Genetic Algorithm based Cluster Head Selection for efficient WSN

Dr. M. V. Vyawahare
Priyadarshini College of Engineering

Dr. P. R. Rothe
Priyadarshini College of Engineering

Abstract— With the development of various applications of wireless sensor networks, they have been widely used in different areas. Energy saving in these networks is an essential challenge. Considering that the energy consumption rate while sensing information and receiving information packets from another node is constant, the sensor nodes consume maximum energy while performing data transmission. Therefore, the routing methods try to reduce energy consumption based on organized approaches. One of the promising solutions for reducing energy consumption in wireless sensor networks is to cluster the nodes and select the cluster head based on the information transmission parameters such that the average energy consumption of the nodes is reduced. A genetic algorithm-based clustering technique, called GA-clustering, is proposed in this article. The searching capability of genetic algorithms is exploited in order to search for appropriate cluster centres in the feature space such that a similarity metric of the resulting clusters is optimized. The chromosomes, which are represented as strings of real numbers, encode the centres of a fixed number of clusters and the network lifetime is increased. Thus, in this study, a novel optimization approach has been presented for clustering the wireless sensor networks using the multiobjective genetic algorithm .The implementation results show that considering the capabilities of the multi-objective genetic algorithm has improved energy consumption, efficiency, data delivery rate, and information packet transmission rate compared to the previous methods.

Keywords— Wireless sensor networks, Cluster Head, stability period, Network Lifetime.

DOI Number: 10.48047/nq.2022.20.19.NQ99218 **NeuroQuantology2022;20(19): 2563-2570**

1. Introduction

WSNs are usually comprised of a large number of sensor nodes equipped with limited energy resources, but they should operate for a long time without charge or battery replacement. To increase the network lifetime and reduce energy consumption of the sensor nodes of the network, clustering techniques have been presented to achieve an efficient relationship among the sensor nodes [1, 2].

In the clustering techniques, the sensor nodes of a network are combined to constitute small separate clusters. Each cluster has a known leader called the cluster head (CH) and other nodes are known as the member nodes (MN). Selecting the

CH is a fundamental challenge that is the topic of this study. The sensor nodes sense the environment information and transmit it to the corresponding CH. The CH nodes collect data from all sensor nodes of the cluster and transmit it to the base station after data aggregation and removing the duplicate data. Thus, the CH has to organize the network, collect data, and transmit data from the sensor nodes to the sink and the base station and it consumes more energy compared to other nodes [3–4].

Given that the main challenge in wireless sensor networks is energy constraint, so, the performance of most applications in the WSN depends on energy consumption [5]. Hence, the main goal of this paper is at saving energy consumption in wireless sensor nodes. Since the amount of energy required to sense data from the environment and receive packets from other sensor nodes in the WSN is constant, therefore, most of the energy consumption is related to sending packets. The farther the next hop is from the current node, the more energy it takes to send data. So, the closest next hop that has the most remaining energy and least distance to sink not only can save energy but also could improve the quality of service (QoS) parameters. Therefore, this study presents a novel optimization approach using the multiobjective genetic algorithm (MOGA) and the gravitational search algorithm (GSA) for clustering the WSNs. In this study, the multiobjective genetic algorithm based on reducing the intracluster distances and reducing the energy consumption of the MNs is used to select the CH and the nearly optimal routing based on the gravitational search algorithm is used to transfer information between the cluster head nodes and the sink node. Considering the capabilities of the multiobjective genetic algorithm and the gravitational search algorithm, the proposed method has improved energy consumption, efficiency, data delivery rate, and information packet transmission rate compared to the previous methods.

The rest of this paper is organized as follows. Section 2 reviews the literature. Section 3 describes the details of the proposed method. Section 4 presents implementation and evaluation of the proposed method. Finally, the paper is concluded in Section 5.

2. Literature Review

In this section, the routing protocols based on swarm intelligence for WSNs are studied. The protocols employed in WSNs, including the ant colony optimization (ACO) algorithm, particle swarm optimization (PSO), are discussed.

The genetic algorithm-based energy efficiency clusters (GABEEC) is used to increase the network lifetime. The GA evaluates all chromosomes by calculating the fitness function. The fitness function has three parameters, including the round in which the first node dies, the round in which the last node dies, and the cluster distance. This algorithm tries to reduce the network lifetime by reducing the distance of the nodes, but the communication at the CHs is reinforced due to transmitting information about the residual energy to the base station (BS). The genetic clustering algorithm (GCA) employs two parameters, including the total transmission distance in a cluster and the number of CHs to achieve a longer lifetime [6]. The genetic algorithm-based energy efficiency clustering hierarchy (GAECH) performs the GA twice and improves CH selection considering the residual energy and total transmission cost [7].

In the network routing, the ACO-based techniques achieve a better overhead due to real-time computations and less control. The ACO has various shortcomings, for example, its performance depends on the previous cycle. In a study, the authors presented a CH selection algorithm using the ACO to construct balanced load clusters. This algorithm uses the residual energy of the node and the distance between the nodes to select the CH [8].

PSO is a subset of swarm intelligence based on a population-based random optimization approach. PSO applies the social behavior of the birds or fishes to the real-world problems. This approach preserves local solutions and global solutions and generates the best fitness of an objective [9]. Also, swarm intelligence is used in the WSN for clustering optimization. In a study, the authors presented a clustering algorithm using PSO. They considered two types of nodes: natural sensor nodes and high-energy nodes. The high-energy nodes operate as CHs while the normal sensor nodes operate as members of the cluster [10]. PSO is also used in the information broadcast protocol.

3. The Proposed Method

A novel optimization approach is presented for WSN clustering using the multiobjective GA algorithm. The multiobjective GA based on reducing the intracluster distance and energy consumption of the MNs is used to find the CHs rate, and reduce the information packet transmission delay.

the proposed method has two main steps as follows:(i)WSN configuration and random node clustering in the network(ii)Determining the optimal CHs using GA

3.1. Initial Clustering

In the first section of the proposed method, the proposed WSN is simulated with 100 sensor nodes and one sink node. The initial parameters are considered based on the standard parameters of similar methods. Finally, the proposed method is compared with other available methods. After initial configuration of wireless sensor nodes in the network, a "Hello" message is transmitted by the sink node to all sensors to identify the nodes and determine the location of the existing nodes of the network. All sensors of the network transmit a routing reply (RREP) to the sink node after receiving the "Hello" message to obtain the exact location of each sensor node for initial clustering of the nodes. Since the energy required to transmit the RREP packets and data is different, different energy consumption constants are considered for each packet type. Thus, the energy consumption of the proposed method is examined accurately. In the next step, after receiving the RREP and identifying the initial location of the wireless sensor nodes, the proposed method clusters the nodes based on the random CHs. Since the initial energy of all nodes at the beginning steps is constant, random selection of the CHs does not interrupt the data transmission and early energy discharge of some sensor nodes does not occur. Then, the wireless sensor nodes are clustered based on their distance from the CH node. The distance of the sensor nodes from the CH node is measured based on the Euclidean distance, In this step, the CH nodes are given to the GA as the initial chromosomes so that the optimal location of the CH is found. In the proposed GA, an optimal node in the selected cluster might not be found as the CH. In this case, the cluster is wound up and upon finding an optimal CH around the cluster, clustering is carried out again.

3.2. Using GA to Determine the Optimal CHs

As mentioned, in the second step of the proposed method, the GA is used to find the new and optimal CHs instead of random CHs selected in the previous step. In the following, the chromosomes are configured and the fitness function of the GA for the proposed method is defined.

3.2.1. Chromosome Encoding

The proposed GA receives an initial set of CHs as input and initiates by generating an initial population of the chromosomes, where each chromosome represents a possible solution for the clustering problem. Therefore, a chromosome is a vector of genes and the numbers inside each chromosome represent the index of a CH node. In Figure 1, an example of the chromosomes of the proposed method is given.

G_1	G_2	G_3	G_4	G_5	G_6	G_7	G_8
5	14	25	39	42	65	9	86

Figure 1
Chromosomes of the proposed method.

As shown in Figure 1, the chromosomes of the proposed method include a vector of genes where each gene represents an index of a CH node. The numbers inserted in each gene represent the CH selected randomly in the initial step. At the beginning, each chromosome is considered as the probabilistic clustering in the WSN that changes by applying the fitness function and mutation and crossover operators. Finally, the chromosome with maximum fitness is selected as the nearly optimal cluster.

3.2.2. Fitness Function of the GA

The fitness function of each chromosome is determined considering the objective function that is a combination of residual energy of the CH, mean intracluster distance, and distance of the CH from the sink. For each possible clustering of the population, the fitness function is considered as the representative of the parameters of interest to balance the residual energy of the CH, mean intracluster distance, and distance of the CH from the sink. Since the scales of the distances between the nodes and energy of the sensor nodes is not the same, the values of the proposed method should be normalized to obtain a unit value as the fitness function; The fitness function used to evaluate the given chromosomes using the proposed objective function,the proper chromosomes are selected from the initial population and other chromosomes are transmitted to the mutation and crossover operators to diversify the population and generate new superior chromosomes. Each chromosome of the new offspring population is checked to see if it is a possible solution for the problem or not (does it minimize the fitness function and satisfy the given constraints or not). The impossible chromosomes that violate the existing constraints are penalized considering their fitness value such that they have a lower probability to be selected for generation and conversion to new chromosomes. The most proper chromosome that represents the nearly optimal clustering solution is preserved at each iteration and sorted based on its optimality. This process continues until the termination condition is met.

3.2.3. The Crossover Operator

The crossover operator is an essential step of the GA to diversify the population and generate new chromosomes. To increase the search domain and public feasible solutions, the GA should apply the crossover operator between two chromosomes (parents) and generate new offspring as the new population. The crossover operator is performed as a random replacement of a number of genes of the first chromosome with the second chromosome. Thus, the parameter that is essential in the crossover operator is called the crossover probability or -crossover, which is defined as follows for random crossover.in which is the maximum number of genes and is the minimum number of genes of the chromosome. The parameter is considered as a random number in the range of and . The value of -crossover is considered as a part of the chromosome that should be exchanged between the first chromosome and the second chromosome; the beginning and ending points are called the crosspoints. The beginning crosspoint might be selected from the beginning of the chromosome or any other part of the chromosome, and the ending crosspoints are added to the -crossover. Figure 2 shows the crossover operator.

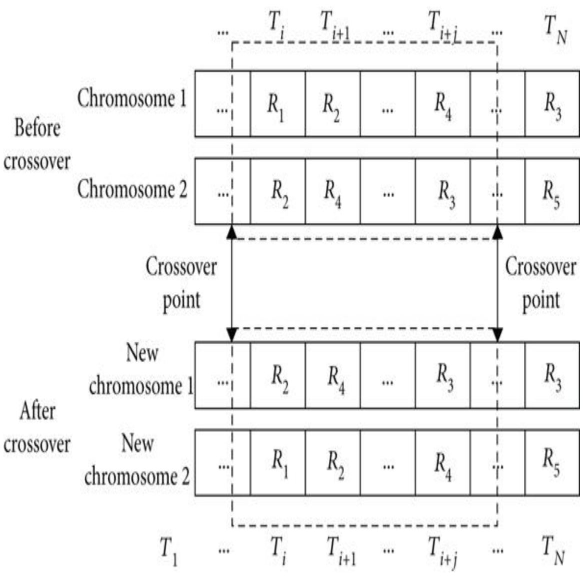

Figure 2
An example of crossover of two chromosomes.

As shown in Figure 2, each gene, , represents a CH node and represents the cluster of each gene for any of the solutions.

3.2.4. Mutation Operator

The mutation operator, like the crossover operator, also plays a significant role in generating a new population and diversifying the chromosomes. This operator can increase the search domain and the possible solutions and generate new offspring as the new population. Thus, in this operator, the probability parameter is of great importance that is considered as a chromosome that should mutate. The -mutate parameter or the mutation probability can be calculated.

The -mutate parameter represents the mutation probability, and the parameter can be considered as a value between zero and one; zero or one can be considered as the first or the last gene of the chromosome. The difference of the mutation and the crossover operators is that the crossover operator changes several genes of a chromosome with another chromosome but the mutation operator only changes the value of one gene to generate a new chromosome. Figure 3 shows the mutation operator.

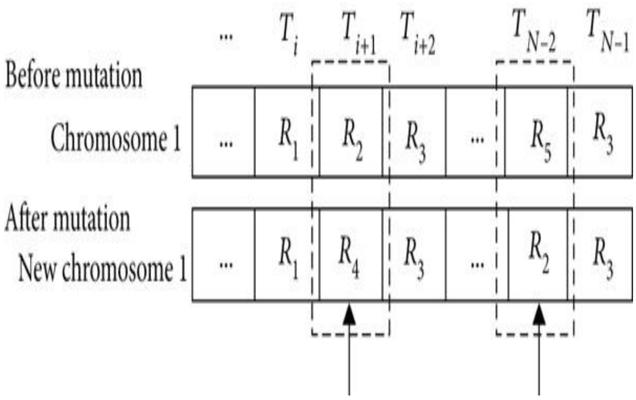

Figure 3
An example of population diversity of the chromosomes.

As shown in Figure 3, there are two chromosomes with different genes that change at and , independently. At , the value of the gene changes from to . Similarly, the value of the gene mutates from to at . The gene mutation might yield excellent results. Sometimes the results might not be satisfactory. However, the gene mutation is essential to preserve population diversity.

3.2.5. Selection Operator

The selection operator, after the crossover and the mutation operators, selects the chromosomes with maximum fitness or the nearly optimal solution among the new population and the chromosomes generated as the next generation for the clustering problem to reduce energy consumption in the WSN. In this case, the CHs assigned to each cluster with maximum residual energy, minimum mean intracluster distance, and minimum distance to the CH are examined to balance the energy consumption of each cluster. Therefore, the efficiency of the clustering algorithm is optimized to balance the energy consumption of the WSN.

4. Implementation of the Proposed Method

To implement the proposed method, first, the WSN is configured based on the standard parameters. The proposed network is implemented in a environment. To implement this scenario, MATLAB 2021a is used. Figure 4 shows the initial configuration of the proposed WSN.

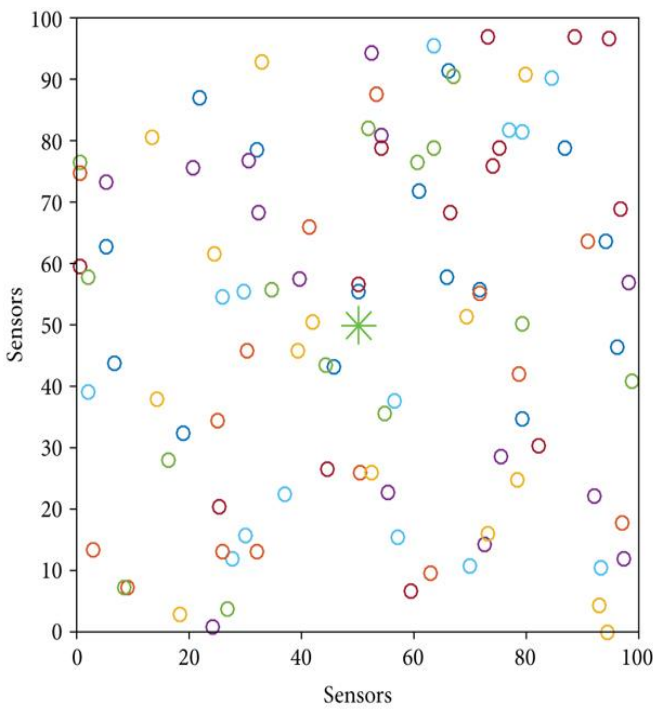

Figure 4
Initial configuration of the proposed WSN.

As shown in Figure 4, the proposed WSN is simulated based on the initial parameters. This network is comprised of 100 sensor nodes that are distributed randomly in the network. The sink node is also at the center of the network that facilitates access to it.

In the first step, the sink node collects information about the location of the network sensors and selects multiple CHs accordingly. The CHs are selected randomly, and the sensor nodes also join the CHs based on their distance and constitute clusters. After formation of the first cluster and transmitting data to the CH, the information about the initial energy and the intracluster distances and distance of the CH from the source can be calculated. In Figure 5, the initial clustering of the WSN is shown.

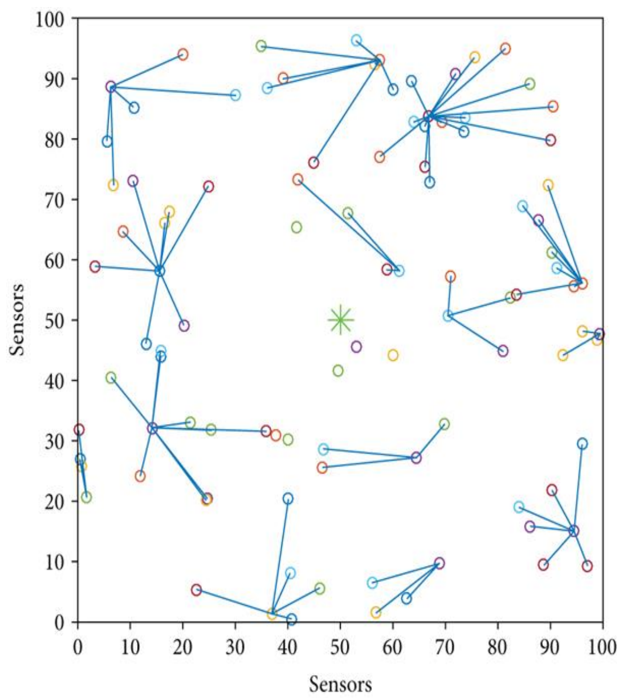

Figure 5.
Initial clustering of the nodes in a WSN.

Information is given as the initial chromosome to the GA, and the second step of the proposed method is implemented. In the first step of the proposed GA, the initial population is evaluated considering the randomly selected CHs. To evaluate the input population of the proposed GA, the fitness function is applied to the initial population. Accordingly, the proposed GA evaluates the CHs in terms of the residual energy, mean intracluster distance, and distance from the sink node. Since no data is transmitted in the initial population, the initial energy of all CH nodes and MNs is the same. Thus, in the first step, the CHs are evaluated in terms of intracluster distance and distance from the sink node. In the following, since the optimal routes are found, information transmission, energy consumption, and residual energy affect the selection of the optimal CHs. The fitness value of each CH in the initial population is calculated. It is seen that some CHs have an excellent fitness but other are weak. Thus, in the next step, the new population is generated based on the crossover and mutation operators and the fitness value of the new population is examined. The new population of the proposed method is a combination of the CHs that have replaced the previous CHs, and their fitness is evaluated.

The clusters of three CHs with indexes of 86, 80, and 80 are broke down, and they are integrated in a cluster with CH 51. Also, it is seen that CH 97 is transferred from the previous population to the new expert population. Figure 6 shows the replacement of the new CHs with the previous CHs.

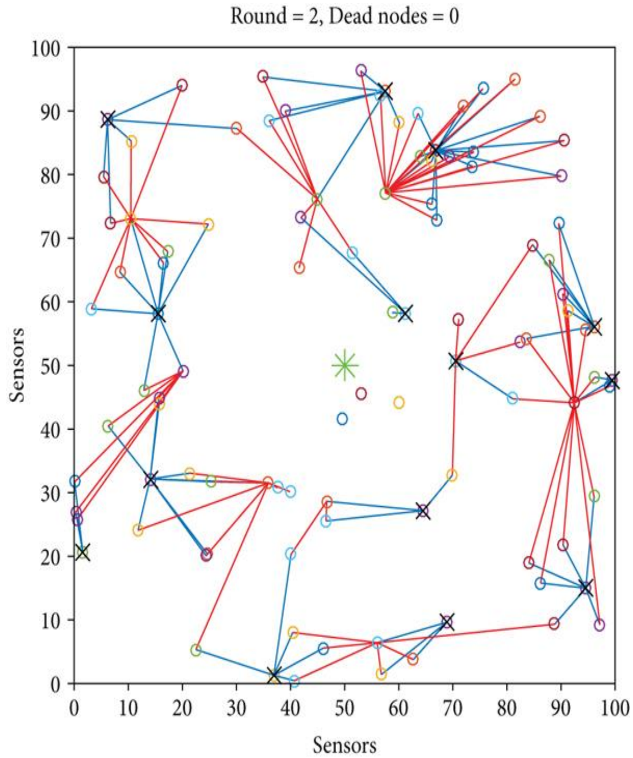

Figure 6
Replacement of the new CHs with previous CHs.

As shown in Figure 6, the previous CHs are represented with a black , the previous connections are represented in blue, and the connections with new CHs are represented in red. The new CHs are selected as the proposed solution for information transmission form sensor nodes, data aggregation, and data transmission to the sink node. Considering the distance of the CH nodes from the sink node, if a CH observes another CH along the direct route to the sink, the packet transmission is carried out in multihops between the CHs and the sink node.

5. Conclusion

The wireless sensor network is one of the most recent environments monitoring and controlling networks that collect information from the environment and aggregate data for network applications autonomously without any infrastructure. Thus, the popularity of these networks has resulted in various challenges, among which the unbalanced energy consumption can be mentioned. Considering the limited energy of the sensor nodes, unbalanced energy consumption might affect all performance measures of the network. Thus, in this study, a novel optimization approach using the multiobjective genetic algorithm has been presented for WSN clustering. In this study, the multiobjective genetic algorithm based on reducing the intracluster distances and the energy consumption of the member nodes is used to select the cluster heads. The implementation results of the proposed method show that considering the capabilities of the multiobjective genetic algorithm, the proposed method has improved the average energy consumption, data delivery rate, and network lifetime significantly compared to the previous methods.

References

1. L. Xu, R. Collier, and G. M. P. O'Hare, "A survey of clustering techniques in WSNs and consideration of the challenges of applying such to 5G IoT scenarios," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1229–1249, 2017.
2. S. R. Nabavi and S. M. Mousavi, "A review of distributed dynamic key management schemes in wireless sensor networks," *Journal of Computers*, vol. 13, no. 1, pp. 77–89, 2018.
3. H. Ouchitachen, A. Hair, and N. Idrissi, "Improved multi-objective weighted clustering algorithm in wireless sensor network," *Egyptian Informatics Journal*, vol. 18, no. 1, pp. 45–54, 2017.
4. S. R. Nabavi, N. Osati Eraghi, and J. Akbari Torkestani, "Temperature-aware routing in wireless body area network based on meta-heuristic clustering method," *Journal of Communication Engineering*, vol. 9, no. 2, 2020.
5. S. R. Nabavi, N. Osati Eraghi, and J. Akbari Torkestani, "Wireless sensor networks routing using clustering based on multi-objective particle swarm optimization algorithm," *Journal of Intelligent Procedures in Electrical Technology*, vol. 12, no. 47, pp. 29–47, 2021.
6. S. Mudundi and H. Ali, "A new robust genetic algorithm for dynamic cluster formation in wireless sensor networks," in *Proceedings of the 7th IASTED International Conferences on Wireless and Optical Communications, WOC*, pp. 360–367, Montreal, QC, Canada, 2007.
7. B. Baranidharan and B. Santhi, "GAECH: genetic algorithm based energy efficient clustering hierarchy in wireless sensor networks," *Journal of Sensors*, vol. 2015, Article ID 715740, 8 pages, 2015.
8. C. K. Ho and H. T. Ewe, "A hybrid ant colony optimization approach (hACO) for constructing load-balanced clusters," in 2005 IEEE Congress on Evolutionary Computation, pp. 2010–2017, Edinburgh, UK, 2005.
9. T. Gui, C. Ma, F. Wang, and D. E. Wilkins, "Survey on swarm intelligence based routing protocols for wireless sensor networks: an extensive study," in *Proceedings of the IEEE International Conference on Industrial Technology*, pp. 1944–1949, Taipei, Taiwan, May 2016.
10. M. Azharuddin and P. K. Jana, "Particle swarm optimization for maximizing lifetime of wireless sensor networks," *Computers & Electrical Engineering*, vol. 51, pp. 26–42, 2016.