JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 36, 777-794 (2020)
DOI: 10.6688/JISE.202007_36(4).0006

A NSGA-II Based Energy Efficient Routing Algorithm for Wireless Sensor Networks

MOHIT KUMAR[1], SONU MITTAL[2] AND MD. AMIR KHUSRU AKHTAR[3]
[1]*School of Computer and Systems Sciences*
[2]*Department of Computer Science and Engineering*
Jaipur National University
Jaipur, Rajasthan, 302017 India
[3]*Department of Computing and Information Technology*
Usha Martin University
Ranchi, Jharkhand, 835221 India
E-mail: [1]mohitsmailbox13@gmail.com; [2]sonum7772@rediffmail.com; [3]akru2008@gmail.com

Energy conservation has remained as one of the most challenging issue in design of WSNs. Over the years many effective clustering and routing algorithms have been proposed for handling this challenge. In this paper we propose NSGA-2 based routing algorithm for WSN. The routing algorithm has been developed by optimizing two parameters namely transmission distance and total number of hops which are conflicting in nature. We made some changes in NSGA-2, thus ensuring a better convergence rate compared to traditional NSGA-2. Theoretical analysis and grievous simulation demonstrate the effectiveness of the proposed algorithm. The obtained results were compared with several popular algorithms to validate the superiority of the proposed algorithm in terms of numerous performance metrics.

Keywords: wireless sensor network, clustering, routing, NSGA-2, transmission distance, hops

1. INTRODUCTION

Wireless sensor networks (WSNs), are one of the most important technologies for acquisition and processing of information and currently they are being used in numerous applications of different areas [1, 2]. The WSNs are equipped with large number of tiny sensor nodes which are deployed in target area for collecting and processing data and then sending them to the base station in either single hop or multi-communication mode. Since for many applications the nodes are deployed in an area where recharging or replacing the batteries is not possible. Therefore, for long run of WSNs their energy consumption should be minimized. Over the years reduction of energy consumption has remained the most pressing issue in WSNs. Many researchers [3-5] have addressed this issue by designing energy efficient clustering and routing algorithms. In several applications of WSNs, instead of choosing cluster heads (CHs) among normal sensor nodes, few high energy nodes named gateways are deployed which acts as cluster heads [6-10]. The gateways play very important role, as they are responsible for receiving and processing data and then sending them to the base station using single or multiple hops. It may be noted that the gateways also run on battery and hence they are also power constrained. Inappropriate formation of cluster and improper route selection by the gateways can lead to increased communication

Received September 3, 2019; revised October 29, 2019; accepted December 9, 2019.
Communicated by Xiaohong Jiang.

latency, and thus it can degrade the performance of WSNs drastically. In multi-hop routing model a gateway needs to forward the data to the base station (BS) through other gateways. Therefore, we can say that other gateways act as relay nodes. So, for minimizing the energy consumption in data routing, a gateway needs to select the best neighbor gateway (relay node). It is important to note that, for a sensor network consisting of m number of gateways, with each gateway having an average of k valid neighbors (one hop). For such network the total number of valid routes will be k^m. Applying brute force approach for large scale WSN will yield very high computational complexity. As a matter of facts selection of route in WSNs is a NP-hard problem. Use of heuristic algorithms for solving such problems with large solution space is always advantageous. Battery life of the gateways can be improved by reducing the transmission distance between sender and receiver. Many researchers [11-13, 16] have solved the multi-objective optimization problem by combining them into single objective scalar function. This approach is known as weighted-sum approach. This approach has few important drawbacks like difficulty of setting suitable value of weighting factors. Beside that even distribution of the weights among objective function does not always result in an even distribution of solutions. In those algorithms, the weighting value assigned to objective functions is chosen in trial and error fashion and it does not give any guarantee that the obtained value using weighting factor is optimum. The proposed algorithm provides a set of solutions, by optimizing both the conflicting objectives. These solutions are known as pareto optimal solutions. In this paper, we have proposed NSGA-2 [14] based routing algorithm. We have optimized two parameters that are conflicting in nature. The proposed algorithm ensures that the routing is energy balanced. The main contribution of this paper is listed as follows.

- Generation of restricted initial population – the generation of chromosomes is restricted by taking into accounts connectivity between the gateways. This is contrary to the generation of randomized chromosomes as used in traditional NSGA-2.
- Energy balanced mutation operation-During mutation phase, the mutation point is chosen in such a manner, that the newly generated child chromosomes ensure better energy balanced routing. This is again in contrary to the traditional NSGA-2 in which mutation point was chosen in random way.
- The proposed routing algorithm finds out the best route from all the gateways to the base station by maximizing the remaining energy of each gateway also keeping trade-off between conflicting objectives.

The above strategies make our algorithm faster than traditional NSGA-2. Our proposed algorithm is both energy efficient and energy balanced. We performed grievous simulation and compared the obtained result with other popular algorithms. The comparison clearly exhibits the superiority of the proposed algorithm as compared to other algorithms. The remaining paper is organized as follows: The related work is discussed in Section 2. System model and Cluster formation is discussed in Sections 3 and 4 respectively. The proposed routing algorithm is presented in Section 5. Result analysis is done in Section 6. Section 7 concludes the paper.

2. RELATED WORK

In recent years a wide number of algorithms were developed addressing clustering and routing issues of WSN. In this section we have received number of papers which are related to our proposed routing algorithm. Selection of optimal route is an optimization problem that is N-P hard in nature. Numerous heuristic and meta-heuristic algorithm [15, 17-21] have been applied addressing this problem. A popular technique named LEACH [22] uses a distributed algorithm to form clusters. For balancing the load amongst the sensor nodes it dynamically rotates the load of the Cluster Head to the nodes. But the main disadvantage with this approach is that it may select a node as CH with low energy which may eventually die quickly. Also, in this approach BS receives the packet from the CH via single hop communication which is not a practical scenario for WSNs that may have large coverage area. Hence, this lead to development of more improved algorithms over LEACH such as PEGASIS [23], HEED [24] *etc*.. In comparison to LEACH, PEGASIS promotes network lifetime. PEGASIS uses a method of chain formation with nodes such that each node communicates with the neighbor and only a single node selected as group head will transmit data to the BS. But this approach is also unsuitable for large networks as it constantly requires adjusting the topology and the data delay is also high. Bari *et al.* [21] proposed an algorithm based on GA for a two-tire WSN in which data routing is done using relay nodes. Roulette-wheel selection method is used for selection of nodes and the fitness function is defined using network lifetime in terms of rounds. Mutation operation is carried out by critical node selection from the relay nodes, which during data transmission dissipated the maximum energy. Mutation is carried out by replacement of node in the next hop of the critical node by the new next-hop relay node or else by the diversion of incoming flow towards the critical node to any other relay node. SK Gupta *et al.* [15] proposed routing algorithm based on GA called GAR. The main focus of this algorithm was to minimize the overall communication distance between the gateways and the BS. In this algorithm tournament selection was used in contrast to Roulette-wheel selection which was used by Bari *et al.* Moreover, these two algorithms only focused on routing of data from the gateways to the BS. But they did not consider the communication between the sensor nodes and their respective gateways. Kuila *et al.* [6] proposed GA based load balanced clustering algorithm for WSNs. In this algorithm cluster formation in done in such a way that the maximum load of each of the gateways is minimized this in turn works for equal and unequal load on nodes. This algorithm has better load balancing and faster convergence than the traditional GA. But on the other hand, it does not take into consideration the residual energy of the nodes. Chiang *et al.* [25] proposed minimum hop routing model [MHRM]. In this algorithm each gateway constructs a path to the BS in such manner that the hop count is minimized so in this model a gateway selects the relay node for data forwarding which is farthest from it. This leads to much more energy dissipation while transmission. Yessad *et al.* [27] proposed a multipath routing protocol for homogenous WSN for balancing the energy consumption of nodes. The path selection is made on the probability of residual energy, communication energy cost and number of paths. But this algorithm is unsuitable in case of a large scale of networks. Also, few routing algorithms have been proposed which take into consideration the fault tolerance issues (Djukic and Valaee *et al.* [27], Intanagonwiwat *et al.* [28]). Direct diffusion routing protocol was the most popular among them which was proposed by Intanagonwiwat. Direct Diffusion protocol uses query driven data delivery. In DD, between source node and BS, multiple node disjoint paths are created. However, it is to be noted that DD is not suitable for those applications which require continuous data delivery as it is query driven. Song *et al.* [33] proposed a pervasive modeling scheme that consisted of features like smart routing and load balancing. But, this model was developed for space and terrestrial networks. Liu *et al.* [29] has addressed the hot spot problem by proposing a cluster based routing protocol in order to maximize the networks lifetime. However, this implementation was done solely for vehicular sensor networks. Song *et al.* [32] proposed dynamic routing by leveraging the advantages of Moving Target Defense. But, the main focus of this paper was privacy enhancement using smart collaborative distribution scheme. Mohamed *et al.* [31] proposed two protocols CZSEP and CZSEP-HN to overcome the problem of large transmission distance between normal nodes and base station (BS) and between cluster heads and BS. But, they did not minimize the number of intermediate nodes in a path from gateway to the BS. Ok *et al.* [30] focused on balancing the energy consumption of the network and presented a distributed energy balance routing scheme (DEBR). But the main concern with DEBR is that the algorithm may select a CH in the next hop for energy balancing which might not have any other CHs in its communication range for transmitting the data. Due to which some of the important data might not be able to reach the sink. Recently Md. Azharuddin *et al.* [13] have proposed a PSO based energy balanced routing algorithm. In this algorithm they have used the weighted sum approach. Although this is a popular approach but the major drawback with this approach is even distribution of the weights among objective functions does not always result in even distribution of solution.

3. SYSTEM MODEL

In our WSN model we assume that the deployment of sensor nodes in a target area is done in random manner and they become stationary after the deployment. The sensor node can be assigned to only those gateways which are falling under its communication range. Each sensor node can be assigned to only one gateway. It is the responsibility of the sensor nodes to sense data in their respective regions and sends it to the gateways to which they are assigned to. The gateways process the data by aggregating them and removing the redundant data. The gateways then send the data to the base station via a single hop or through multiple hops. Apart from processing the data the gateways also act as relay nodes in data forwarding. All the communications are done over wireless links. In the proposed algorithm we have used following terminologies.

Preliminaries

In this section, we discuss about the network model, various assumptions and terminologies used. This will give a better understanding of the proposed algorithms.

3.1 Energy Model

In this paper, the radio energy model for calculating energy dissipation is used [22]. Here the energy dissipated by the transmitter is used to run the power amplifier and the radio electronics and the energy dissipated by the receiver is used to run the radio electronics. The energy consumed by the sensor nodes relies upon the amount of data and the distance which it needs to travel. In this energy model, energy dissipated by the sensor nodes is proportional to (d^2) when the propagation distance (d) is less compared to threshold distance (d_0) or else it is proportional to (d^4). The same is given in the equation below.

$$E_{TX}(l,d)=\begin{cases} l\times E_{elec}+l\times \varepsilon_{fs}\times d^2, \text{ for } d<d_0 \\ l\times E_{elec}+l\times \varepsilon_{mp}\times d^4, \text{ for } d\geq d_0 \end{cases} \tag{1}$$

$$E_{RX}(l,d)=l\times E_{elec} \tag{2}$$

Where E_{TX} is the amount of energy utilized in transmitting any message of length l bit over the distance d, E_{RX} is the amount of energy required in receiving the message, ε_{fs} is the free space model of the transmitter amplifier and ε_{mp} is the multi-path model of transmitter amplifier. Please note that E_{elec} depends on multiple factors which includes modulation, spreading of signals, filtering and digital encoding. Also, in the model represented here both the multi-path and free space models were used, which solely depends on distance between receiver and transmitter.

3.2 Basic Terminologies

For the ease of understanding in the proposed algorithm we have used following terminologies.

1. The set of sensor nodes is denoted by $S=\{S_1, S_2, S_3, \ldots, S_N\}$.
2. The set of gateways is denoted by ¥ = $\{g_1, g_2, g_3, \ldots, g_M\}$ and g_{M+1} indicates the base station (BS), $N>M$.
3. $E_{res}(g_j)$ denotes the residual energy of gateway g_j.
4. dist(s_i, g_j) denotes the Euclidian distance between sensor node s_i and the gateway g_j.
5. node_degree(g_j) denotes the total number of sensor nodes assigned to gateway g_j.
6. dist(g_j, BS) denotes the Euclidian distance between gateway g_j and the base station (BS).
7. comm(s_i) denotes the set of gateways which are falling under maximum communication range(C_s) of sensor node s_i.

$$\text{comm}(s_i)=\{g_j|\text{dist}(s_i, g_j)\leq C_s \wedge \forall g_j\in ¥\} \tag{3}$$

8. comm(g_j) denotes the batch of gateways which are within the communication range of g_i. Here d_{max} denotes the maximum communication range of the gateways. Hence, distance between gateways, dist(g_i, g_j) will always be less than or equal to d_{max}. In other words,

$$\text{comm}(g_j)=\{g_j|\forall g_j\in(¥+g_{M+1})\wedge \text{dist}(g_i, g_j)\leq d_{max}\}. \tag{4}$$

9. Maximum Hop Count (Max_Hop) of the gateways is defined as the maximum number of hops that can be taken for reaching the BS from any gateway. Hence,

$$\text{Max\_Hop}=\{\text{Max}(\text{Hop\_Count}(g_j)|\forall i, 1\leq i\leq M, g_i\in ¥\}. \tag{5}$$

10. Hop_Count(g_j) denotes the total number of next hops required to reach the base station (BS) from gateway g_i in data routing phase. If the communication between the gateway

g_i and BS is direct then Hop_Count(g_j) is one. Therefore Hop_Count(g_j) can be defined as:

$$\text{Hop\_Count}(g_i) = \begin{cases} 1 & \text{Nxt\_hop}(g_i) = g_{M+1} \\ 1 + \text{Hop\_Count}(g_i) & \text{Nxt\_hop}(g_i) = g_j \end{cases}. \tag{6}$$

11. Nxt_Hop(g_i) denotes the batch of gateways that might be selected for data forwarding by the gateway g_i. The next hop relay nodes should be towards the base station (BS) Hence,

$$\text{Nxt\_Hop}\ (g_i) = \{g_j | \forall g_j\ \text{comm}(g_i) \wedge \text{dist}\ (g_j, g_{M+1}) \leq \text{dist}\ (g_i, g_{M+1})\}. \tag{7}$$

12. Max_dist denotes the maximum distance between two gateways in the routing path. It can be defined by

$$\text{Max\_dist} = \text{Max}\{\text{dist}(g_i, \text{Nxt\_hop}(g_i) | \forall i, 1 \leq i \leq M, g_j \in ¥\}. \tag{8}$$

4. CLUSTER FORMATION

In the proposed algorithm, we have used weight function for the formation of cluster. The weight function is built by making use of local information, and by using this function every node s_i will be assigned to their corresponding gateways. The weight function will depend upon following parameters:

1. Gateway residual energy: A sensor node s_i should join that gateway which is having highest residual energy among all the gateways which are falling under its communication range. Hence,

$$\text{GS\_Weight}\ (s_i, g_j) \propto E_{res}(g_j). \tag{9}$$

2. Distance of gateway from sensor node: A sensor node dissipates maximum energy during transmission. The shorter distance between sensor node and gateway ensures lesser energy consumption.

$$\text{GS\_Weight}(s_i, g_j) \propto \frac{1}{dist(s_i, g_j)} \tag{10}$$

3. Gateway node degree: A sensor node s_i should join gateway g_j that has lowest node degree compared to all the gateways which are falling under its communication range.

$$\text{GS\_Weight}(s_i, g_j) \propto \frac{1}{node\ degree(g_j)} \tag{11}$$

4. Distance from the gateway to the Base Station: The gateways which are nearer to the sink are overburdened because they are responsible for data forwarding to the BS. Therefore, their energy diminishes faster as compared to other gateways. So, they should

be lightly loaded. Thus, the selection weight of gateways which are far away from the base station should be greater as compared to gateways which are nearer.

$$\text{GS\_Weight}(s_i, g_j) \propto \text{dist}(g_j, \text{BS}) \tag{12}$$

Combining above equations we get,

$$\text{GS\_Weight}(s_i, g_j) \propto \frac{E_{res}(g_j) \times dist(g_j, BS)}{node\ degree(g_j) \times dist(s_i, g_j)},$$

$$\text{GS\_Weight}(s_i, g_j) = P \times \frac{E_{res}(g_j) \times dist(g_j, BS)}{node\ degree(g_j) \times dist(s_i, g_j)}. \tag{13}$$

Where P is proportionality constant. We assume that $P = 1$ (without loss of generality). For the formation of cluster, for each sensor node GS_Weight is calculated using Eq. (13) and the sensor node joins gateway which is having the highest value.

Cluster Formation Algorithm

Input:

(1) A set of sensor nodes $S = \{s_1, s_2, \ldots, s_n\}$ and the set of gateways ¥ = $\{g_1, g_2, \ldots, g_m\}$
(2) d_{ij} = Euclidian distance (s_i, g_j) where $g_j \in \text{Comm}(s_i)$
(3) $E_{Res}(g_j)$: Residual energy of gateways
(4) B_{set} and U_{set}: Bounded set and Unbounded set

Output: An Assignment A: $s \rightarrow g$ such that overall energy consumption of the sensor node is minimized.

1. Sort the sensor nodes in increasing order in the basis of number of possible gateways to which they can be assigned to.
2. While ($B_{set} \neq$ Null)
   2.1 Remove s_i from S and then assign it to corresponding gateway such that $g_j \in \text{Comm}(s_i)$
   2.2 Delete S_i from B_{set} as well as S
   End While
3. Sort U_{set} in the basis of cardinality in non-decreasing order.
4. While ($U_{set} \neq$ Null)
   4.1 Select the first sensor node from U_{set} say s_i
   4.2 Designate s_i to g_j such that $\text{GS\_Weight}(s_i, g_j) = \text{MAX}\{\text{GS\_Weight}(s_i, g_j) \mid \forall g_j \in \text{Comm}(s_i)\}$
   4.3 Take out s_i from U_{set} and S
   End While

Stop

Fig. 1. Cluster formation algorithm.

5. PROPOSED ROUTING ALGORITHM

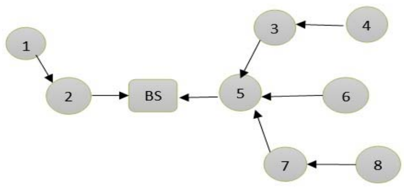
We now present our NSGA-2 based routing algorithm. In the upcoming section chromosome representation, initialization of population, fitness function, non-dominated sorting along with crossover and mutation are discussed in detail.

5.1 Chromosome Representation

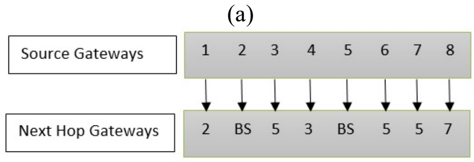
In the routing phase, every gateway is supposed to send the aggregated data to the base station via a single hop or multiple hops. Chromosomes are represented as sequence of gateways, indicating the next hop gateway chosen by the set of gateways for data forwarding. Considering a chromosome if its '*i*th' gene value is assumed to be j, then it will imply that the gateway g_i has selected gateway g_j for the next hop. Also, the number of gateways is same as the length of each chromosome. It may be noted that the same value j for any gene position can be repeated. This is because two different gateways might choose the same gateway for next hop.

5.2 Initial Population Generation

A set of chromosomes that are generated randomly is known as initial population. Each and every chromosome is a series of gateways that corresponds to a valid routing path. The generation of valid chromosome is done in such a way that the value of '*i*th' is say '*j*' and is selected randomly such that g_j belongs to Next_hop(g_i). It may be noted that the NSGA-2 approach for initial population generation does not operate on any particular algorithm or in other words initial population is generated randomly. The major drawback of simple NSGA-2 approach is that, it generates large number of invalid chromosomes. In our approach, gateway g_j selects next hop gateway from only among those gateways that falls within its communication range. This ensures that no invalid chromosome is generated and makes the proposed algorithm converge faster compared to traditional NSGA-2. Fig. 2 (a) shows directed acyclic graph $G(V, E)$ where edge $g_i \rightarrow g_j$ represents g_j is chosen as next hop gateway by g_i for data forwarding. Fig. 2 (b) shows representation of chromosome. Each valid chromosome must contain valid routing path *i.e.* every gateway must contain one or more gateways that forwards the aggregated data directly to the base station. Fig. 3 shows two chromosomes that are generated from Table 1. During the generation of initial population, the algorithm does not try to minimize the energy con-sumption of the gateways nor it tries to find the best route to the base station.



(a)



(b)

Fig. 2. (a) A sub-graph of WSN; (b) Chromosome representation; Circle Gateway, square base station, arrow direction of flow of data.

Gateway Number	1	2	3	4	5	6	7	8
Chromosome 1	2	BS	5	3	BS	5	5	6
Chromosome 2	2	BS	5	6	BS	5	5	7

Fig. 3. Chromosome generated using Table 1.

Table 1. Gateways with possible next hop gateways.

Gateways	Next hop possible gateways
g_1	$\{g_2\}$
g_2	{BS}
g_3	$\{g_5\}$
g_4	$\{g_3, g_6\}$
g_5	{BS}
g_6	$\{g_3, g_5\}$
g_7	$\{g_5\}$
g_8	$\{g_6, g_7\}$

5.3 Derivation of Fitness Function

Each chromosome is evaluated based on some fitness functions. For building fitness function we have taken two objectives into account. The first one being the minimization of maximum number of hops/forwards that is used by the gateways for forwarding data from source gateway to the base station. The second objective is minimization of maximum distance between two gateways. Therefore, the two objectives are as follows:

Objective 1: Minimize Max_hop = Max{Total no of hops $(g_i)|\forall i, 1 \leq i \leq m\}$ (14)

Objective 2: Minimize Max_dist = Max{Dist$(g_i, g_j)|g_j \in$comm$(g_i), \forall i, 1 \leq i \leq m$ and $i \neq j\}$ (15)

It is important to note that the above objectives are conflicting in nature. In this paper multi-objective optimization is demonstrated using NSGA-2. The algorithm produces a set of solutions by optimizing the two conflicting objectives. The set of compromised solutions are known as pareto optimal solutions.

5.4 Non-dominated Sorting and Selection

The initial generated population is evaluated based on all objective functions. The chromosomes are then sorted based on non-domination approach. The obtained solution is classified into different fronts. Fitness (rank) is assigned to each non-dominated solution (one for the best front, two for the next best and so on). During selection process if two solutions have same non-domination rank then, crowding distance sorting is applied on them for judging which solution is less crowded. Crowding distance sorting is a strategy in which a solution residing on least crowded region on the front is chosen over other solutions. This ensures the diversity in set of solutions. The selection mechanism determines which of the chromosomes will mate to produce child chromosome. Thus, it plays very important role in improving quality of population. We have used roulette wheel selection method for the selection of chromosome with best fitness value from the current population. The selection operation is followed by crossover and mutation.

5.5 Crossover

Chromosome operation is used to generate new offspring's from a set of selected parents. Here, we have used 1-point crossover operation. In order to ensure that each and every newly generated chromosomes are valid, we have applied one constraint that the value of 'j' of the 'ith' gene will be such that $g_j \in$Next_hop(g_i) where $i \neq j$. Hence for each gateway the next hop gateway is valid. Fig. 4 shows the crossover operation in which the down arrow indicates the crossover point.

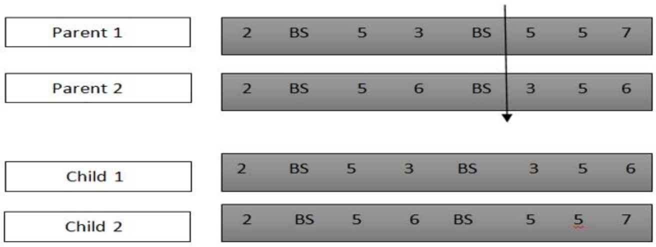


Fig. 4. Crossover.

5.5.1 Lemma: The crossover operation produces valid child chromosomes.

***Proof*:** A valid chromosome is one which consists of a path from every gateway to the base station without forming any cycle.

In Section 5.2 it is mentioned that the value 'j' of the 'ith' gene will be such that $g_j \in$ Next_hop(g_i) where $i \neq j$. Hence every selected next hop relay node will be valid. In the proposed algorithm the value of gene position after a point are interchanged with each other. Since, all the parents are valid hence the newly generated chromosomes will also be valid.

5.6 Mutation

The beauty of the proposed algorithm lies in the fact that the mutation is applied at specific gene location unlike normal NSGA-2 algorithm. The main objective of this algorithm is to balance the energy consumption of the gateways by ensuring that the traffic passing through the gateways is evenly distributed across all gateways. The idea is illustrated with example in the figures below. Fig. 5 (a). represents the chromosome representation. It is shown that the maximum traffic is flowing through gateway g_3. Therefore, 3 are appearing at multiple gene locations. In the proposed mutation operation, out of three identified gene locations, one location is chosen and is replaced by another gene value. The chosen location must have alternate next hop option. It is important to note that the new gateway which replaces the old one must fall within the communication range of gateway whose next hop was changed. Fig. 5 (b) represents the chromosome after the mutation operation.

Example 1: Consider the chromosome representation shown in Fig. 5 (a). It is shown that the gateway g_3 is chosen as next hop by three gateways. Hence, g_3 is appearing at three different gene locations. In the proposed method, during mutation operation out of the three gene location occupied by gateway g_3, one is randomly chosen and is replaced by another gateway. It is important to note that the new gateway which will replace the older gateway must be from next hop gateway list of the corresponding gateway. Gateway g_3 is occupying three gene locations 1, 4 and 8. Out of these three positions one is randomly selected. In the shown Fig. 5 (b) gene location 4 is randomly selected and at that position g_3 is replaced by g_5 (refer Table 1).

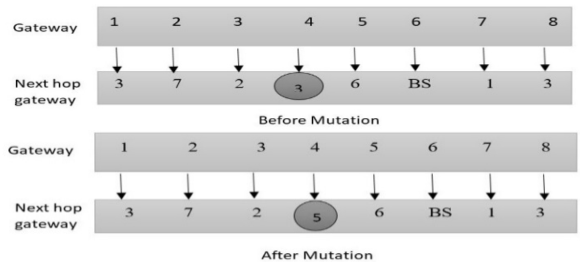


Fig. 5. Representation of chromosome before and after the mutation operation.

5.6.1 Lemma

Lemma: The chromosome generated by the mentioned above process of mutation is valid.

Proof: As per the chromosome representation, the gateway with maximum traffic is identified and is replaced by another gateway with lesser traffic. During mutation process the identified gateway is appearing at more than one gene location. Out of these identified locations, one is randomly chosen and replaced. Also, it is taken care that the new gateway which replaced the older one always falls under the communication range. It is important to note that the generated offspring's by the crossover operations are all valid and the validity of these offspring's was not hampered in the mutation process by the new gateway replacement.

5.6.2 Remark

It is important to note that the mutation strategy used in our approach makes our algorithm to converge faster as compared to traditional NSGA-2. The reason for this is, in traditional NSGA-2 the point of mutation is chosen randomly which may lead to generation of poor quality chromosome and hence its convergence rate is slower.

5.7 Elitism

In this strategy, the best chromosomes obtained are carried forward in the next generation. This ensures the quality of result obtained does not get degraded during the process of evolution.

NSGA-2 based Routing Algorithm

1. Set $G = 0$, $S =$ Pop_Size
2. Generate initial population /* As described in Section 5.2 */
3. $P_f =$ Calculate fitness value of generated population /* Using Eqs. (14) and (15) */
4. Repeat
5. $Q_f =$ Create offspring population from P_f by applying 1-point crossover and mutation operation. /* Refer Sections 5.5 and 5.6 */
6. Assign fitness value to each newly generated chromosome
7. $M_f = Q_f \cup P_f$ Create mating pool by combining parent and offspring population.
8. $F =$ Apply non-dominated sorting (M_f) and put all the obtained solutions in identified fronts $\{F_1, F_2, F_3, \ldots, F_n\}$.
9. Calculate the crowding distance of each solution.
10. Perform Selection (Roulette Wheel) in the basis of rank. If chosen solutions are having same rank then comparison is made in the basis of crowding distance. The criteria of selection are lower rank and higher crowding distance.
11. $G = G + 1$. Increment the Generation Count.
12. Until $G <$ Max_Gen.

Fig. 6. NSGA-2 based routing algorithm.

6. PERFORMANCE EVALUATION

6.1 Simulation Setup

We performed simulation using MATLAB R2012b and C programming language. The parametric values used in performing simulation are shown in Table 2. In the simulation run different WSN scenarios were created by varying the number of sensor nodes and gateways ranging from 200-600 and 30-70 respectively. The experiment results were obtained by considering the clustering and routing in a combined manner. But, in the comparisons made in the basis of number of hops and distance covered in round only routing is considered.

Table 2. General parameters with values for simulation.

Name of Parameter	Values
Target Area	400m × 400m
Number of sensor nodes	200-600
Number of Gateways	30-70
Initial Energy (sensor nodes)	2J
Initial Energy (Gateways)	10J
Communication Range (sensor nodes)	80m
Communication Range (Gateways)	120m
E_{elec}	50nJ/bit
E_{da}	5nJ/bit
$\in_{fs}$	10pJ/bit/m^2
$\in_{mp}$	0.0013pJ/bit/m^4

d_0	87.7m
Size of Packet	4000bits
Size of Message	200bits
Population Size	20
Generations	100
Crossing Probability	0.5
Mutation Probability	0.5
Selection of Parents	Roulette wheel
New Generation Selection	Elitist

6.2 Simulation Result

For the sake of comparison, various other popular algorithms were simulated. The results obtained were compared with PSO based energy balanced clustering and routing algorithm developed by Azhar *et al.*, another PSO based approach given by Kuila *et al.*, a GA based algorithm proposed by Kuila *et al.* and LDC by Bari *et al.*

In view of comparison we have used following performance metrics.

1. Network Lifetime: The lifetime of a network for WSNs can be defined in numerous ways. In our simulation we have defined the network lifetime as total number of rounds till the first gateway depletes its energy fully and dies. In order to increase the network life-time the depletion of energy among the gateways should be balanced. The proposed algorithm builds a trade-off between inter-cluster distance and total number of hops. In the simulation the number of sensor nodes were varied from 200-800 and the number of gateways used were 50. The comparison of the proposed algorithm with other existing algorithms with respect to network lifetime can be seen in Fig. 7. The lifetime of the proposed algorithm is better than LDC, GALBCA and PSO based algorithms presented by Kuila and Azhar respectively.

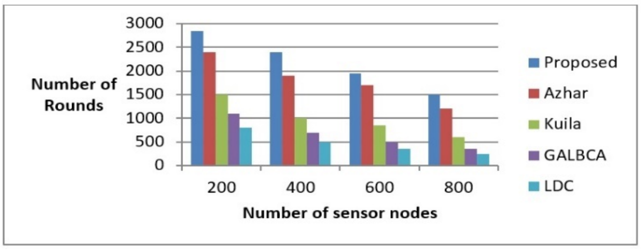


Fig. 7. Comparison in terms of network lifetime.

2. Difference between First Gateway Die (FGD) and Last Gateway Die (LGD): It is calculated in terms of total number of rounds. Fig. 8 shows comparison of results in the basis of difference in death of first gateway and the last gateway. It is important to note that lower duration indicates better energy balancing of the gateways. In the simulation the number of sensor nodes used were varied from 200-600 and the number of gateways used were 80. Fig. 8 demonstrates the superiority of the proposed algorithm in terms of energy balancing.

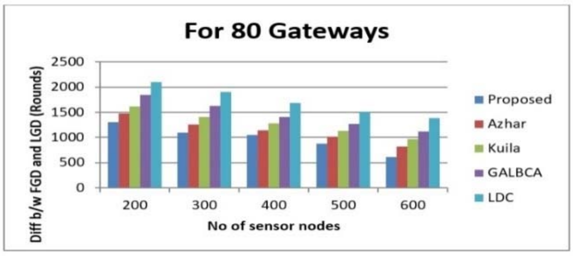


Fig. 8. Comparison in terms of difference between death of first gateway and the last gateway.

3. Energy Consumption: For measuring energy consumption we ran the proposed algorithm with 500 sensor nodes and 80 gateways. In this paper we have mainly focused on the routing problem of WSNs. Therefore, the energy consumed by the gateways in routing is only considered and the comparison with other algorithms is made considering this factor. Energy consumption here can be defined as energy consumed by the gateways in forwarding data from itself to the base station via a single hop or multiple hops. Fig. 10 shows the comparison between proposed algorithm and some existing algorithms. Since, the proposed algorithm operates for more number of rounds it's energy consumption becomes greater than other algorithms at later stages.

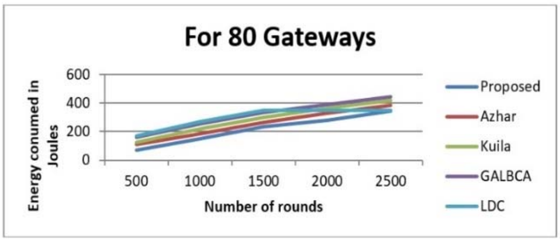


Fig. 9. Comparisons in terms of energy consumption.

4. Energy imbalance factor: For judging quality of energy balancing, we calculate the standard deviation of the energy consumed by the gateways (for the purpose of routing only) and plot against number of rounds. It is calculated using the given Eq. (16)

$$\text{EIF} = \tfrac{1}{m}\sqrt{\sum_{i=1}^{m}(E_{avg} - E_{con}(i))^2}. \tag{16}$$

Where m is total number of alive gateways, E_{avg} is the average residual energy consumption of the gateways, $E_{con}(i)$ denotes the energy consumed by gateway 'i' in the present round. Fig. 10 shows the proposed algorithm is more balanced as compared to other existing algorithms.

5. Comparative analysis based on total number of hops taken and distance covered in rounds: For next comparison we executed two more routing algorithms *i.e.* GAR [15] and MHRM [25]. Fig. 11 shows the comparison between different algorithms on the basis of total number of hops taken. MHRM takes minimum number of hops because of its next hop selection criteria. Next hop in MHRM is selected based on maximum

possible distance. The network lifetime is very low in MHRM as compared to other algorithms. This is due to long distance transmission which in turn increases the energy consumption. On the other hand GAR, concentrates on reduction of transmission distance but does not focus on reducing number of hops. Since GAR optimizes single objective *i.e.* distance covered in a round, it gives slightly better result as compared to our algorithm. The proposed algorithm is multi-objective and it builds a tradeoff between the mentioned above factors. Figs. 11 and 12 show the comparison in terms of total number of hops and total distance covered respectively.

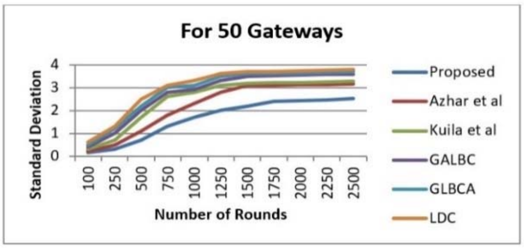


Fig. 10. Comparison in terms of energy balancing.

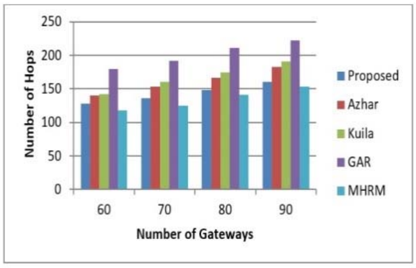


Fig. 11. Comparison in terms of total number of hops.

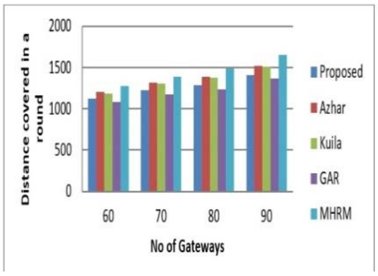


Fig. 12. Comparison in terms of total distance covered in a round.

7. CONCLUSION

This paper addresses the most important design issues in WSN. An energy efficient routing algorithm based on non-dominated sorting genetic algorithm (NSGA-2) is presented in this paper. In the proposed algorithm we have optimized two parameters which are conflicting in nature. In order to make our routing algorithm energy balanced we have made some changes in traditional NSGA-2. Furthermore, the problem of calculation of suitable weight values allocated to each objective function is eliminated completely. In the proposed algorithm the energy consumed by the gateways are significantly balanced, thus improving the lifetime of the network. The proposed algorithm converges faster as compared to traditional NSGA-2. For cluster formation we have derived weight function based on several parameters. We have performed grievous simulation and the results obtained are very encouraging. The proposed algorithm outperforms several existing algorithms in terms of network lifetime, rate of convergence, power imbalance factor *etc.* We have not addressed the fault tolerance issue in the proposed algorithm and our next endeavor will be design of clustering and routing algorithm with fault tolerance for dynamically changing WSN.

REFERENCES

1. I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, Vol. 38, 2002, pp. 393-422.
2. J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, Vol. 52, 2008, pp. 2292-2330.
3. A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications*, Vol. 30, 2007, pp. 2826-2841.
4. K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Networks*, Vol. 3, 2005, pp. 325-349.
5. M. Kumar and D. Kumar, "Path cost based load and energy balanced clustering & routing algorithm," *International Journal of Applied Engineering Research*, Vol. 12, 2017, pp. 5690-5696.
6. P. Kuila, S. K. Gupta, and P. K. Jana, "A novel evolutionary approach for load balanced clustering problem for wireless sensor networks," *Swarm and Evolutionary Computation*, Vol. 12, 2013, pp. 48-56.
7. C. P. Low, C. Fang, J. M. Ng, and Y.-H. Ang, "Efficient load-balanced clustering algorithms for wireless sensor networks," *Computer Communications*, Vol. 31, 2008, pp. 750-759.
8. A. Bari, *et al.*, "A genetic algorithm based approach for energy efficient routing in two-tiered sensor networks," *Ad Hoc Networks*, Vol. 7, 2009, pp. 665-676.
9. P. Kuila and P. K. Jana, "A novel differential evolution based clustering algorithm for wireless sensor networks," *Applied Soft Computing*, Vol. 25, 2014, pp. 414-425.
10. P. Kuila and P. K. Jana, "Heap and parameter-based load balanced clustering algorithms for wireless sensor networks," *International Journal of Communication Networks and Distributed Systems*, Vol. 14, 2015, pp. 413-432.
11. P. C. Rao, P. K. Jana, and H. Banka, "A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks," *Wireless Networks*, Vol. 23, 2017, pp. 2005-2020.
12. P. Kuila and P. K. Jana, "Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach," *Engineering Applications of Artificial Intelligence*, Vol. 33, 2014, pp. 127-140.
13. Md. Azharuddin and P. K. Jana, "PSO-based approach for energy-efficient and energy-balanced routing and clustering in wireless sensor networks," *Soft Computing*, Vol. 21, 2017, pp. 6825-6839.
14. K. Deb, *et al.*, "A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II," in *Proceedings of International Conference on Parallel Problem Solving from Nature*, 2000, pp. 849-858.
15. S. K. Gupta, P. Kuila, and P. K. Jana, "GAR: an energy efficient GA-based routing for wireless sensor networks," in *Proceedings of International Conference on Distributed Computing and Internet Technology*, 2013, pp. 267-277.

16. 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, 2015, pp. 2403-2423.
17. S. M. Guru, S. K. Halgamuge, and S. Fernando, "Particle swarm optimisers for cluster formation in wireless sensor networks," in *Proceedings of IEEE International Conference on Intelligent Sensors*, *Sensor Networks and Information Processing*, 2005, pp. 319-324.
18. N. M. A. Latiff, C. C. Tsimenidis, and B. S. Sharif, "Energy-aware clustering for wireless sensor networks using particle swarm optimization," in *Proceedings of IEEE 18th International Symposium on Personal*, *Indoor and Mobile Radio Communications*, 2007, pp. 1-5.
19. P. C. Rao and H. Banka, "Energy efficient clustering algorithms for wireless sensor networks: novel chemical reaction optimization approach," *Wireless Networks*, Vol. 23, 2017, pp. 433-452.
20. E. A. Khalil and A. A. Bara'a, "Energy-aware evolutionary routing protocol for dynamic clustering of wireless sensor networks," *Swarm and Evolutionary Computation*, Vol. 1, 2011, pp. 195-203.
21. A. Bari, *et al.*, "A genetic algorithm based approach for energy efficient routing in two-tiered sensor networks," *Ad Hoc Networks*, Vol. 7, 2009, pp. 665-676.
22. W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, Vol. 1, 2002, pp. 660-670.
23. S. Lindsey and C. S. Raghavendra, "PEGASIS: Power-efficient gathering in sensor information systems," in *Proceedings of IEEE Aerospace Conference*, Vol. 3, 2002, pp. 1125-1130.
24. O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, Vol. 4, 2004, pp. 366-379.
25. S.-S. Chiang, C.-H. Huang, and K.-C. Chang, "A minimum hop routing protocol for home security systems using wireless sensor networks," *IEEE Transactions on Consumer Electronics*, Vol. 53, 2007, pp. 1483-1489.
26. S. Yessad, *et al.*, "Balanced energy efficient routing protocol for WSN," in *Proceedings of IEEE International Conference on Communications and Information Technology*, 2012, pp. 326-330.
27. P. Djukic and S. Valaee, "Reliable packet transmissions in multipath routed wireless networks," *IEEE Transactions on Mobile Computing*, Vol. 5, 2006, pp. 548-559.
28. C. Intanagonwiwat, *et al.*, "Directed diffusion for wireless sensor networking," *IEEE /ACM Transactions on Networking*, Vol. 11, 2003, pp. 2-16.
29. Y. Liu, *et al.*, "Multi-layer clustering routing algorithm for wireless vehicular sensor networks," *IET Communications*, Vol. 4, 2010, pp. 810-816.
30. C.-S. Ok, *et al.*, "Distributed energy balanced routing for wireless sensor network," *Computers & Industrial Engineering*, Vol. 57, 2009, pp. 125-135.
31. F. A. Mohamed, *et al.*, "Energy-efficient circle zones stable election protocol with helper nodes for heterogeneous WSNs," *IET Wireless Sensor Systems*, Vol. 9, 2019, pp. 313-322.

32. F. Song, *et al.*, "Smart collaborative distribution for privacy enhancement in moving target defense," *Information Sciences*, Vol. 479, 2019, pp. 593-606.
33. F. Song, *et al.*, "Modeling space-terrestrial integrated networks with smart collaborative theory," *IEEE Network*, Vol. 33, 2019, pp. 51-57.

Mohit Kumar is an Assistant Professor in the Department of Computer Science and Engineering, Cambridge Institute of Technology, Ranchi, Jharkhand. His research interest includes wireless sensor networks, ad-hoc networks and network security. He is currently pursuing Ph.D. from School of Computer and Systems Sciences at Jaipur National University, Jaipur, Rajasthan, India. He received his M-tech from Jaypee University of Information Technology, Soaln, H.P, India. He obtained his B.E from Pune University, Pune, Maharashtra, India.

Sonu Mittal is working as an Associate Professor in Computer Science and Engineering Department of Main Campus, Jaipur National University, Jaipur. He obtained his M.Tech. (Computer Science) in 2007 and Ph.D. (Computer Science and Engineering) in 2015. Dr. Mittal has authored/co-authored a book on computer networks and published more than 28 papers in international journals and conferences. He has chaired 5 International conferences and has been actively engaged as reviewer for various national and international forums. His major areas of research interest are software engineering, computer networks and machine learning.

Mohammad Amir Khusru Akhtar is an Associate Professor in the Faculty of Computing and Information Technology, Usha Martin University, Ranchi, Jharkhand, India. His research interest includes mobile ad-hoc network, natural language processing, deep learning and feature engineering. He is the author of over 20 peer-reviewed publications. He received his Ph.D. from the Department of Computer Science and Engineering at Birla Institute of Technology, Mesra, Ranchi, India in 2015. He received his M. Tech degree from the Department of Computer Science and Engineering at Birla Institute of Technology, Mesra, Ranchi, India in 2009.