1150 sec, 1020 sec, 875 sec, 725 sec and 632 sec respectively.

In WSN, the energy consumption has a linear relation with distance; the Average Energy Consumption (AEC) increases with an increase in distance and vice versa. With an increase in the number of rounds, the SN's energy gets reduced due to the continuous transmission of data to CH followed by BS. This is a general relation which holds for all the methods as represented by increasing characteristics in Fig. 7. With an increase in energy consumption at nodes, the remaining energy left at them will become less. Though the proposed approach also faced the same problem, it has less effect compared to the existing methods. Since the proposed approach focused to assign a new CH at every round by checking the residual energy status of CH, it has obtained less AEC compared to the conventional approaches. For a CH with higher node density, its AEC will be too high and at this phase, the proposed MOCRNP introduced relay node to help to the CH by which the CH can maintain an energy balance. Hence the proposed approach has less AEC even compared to conventional approaches, even though the number of rounds increased. LEACH and LEACH-GA have not focused in an optimized CH selection and hence they have faced a higher AEC. Next, even though the remaining methods such as GAEEC, GDFEC and GABEEC focused on the optimization through GA, they didn't consider multiple constraint at CH selection as well as no relay node deployment concept for distant communications or larger network areas. On an average, the proposed MOCRNP gained an average energy consumption of

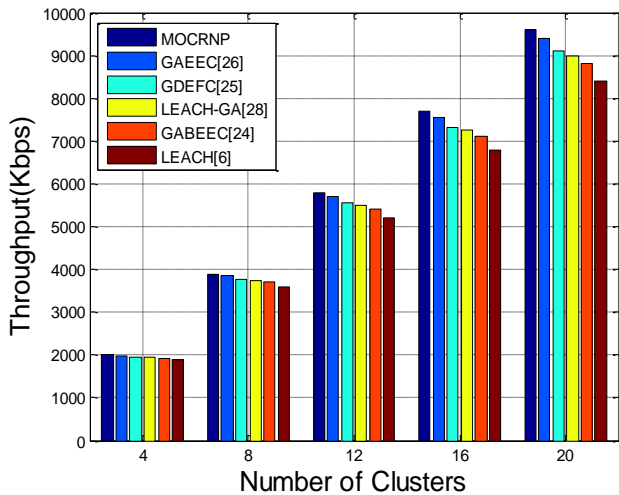


Figure. 9 Throughput vs. cluster count

57J while the GAEEC, GDFEC, LEACH-GA, GABEEC and LEACH have gained 66J, 71J, 80J, 86J and 95J respectively.

With an increase in the Rounds number, the SN's RE gets reduced due to the continuous transmission of data to CH and followed by BS. This is a general relation which holds for all the methods as represented by decreasing characteristics in Fig. 8. Since our main intention in this work is to attain an improved network lifetime we have to take care about the residual energy of a node. For a given sensor node, if its residual energy is more, then only it can sustain for more time. For this purpose, at every round of optimization through GA, MOCRNP checks for residual energy status of CH and if its found to be less, then additional RNs are deployed between CH and BS. Furthermore, the node density associated with CH is also reduced and the nodes which came out from one CH range are assigned for new CH which have higher RE. Hence the proposed approach has gained more Average Residual Energy compared to the conventional approaches. Even though the conventional approaches are also focused in this direction, they didn't view the energy consumption in multiple point of views as it was done in MOCRNP. On an average, the residual energy left at proposed MOCRNP is 154.1667Joules while the residual energy left at GAEEC, GDFEC, LEACH-GA, GABEEC and LEACH is 141.6667J, 129.1235J, 121.3124J, 116.4451J and 109.4157J respectively.

As the network size increases, there is a need of more number of clusters such that maximum nodes will get covered and can forward the information to the BS with the help of CHs. For a less number of clusters, all sensor nodes may or may not get covered and hence they can't transmit the information to BS, resulting in a less throughput. As the number of

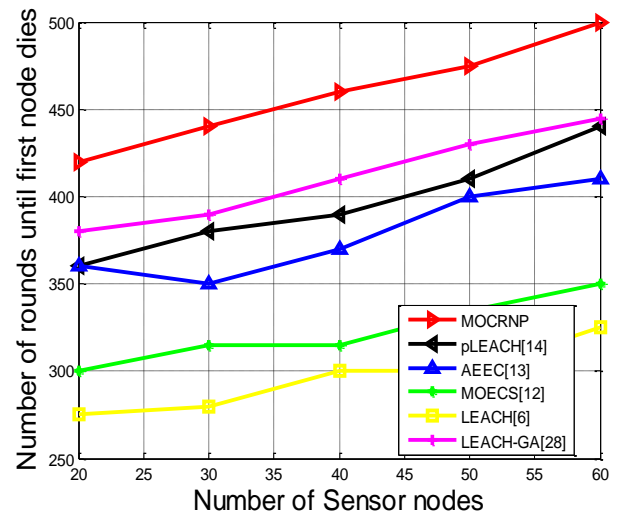


Figure. 10 Number of rounds until the first node dies vs. number of SNs

clusters increases, almost all sensor nodes will come into the range of at least one CH and then forward their data to BS, resulting in throughput at higher cluster count, as shown in Fig. 9.

Moreover, at every instant of clusters, the proposed MOCRNP has achieved a better throughput compared to the conventional approaches. Because, MOCRNP has also focused on the Coverage and connectivity constraints by which every node in the network will maintain at least one possible path to base station directly, through CHs or through relay nodes. For example, the sensor nodes which are more nearer to the base station won't require any additional assistance. Unlike, for distant and corner nodes, the assistance is mandatory, which is ensured in MOCRNP. LEACH has very less throughput and then for GABEEC which is an extension of LEACH. Even though LEACH-GA considered the Intra-Cluster distance, there is no choice for sensor nodes which are not covered by CHs. Further the GAEEC and GDFEC won't discuss about the coverage and connectivity constraints. On an average, the proposed MOCRNP gained an average throughput of 5800 Kbps while the GAEEC, GDFEC, LEACH-GA, GABEEC and LEACH have gained 5696 Kbps, 5542 Kbps, 5488 Kbps, 5384 Kbps and 5180 Kbps respectively.

In the optimization process, as the time increases, the optimization of fitness function increases. In this work, the optimization is accomplished for Optimal CH selection and at every round the fitness function is measured through energy, distance and node density. Considering this many factors makes the SN to live for more time because for every SN in the network, the FF assigns new CH if the current CH have less residual energy status. In this manner, for

every SN the number of rounds up to which it can survive will be more, as shown in Fig. 10.

Moreover, as the number of SNs increases, every node in the network can have multiple CH options and in every round they can choose an optimal CH which has more energy and forwards data to BS. Even at worst case, the SNs have an option in the form of RNs and hence the total number of rounds is more. Compared to the LEACH and its subsequent approaches such as pLEACH, AEEC and MOECS, the proposed MOCRNP has larger number of rounds until the first node dies, shows an improved performance. The improvement is observed through the following numerical results; in the proposed model, the average number of rounds taken for a first node to get die is approximately 459, while the tradition methods such as pLEACH, AEEC, MOECS, LEACH and LEACH-GA have taken 396, 378, 323, 296 and 411 rounds. The higher number of rounds denotes a higher lifetime and vice versa. From these numerical results we can understand that the proposed model has better network lifetime compared to existing methods.

5. Conclusion

In this paper, we have proposed an energy, coverage and connectivity balanced clustering algorithm and relay node placement algorithm through Genetic Algorithm and Firefly Algorithm respectively. Under clustering process, we have focused on the energy consumption of CH as well as Cluster node and at the optimization through GA, the fitness function searches for an optimal CH which has more residual energy, minimum distance to base station and maximum nod density. Next, at the relay node placement, the firefly algorithm focused on the connectivity and coverage such that for the clusters heads which are far away from base station can connect to base station with optimal coverage. We have conducted several experiments using different scenarios of the WSN and it has been shown that the proposed MOCRNP scheme outperforms the conventional approaches, namely LEACH, LEACH-GA, and GABEEC in terms of Network lifetime, average energy consumption, average residual energy, and throughput. From the comparative analysis, the observed concrete results are depicted below;

1. In the perspective of energy consumption, the proposed approach has gained a reduced energy consumption of approximately 22% of energy consumed by existing methods;

2. In the perspective of throughput, the proposed approach has achieved an improved throughput of 342 Kbps when compared with existing methods.

3. Finally, the improved network lifetime of proposed approach is noticed as approximately 500 sec when compared with existing methods.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, methodology, software simulation, investigation and original draft of this paper by N. R. Amgoth. Validation, formal analysis and supervision by R. N. Bhukya. Visuvalization, data curation, review and editing by N. D. Lavadya.

References

- [1] J. S. Leu, T. H. Chiang, M. C. Yu, and K. W. Su "Energy Efficient Clustering Scheme for Prolonging the Lifetime of Wireless Sensor Network with Isolated Nodes", *IEEE Communications Letters*, Vol. 19, No. 2, pp. 259-262, 2015.
- [2] I. F. Akyildiz, W. Su, S. Yogesh. S. W. McLaughlin "Wireless sensor networks: a survey", *Journal of Computer Networks*, Vol. 38, No. 4, pp. 393-422, 2002.
- [3] A. A. Nurul, M. Drieberg, and P. Sebastian. "Deployment of MICAZ mote for Wireless Sensor Network applications", In: *Proc. Of International Conf. on Computer Applications and Industrial Electronics*, Penang, Malaysia, pp. 303-308, 2011.
- [4] M. A. M. Vieira, C. N. Coelho, D. C. Da Silva, and J. M. da Mata, "Survey on wireless sensor network devices", In: *Proc. of IEEE Conf. on Emerging Technologies and Factory Automation*, Lisbon, Portugal, pp. 537-544, 2003.
- [5] B. P. Deosarkar, N. S. Yadav, and R. P. Yadav, "Cluster head selection in clustering algorithms for wireless sensor networks: A survey", In: *Proc. of International Conf. on Computer, Communication and Networking*, St. Thomas, IV, USA, pp. 1-8, 2008.
- [6] W. R. Heinzelman, P. C. Anantha, and B. Hari. "Energy-efficient communication protocol for wireless micro-sensor networks", In: *Proc. of 33rd Annual Hawaii International Conf. on System Sciences*, Maui, HI, USA, pp. 1-10, 2000.
- [7] S. Lindsey and S. R. Cauligi, "PEGASIS: Power-efficient gathering in sensor information

- systems”, In: *Proc. of IEEE Conf. on Aerospace, Big Sky, MT, USA*, pp. 3, 2002.
- [8] Y. F. Huang, N. C. Wang, and M. C. Chen, “Performance of a hierarchical cluster-based wireless sensor network”, In: *Proc. IEEE International Conf. on Sensor Networks, Ubiquitous, and Trustworthy Computing*, Taichung, Taiwan, pp. 349-354, 2008.
- [9] S. Tyagi and N. Kumar, “systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks”, *Journal of Network and Computer Applications*, Vol. 36, No. 2, pp. 623-645, 2013.
- [10] O. Younis and S. Fahmy, “HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks”, *IEEE Transactions on Mobile Computing*, Vol. 3, No. 4, pp. 366-379, 2004.
- [11] MaoYe, C. Li, G. Chen, and j. Wu, “EECS: an energy efficient clustering scheme in wireless sensor networks”, In: *Proc. of the 24th IEEE International Conf. on Performance, Computing, and Communications*, Phoenix, AZ, USA, pp. 535-540, 2005.
- [12] N. Aslam, P. William, R. William, and S. S. kumar, “Energy efficient cluster formation using a multi-criterion optimization technique for wireless sensor networks”, In: *Proc. of 4th Annual IEEE Conf. on Consumer Communications and Networking*, Piscataway, NJ, pp. 650-654, 2007.
- [13] O. B. jargal and Y. Kwon, “AEEC – adaptive and energy efficient clustering algorithm for content based wireless sensor networks”, In: *Proc. of 2nd International Conf. on Computer Science and Its Applications*, Jeju, Korea (South), pp. 1-6, 2009.
- [14] H. Gou, Y. Yoo, and H. Zeng, “A partition-based LEACH algorithm for wireless sensor networks”, In: *Proc. of the Ninth IEEE International Conf. on Computer and Information Technology*, Xiamen, China, pp. 40-45, 2009.
- [15] C. W. Tsai, T. P. Hong, and G. N. Shiu “Metaheuristics for the lifetime of WSN: A review”, *IEEE Sensors Journal*, Vol. 16, No.9, pp. 2812-2831, 2016.
- [16] Riham S. Y. Elhabyanand and Mustapha C. E. Yagoub, “Two-tier particle swarm optimization protocol for clustering and routing in wireless sensor network”, *Journal of Network Computer Applications*, Vol. 52, No. X, pp. 116-128, 2015.
- [17] S. K. Gupta and P. K. Jana, “Energy efficient clustering and routing algorithms for wireless sensor networks: GA based approach”, *Wireless Personal Communications*, Vol. 83, pp. 2403-2423, 2015.
- [18] R. K. lingam and S. N. cheliyan, “Genetic algorithm based fault tolerant clustering in wireless sensor network”, *IET Communications*, Vol. 11, No. 12, pp. 1927-1932, 2017.
- [19] T. Bhatia, S. Kansal, S. Goel, and A. K. Verma, “A genetic algorithm based distance aware routing protocol for wireless sensor networks”, *Computers and Electrical Engineering*, Vol. 56, pp. 441-455, 2016.
- [20] D. Karaboga, B. Gorkemli, C. Ozturk, and N. Karaboga, “A comprehensive survey: artificial bee colony (ABC) algorithm and applications”, *Artificial Intelligence Review*, Vol. 42, pp. 21-57, 2014.
- [21] D. Karaboga, S. Okdem, and C. Ozturk, “Cluster based wireless sensor network routing using artificial bee colony algorithm”, *Wireless Networks*, Vol. 18, pp. 847-860, 2012.
- [22] J. R. Parvin, and C. V. Nayaki, “Particle swarm optimization-based clustering by preventing residual nodes in wireless sensor networks”, *IEEE Sensors Journal*, Vol. 15, No. 8, pp. 4264-4274, 2015.
- [23] A. Mohmmad and P. K. Jana, “PSO-based approach for energy-efficient and energy-balanced routing and clustering in wireless sensor networks”, *Soft Computing*, Vol. 21, pp. 6825-6839, 2017.
- [24] S. Bayraklı and S. Z. Erdogan, “Genetic algorithm based energy efficient clusters (GABEEC) in wireless sensor networks”, In: *Proc. of 3rd International Conf. on Ambient Systems, Networks, and technologies*, Niagara Falls, Ontario, Canada, pp. 247-254, 2012.
- [25] S. S. Kumar and S. Vishwas, “GDFEC protocol for heterogeneous wireless sensor network”, *Computational Intelligence in Data Mining*, Vol. 1, pp. 345-354, 2015.
- [26] S. P. Singh and S. C. Sharma, “Genetic Algorithm Based Energy efficient clustering (GAEEC) for Homogeneous Wireless Sensor Networks”, *IETE journal of Research*, Vol. 64, No. 5, pp. 1-12, 2018.
- [27] P. Parwekar and S. Rodda, “Optimization of Clustering in Wireless Sensor Networks Using Genetic Algorithm”, *International Journal of Applied Metaheuristic Computing*, Vol. 8, No. 4, pp. 332-343, 2017.
- [28] V. Pal, Yogitab, G. Singh, and R. P. Yadav, “Cluster Head Selection Optimization Based on Genetic Algorithm to Prolong Lifetime of Wireless Sensor Networks”, In: *Proc. of*

3rd International Conf. on Recent Trends in Computing, Ghaziabad, India, pp. 1417-1423, 2015.

- [29] X. Yuan, E. Mohamed, K. Hamdy, E. Minir, and A. M. Riad, "A genetic algorithm-based dynamic clustering method towards improved WSN longevity", *Journal of Network System and Management*, Vol. 25, pp. 21-46, 2017.
- [30] E. Mohamed, F. Ahmed, N. Zhou, M-M. Wang, S. Abdalla, and J. Batle, "Dynamic multi-hop clustering in a wireless sensor network: Performance improvement", *Wireless Personal Communications*, Vol. 95, pp. 1-21, 2017.
- [31] N. R. Amgoth, R. N. Bhukya, and N. D. Lavadya "Location Aware Relay Node Placement for Energy Constrained Wireless Sensor Networks", *International Journal of Advanced Trends in Computer Science and Engineering*, Vol. 8, No. 3, pp. 519-526, 2019.