Designing a new clustering-based routing algorithm using a genetic algorithm to expand the lifetimeof the IoT network based on wireless sensors

Behzad Habili, Department of Computer Engineering, Iranian university, Tehran, Iran.

Fatemehkhezri, Department of Computer Engineering, south Tehran Branch, Islamic Azad University, Tehran, Iran.

Abstract

Background: Energy is one of the main factors forcommunication in the Internet of Things based on sensors. Sensors are a low and non-rechargeable energy source because sensors are commonly used in places and applications where they cannot be recharged.

Purpose: The most important goal of this study is to reduce the energy consumption of sensors and increase the lifetime of the Internet of Things network using genetic algorithm while selecting a clusterhead and routing between clusterheads.

Methods: In this paper, after distributing nodes in the network, the genetic algorithm is used to select the cluster heads and also to create a route between the cluster heads to the base station. After selecting the cluster head by the genetic algorithm, the normal nodes join thecluster head and send their data to it. After collecting the data and processing them, the data is transferred to the base station from the route specified by the genetic algorithm.

Results: The proposed algorithm is implemented with MATLAB simulator and compared with LEACH, MB-CBCCP and DCABGA protocols. The simulation results show that the proposed algorithm works better than the mentioned algorithms in different environments.

Conclusion: Due to energy constrained and the fact that in most applications the batteries can not be recharged, the use of optimization algorithms in designing and implementation of routing and clustering algorithms has a significant impact on expanding the lifetime of these networks.

Keywords: IoT, IoT clustering, IoT routing, Genetic algorithm

1.Introduction

The Internet of Things is one of the new technologies nowadays. In the IoT, smart objects connect with each other, data is collected, and specific requests from users are met with the various data provided. The development of energy efficient schemes for the IoT is a challenging problem, as IoT becomes more complex due to its large scale, and current wireless sensor network techniques cannot be applied directly to the IoT [1].

The Internet of Things refers to the precise communication between the digital and physical worlds [2]. Various researchers have described IoT in different ways [3], [4]:

"A dynamic global network infrastructure with the ability to automatically adjust based on interactive communication standards and protocols that physical and virtual" objects "have identities, physical properties and virtual characters and use intelligent interfaces, andintegrates with the information network seamlessly".

In sensor-based IoT networks, the sensor is the main power source for the nodesof the battery. And in most applications, due to the large number of sensors and the unavailability of sensors, switching their energy source is impossible or very difficult. Therefore, energy saving is one of the factors that should be considered[5], [6].

There are many approaches to reduce energy consumption in the sensor-based IoT network, and one of the most widely used and effective approaches is the improvement of clustering and routing algorithms in these networks, which has recently been considered by many researchers. Due to the lack of energy source in sensor-based IoT networks, designing of energy efficient algorithms to reduce the energy consumption of sensors is very important.

The network can be divided into small sections where the sensors are divided into parts that are called clusters. In each cluster there is a cluster head whose task is to collect data from member nodes and transfer the collected data to the base station. The advantages of clustering are scalability of the network, localization of route settings, preservation of communication bandwidth by reducing relay packets, reduction of energy consumption rate and stability of network topology [7].

Clustering supports scalability and can be expanded to any level. The communications between the sensor nodes inside the clusters are local.Sensors are only involved in communicationwith their cluster head and are not affected by changes between the cluster head [9] [10].

The problem of routing is as important as clustering. In general, routing answers the following question: How does an entity get from source to destination? In the Internet of Things, the entity is a data packet, and the sourceand destination of the data packet are two computing devices [10]. Packets must pass through intermediate nodes before reaching their destination. This method is known as multi-stage routing, which reduces energy consumption in nodes because nodes do not need to send their data over long distances[11] [12].

The rest of this paper is organized as follows: In Section 2, a review of previous studies is given, and in Section 3, the proposed algorithm is describedin detail, and in Section 4, the results obtained from the simulationsare analyzed, and in the final section,the conclusion is given.

2-Previous Studies(Literature Review)

Network node clustering is a successful control and topologydesign method that can be used to increase network productivity. The two main steps in cluster-based routing are clusterheadselection and routing throughcluster head [13] [14]. By using cluster heads to collect data and send this information from cluster heads to central stations,it is possible to save the energy [14]. Therefore, selecting the appropriate cluster head among the nodes can improve energy efficiency and increase the lifetime of the sensor network.

In this section, we will review some routing and clustering protocols that have been considered in recent years.

Low-Energy adaptive clustering Hierarchy(LEACH) is a distribution algorithm that locally selects the clusterhead. In this method, the cluster heads are selected based on the possible model. All data processing, such as data integration and collection, is done locally within the cluster. In LEACH, the shape of clusters is formed by using adistributedalgorithm and nodes make decisions independently without any centralized control. First, a node randomly decides to convert into a p-probability cluster head and publishes its decision. Any non-cluster head node can become a desired cluster head member according to the minimum energy required to communicate with the cluster head . The cluster headrotates alternately between the cluster nodes to maintain load balance. This rotation is performed by taking each node i to select a random number T (i) between 0 and 1. Each node iconvert into a cluster head for the current rotation if the number T (i) is less than the following threshold:

$$T(i) = \frac{p}{1-(r.mod\frac{1}{p})}, i \in G$$

Equation (1)

Where p is the desired percentage of cluster headnodes in the sensor population, r is the number of rounds and G is the set of nodes that are not cluster headin the 1/ppreviousround. LEACH forms the cluster topology of a hop that each node can be transferred directly to the cluster head and then transferred to the base station. See Figure 1.

Limitations of LEACH: Although LEACH savesenergy in nodes and expands network lifetime, it still has some drawbacks (limitations) [15], [16]:

- LEACH is appropriate for small size networks
- At the beginning of each round, LEACH randomly selects cluster headregardless of the energy remaining from these nodes.
- Eachcluster head communicates directly with the base station using LEACH, regardless of whether the distance is close or not.
- Cluster heads can be concentrated in one place and therefore nodes are separated (without cluster heads).
- There is no mechanism in LEACH to ensure that the selected cluster heads are evenly distributed over the network.

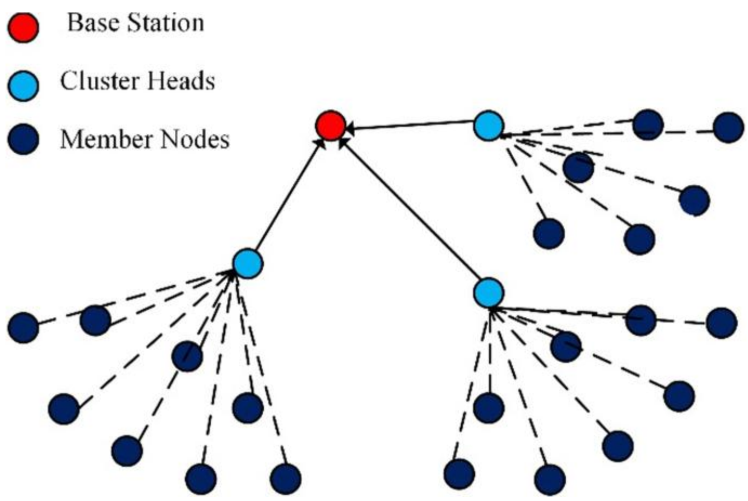

Figure 1: How the LEACH algorithm works

The MB-CBCCP protocol is introduced by Rani et al. [16] for the WSN-based IoT network. In this method, the network environment is layered for clustering operations. Then, in each cluster, the gate nodes, thecluster head, and the coordinator are identified. Then, in each cluster, the normal nodes become members of the nearest gate node, and the gate nodes are connected to the cluster heads. The cluster heads are then connected to the coordinators and the coordinator nodes are connected to the nearest coordinator node of the higher cluster. In the next step, the last layer coordinators are connected directly to the sink, and at the end, the last layer cluster head is connected directly to the base station.

In this method, RNs are selected randomly and no distance parameter is considered for the next RN selection. As a result, some RNs may overlap with each other. Hence, this incurs additional system costs (due to more node selection as RN) and the use of inefficient resources. The number of RNs selected can vary from cluster to cluster and depends on the density of the cluster nodes.

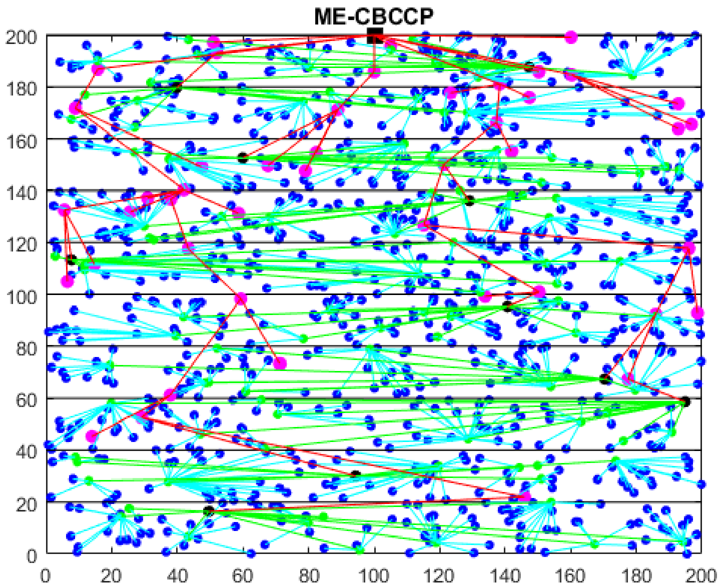

Figure 2: How MB-CBCCP works

Dynamic clustering based on DCRN genetic algorithm: In this paper [17], a dynamic clustering algorithm using genetics is presented. In this algorithm, first the nodes are randomly placed in an environment and the sinkis placed statically outside the environment. Next, the clusters are selected by a genetic algorithm whoseFitness function includes: 1- Set of distance of all cluster heads to base station 2-Cluster distance: Set of distance of all nodes to their cluster heads 3-Standard deviation of cluster distance 4-Transmission energy. Once the clusterheads are identified, the normal nodes become the membersof closest cluster head.

After this step, the information is collected by normal nodes and sent to the cluster head, and then the cluster heads send the information as a strep tag to the base station. In this method, the cluster heads send their information to the base station as a strep tag, which in turn increases the energy consumption of the cluster head that are far from the base station.

3-System model

Before describing the proposed algorithm, we state the hypotheses considered in this method:

- In energy simulation, all nodes are 1 joule.
- Nodes are randomly and uniformly distributed in the environment.
- The base station (sink) can be anywhere in the network environment. However, we consider it outside the environment.
- Not all sensor nodes need to be fixed after distribution. However, mobility here does not involve much change of original location by remote control. It only includes ground displacements such as erosion or displacements caused by foreign objects that cause in situ changes
- Nodes are able to adjust the transmission power based on the distance of the receiving nodes.

3-1-Energy consumption model

Energy consumption in a sensor-based IoT network consists of three parts: data transmission, data reception, and data processing. The energy model is given in Equation 2 [16], [18] [19].

Equation (2)
$$\begin{cases} P_T(K) = E_{elec} * K + E_{amp} * d^{\gamma} * K \\ P_R(K) = E_{elec} * k \\ P_{cpu}(K) = E_{cpu} * k \end{cases}$$

P_T, $P_R$and P_{cpu}represent the energy used to send, receive and process k-bits of data, respectively.E_{elec}, $E_{amp}$and E_{cpu}represent the energy consumption (nJ/bit) for each bit transmitted in the radio radius, the energy required for transmission with a radius greater than $E_{elec}$and the energy required for processing each bitrespectively. According to Equation 2, the total energy consumption of k bits is as equation 3 [7] [20] [21].

Equation (3)$P_{\text{کل}} = P_{\text{ارسال}} + P_{\text{دریافت}} + P_{\text{پردازش}} = k(2E_{elec} + E_{cpu} + E_{amp} \times d^{\gamma})$

Equation 3 shows that energy consumption is directly related to data length. If the data sent first is less, less energywill be consumed. If the transmission distance is less than the threshold, the energy consumption will be related to d^2. If the transition distance is greater than the threshold, it is related to d^4. Therefore, the shorter the transmission distance, the lower the energy consumption.

2-2-Explaining the proposed method

The main purpose of this study is to reduce the energy consumption of IoT-based sensor networks. Therefore, this paper presents an approach based on optimization algorithms to reduce energy consumption in the network. The goal of optimization methods is ultimately to find the best possible answer (close to optimal) in an acceptable time. In this section, we will explain our proposed algorithm.

Proposed Algorithm

```
Select nodes in sensing area for clustering
1.  For k=1:number of clusters
2.      Calculate remain energy, density, centerality and CH distance of nodes;
3.      Select clusterhead with CHGeneticAlgorithm();
4.  End_For
5.  For k=1:number of clusters
6.      For i=1:number of nodes
7.          If node_i is normal node
8.              Node_i joins to nearest clusterhead_k;
9.          End_IF
10.     End_For
11. End_For
Routing to send cluster head information
1   Route=RouteGeneticAlgorithm ();
```

```
For i= clusterhead
    Clusterhead _i joins to route;
End_For
Last node of route, connects to sink;
```

Genetic algorithm is a family of computational models inspired by the concept of evolution. These algorithms encode potentialSolutionsor candidate Solutionsor possible hypotheses for a particular problem in a Chromosome-like data structure. The genetic algorithm preserves vital information stored in chromosome-like data structures by applying recombination operators to chromosome-like data structures.

The implementation of a genetic algorithm usually begins with the production of a population of chromosomes (the initial population of chromosomes in genetic algorithms is usually randomly generated and is bound to upper and lower of the problem variables). Next, the data structures produced by the chromosomes are evaluated, and the chromosomes that can better represent the optimal solutionto the problem (target) have a better chance of reproduction than the weaker solutions. In other words, more reproductive opportunities are allocated to these chromosomes. The goodness of a solutionis usually measured by the population of the current candidate's solutions.

A circulation diagramof evolutionary algorithms and the structure of the genetic algorithmare presented below.

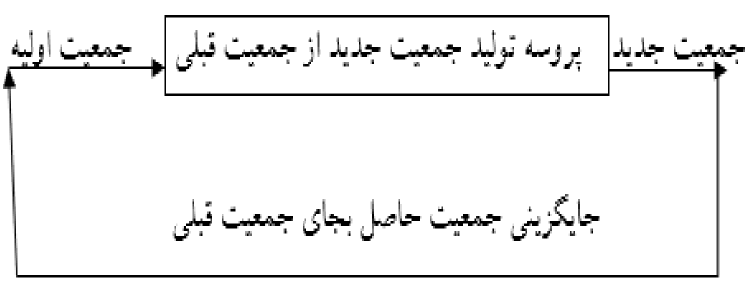

Figure 3: The production process continues until the optimal solutionis obtained

(Most of the initial population is randomly generated)

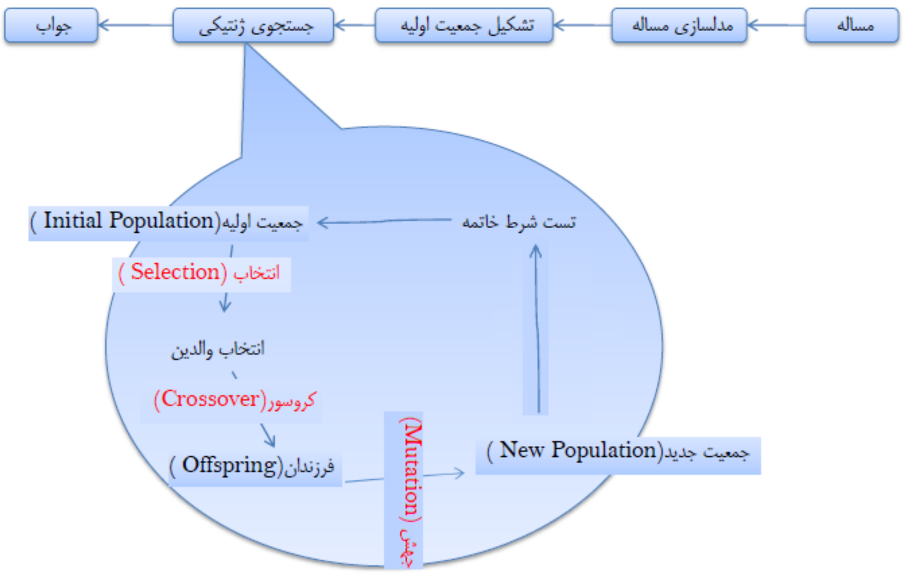

Figure 4: General structure of the genetic algorithm

In the proposed algorithm, the algorithm starts working after the sensors are distributed in the environment. In the proposed method, in the base station, the operation of determining the cluster heads is done by the genetic algorithm. Also the routing between the cluster headsto the base station is done by the genetic algorithm. The proposed algorithm works on a round basis and each round consists of two phases, the first phase is called *Setup* and the second phase is called *steady state*.

In each round, the cluster headsare determined using a genetic algorithm, and the routing between the clusters to the base station is determined by the genetic algorithm, and then a message is sent by the base station to the cluster headsso that the cluster headsare aware of their role. Also they become aware of the specified tree (route) from the cluster headsto the base station. Normal nodes then join cluster headsby receiving messages from them and become members of the cluster headsthat use less energy to connect with them. Then, after all the sensors are connected to the cluster heads, the TDMA[1] scheduling operation is performed by the cluster headsto prevent the data from colliding during the transfer. Also, with the identification of the cluster heads, the operation of determining the route from the cluster headsto the base station is performed by the genetic algorithm and the cluster headsare notified to be aware of the

1. Time-division multiple access

desired route for transferring information to the base station in the same round. After this phase, the second stage called the steady state begins.

Cluster headsdetermine by genetic algorithm based on the cost function that consists of residual energy parameters, density (is equal to the ratio of neighboring nodes to total nodes, and the higher the density,the more appropriate the node is forbeing a cluster head) and centrality (meansthe centrality of the nodeto its neighbor nodes and is the sum of the total distance of the node from its neighbors. Low valueindicates that node needs less energy to be a cluster headand it is more appropriate to be a cluster head). After selection of the cluster head, othersensors can join one of the clusters by communicationwith a cluster head and based on distance and energy consumption for communication.

Cost function equation:

Equation (4) $Cost = -(E_R) - (D = \frac{N_n}{N_T}) + (C = \sum d_i)$

E_R is the residual energy of the nodes. The density of nodes is equal to $D = \frac{N_n}{N_T}$ which N_n represents the neighboring nodes and N_T represents the total nodes. The centrality is also equal to $C = \sum d_i$ which d_i shows the distance to the neighboring node.

In the proposed algorithm, the population size is consider to be 10 and the number of iterations of the algorithm to be 20. 10% of the population is selected for mutation and 20% of the population for fertilization.A cost function is get for all members of the population. In each iteration, the parent algorithm is specified. Parents are selected according to their cost. This means that the lower the cost of a parent, the more likely it is to be chosen as a cluster head. These parents are merged in pairs and two children are produced and the cost of each child is recalculated. And the better child is selected,then for 20% of the nodes,chromosomes are selected and mutation is done upon them.And in the end, the best chromosome, which costs less, is selected as the cluster head. And the base station informs the cluster heads of their role, then the cluster heads send a message that they are cluster heads on the network, and every normal node that receives this message, based on the cluster head's energy, becomes the member of therelevant cluster head and announce its membership to the relevant cluster head. As mentioned earlier, routing is done using a genetic algorithm, which we will explain below. In a population-based genetic algorithm, a population has a set of chromosomes, each of them has a set of genes.We have an early population that consists of several chromosomes. The number of genes on chromosomes is equal to the number of cluster heads minus one, whose field value is given by random numbers between zero and one. And its cost value is equal to the amount of network energy consumption when creating the route between the cluster heads to the base station, which is created using the prufer algorithm of the desired tree. The cost value is obtained from the following equation.

Equation (5)

$$\mathrm{Cost} = \sum_{i=1}^{i=k} E_{Consumed}$$

$E_{Consumed}$ is the energy used by the cluster heads on the route to the base station to send information to the base station.

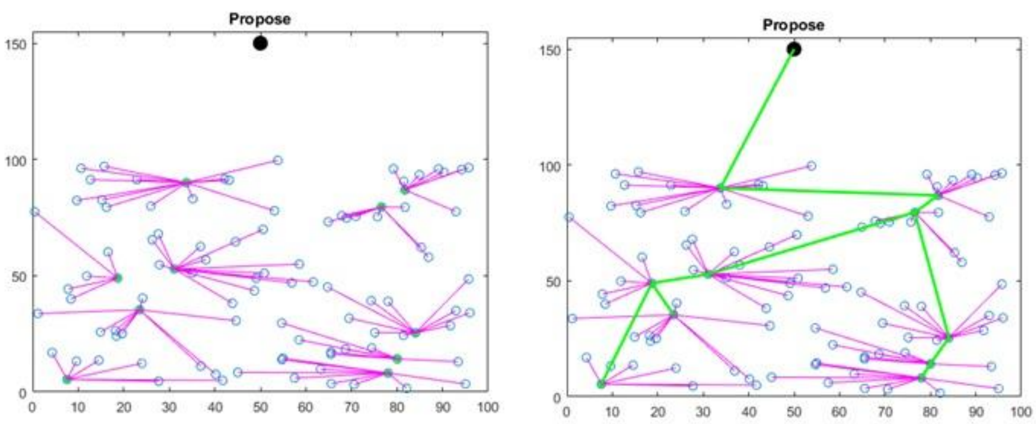

Figure 5: Recruitment of clusters and determination of the route between cluster heads by genetic algorithm

In our case, each chromosome can be asolution. The population is first randomly assigned. We set the number of iterations of the genetic algorithm to 100 to find the most appropriateroute. In the proposed algorithm, 10% of the population is considered for mutation and 20% of the population for selection.

The next step is selecting parents based on the roulette wheel algorithm, selecting 20% of the chromosomes that are lower in cost (lower network energy consumption) as parents. And the cross over operation on these offspring is done as a single point and the new chromosomes are known as the offspring of the next generation and after this operation a mutation occurs which is done randomly on the genes of the chromosomes and after this new population replaces the previous population and this operation is performed up to 100 times to select the best chromosome as the route from the cluster heads to the base station.

In the end, the lower cost chromosome algorithm is selected as a solution to our problem, and at the end, a route between the cluster headsto the base station is formed by the Prufer code, which is announced to the cluster heads by the base station in the *setup* phase.

4-Findings

The proposed protocol is simulated with MATLAB software and compared with LEACH, MB-CBCCP and DCABGA protocols.

Table 1: Simulation parameters

Parameter	Value	Parameter	Value
E_0	2 J	ε_{fs}	10 pJ/bit/m^2
E_{elect}	5 nJ/bit	ε_{amp}	0.0013 pJ/bit/m^4
E_{DA}	5 nJ/bit/message	L_D	4000 bits
d_{break}	87.7 m	L_c	16 bits

In the first experiment, 300 sensor nodesare randomly distributed in an area of 400 x 400 square meters. In this test, the initial energy of all sensors is the same and two joules, and the base station is fixed and in position (300 x 600).

Table 2: Comparison of previous methods with the proposed algorithm

Death of the 100% Nodes	Death of the first node	Lifetime / Method	
248	98	LEACH	400*400 m^2
280	110	MB-CBCCP	
325	128	DCABGA	
438	173	Propose	

As it can be seen from Figure 11, the proposed method improves network energy consumption compared to previous algorithms. The proposed algorithm increases the network lifetime by about 30% compared to the DCABGA algorithm, 46% compared to the MB-CBCCP and 58% compared to the LEACH.

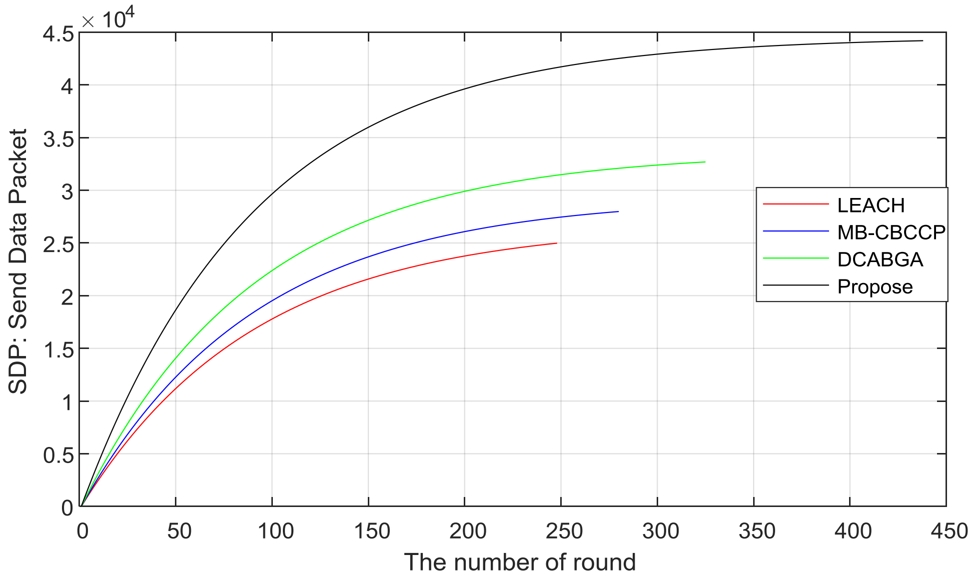

Figure 6: Number of packets sent to the base station

As it can see in Figure 6, the number of packets sent to the base station in the proposed method is about $4*10^4$, in DCABGA it is about $3.7*10^4$, in MB-CBCCP it is about $2.6*10^4$, and it is about $2.1*10^4$ in LEACH, which indicates that the number of packets sent to the base station in proposed method isabout 30% higher than DCABGA, 14% higher than MB-CBCCP and 59% better than LEACH algorithm.

Then we perform the test with the same number of sensors in 600 * 600 m^2 environments and base station (400 * 800).

Table 3: Comparison of different methods

Death of the 100% Nodes	Death of the first node	**Lifetime** / **method**	
86	46	LEACH	500*500 m^2
103	51	MB-CBCCP	

120	60	DCABGA	
162	81	Propose	

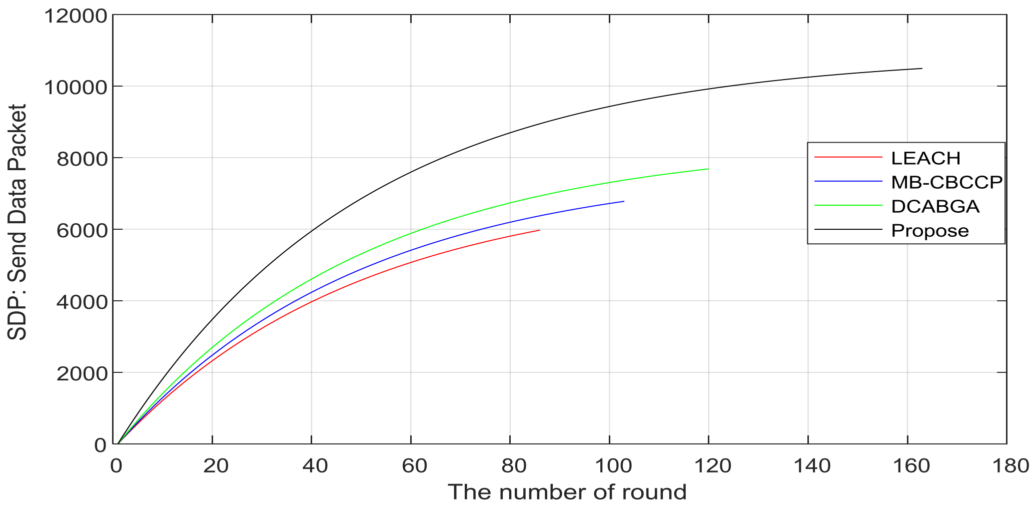

Figure 7: Number of packetssent to the base station in the environment of 500 * 500

As it can be seen infigures above, the proposed method improves network lifetimecompare to DCABGA by about 31%, compare toMB-CBCCP by about 42%, and compare toLEACH by about 60%, and sendingpackets to the base station increases by about 30% compared to DCABGA, compared to MB-CBCCP by 42% and compared to LEACHby 52%.

Conclusion

In this paper, a clustering-based routing method based on genetic optimization algorithms is proposed to optimize the lifetime of the sensor-based IoT network and to improve energy consumption. The proposed algorithm originally used the genetic algorithm for clustering and routing cluster head to the base station. Based on the experiments performed,the proposed algorithm improve the energy consumption and expand lifetime ofwireless sensor IOT network. In future research, other multi-objective optimization and machine learning algorithms can be used to develop the proposed protocol.

References

[1]P. Nayak and A. Devulapalli, "A Fuzzy Logic-Based Clustering Algorithm for WSN to Extend the Lifetime Network," IEEE Sens. J., 2016, doi: 10.1109 / JSEN.2015.2472970.

[2]L. Atzori, A. Iera, and G. Morabito, “The Internet of Things: A survey,” Comput. Networks, 2010, doi: 10.1016 / j.comnet.2010.05.010.

[3]P. P. Ray, “An Internet Survey of Things Architects,” Journal of King Saud University - Computer and Information Sciences. 2018, doi: 10.1016 / j.jksuci.2016.10.003.

[4]P. Sethi and S. R. Sarangi, “Internet of Things: Architectures, Protocols, and Applications,” Journal of Electrical and Computer Engineering. 2017, doi: 10.1155 / 2017/9324035.

[5]J. Y. Chang, “A Distributed Cluster Computing Energy-Efficient Routing Scheme for Internet of Things Systems,” Wirel. Pers. Commun., Vol. 82, no. 2, pp. 757–776, 2015, doi: 10.1007 / s11277-014-2251-8.

[6]H. Farman et al., “Multi-criteria based zone head selection in Internet of Things based wireless sensor networks,” Futur. Gener. Comput. Syst., 2018, doi: 10.1016 / j.future.2018.04.091.

[7]M. Sedighimanesh, H. Zandhesami, and A. Sedighimanesh, “Presenting the Hybrid Algorithm of Honeybee - Harmony in Clustering and Routing of Wireless Sensor Networks,” Int. J. Sensors, Wirel. Commun. Control, 2018, doi: 10.2174 / 2210327908666181029094346.

[8]S. Srivastava, M. Singh, and S. Gupta, “Wireless Sensor Network: A Survey,” in 2018 International Conference on Automation and Computational Engineering, ICACE 2018, 2019, doi: 10.1109 / ICACE.2018.8687059.

[9]C. Gherbi, Z. Aliouat, and M. Benmohammed, “A survey on clustering routing protocols in wireless sensor networks,” Sensor Review. 2017, doi: 10.1108 / SR-06-2016-0104.

[10] K. M. sedighimaneshmohammad, ZandhessamiHessam, Alborzi Mahmood, “Energy Efficient Routing-Based Clustering Protocol Using Computational Intelligence Algorithms in Sensor-Based IoT,” J. Inf. Syst. Telecommun., Vol. 33, no. 9, 2021, doi: 10.52547 / jist.9.33.55.

[11]N. Nasser, L. Karim, A. Ali, M. Anan, and N. Khelifi, “Routing in the internet of things,” in 2017 IEEE Global Communications Conference, GLOBECOM 2017 - Proceedings, 2018, doi: 10.1109 /GLOCOM.2017.8253955.

[12] T. J. Saleem, “A Detailed Study of Routing in Internet of Things,” Int. J. Eng. Sci. Innov. Technol., Vol. 5, no. 3, 2016.

[13]G. Xie and F. Pan, "Cluster-Based Routing for the Mobile Sink in Wireless Sensor Networks with Obstacles," IEEE Access, 2016, doi: 10.1109 / ACCESS.2016.2558196.

[14]K. Thangaramya, K. Kulothungan, R. Logambigai, M. Selvi, S. Ganapathy, and A. Kannan, "Energy aware cluster and neuro-fuzzy based routing algorithm for wireless sensor networks in IoT," Comput. Networks, 2019, doi: 10.1016 / j.comnet.2019.01.024.

[15] N. Chakchouk, "A Survey on Opportunistic Routing in Wireless Communication Networks," IEEE Commun. Surv. Tutorials, vol. 17, no. 4, pp. 2214–2241, 2015, doi: 10.1007 / 978-1-4419-0304-4_2.

[16]S. Mottaghi and M. R. Zahabi, "Optimizing LEACH clustering algorithm with mobile sink and rendezvous nodes," AEU-International J. Electron. Commun., Vol. 69, no. 2, pp. 507–514, 2015.

[17]S. Rani, S. H. Ahmed, and R. Rastogi, "Dynamic clustering approach based on wireless sensor networks genetic algorithm for IoT applications," Wirel. Networks, 2020, doi: 10.1007 / s11276-019-02083-7.

[18]W. Akkari, B. Bouhdid, and A. Belghith, "LEATCH: Low Energy Adaptive Tier Clustering Hierarchy," Procedia Comput. Sci., Vol. 52, pp. 365–372, 2015.

[19]M. Sedighimanesh, H. Z. Hesami, and A. Sedighimanesh, "Routing algorithm based on clustering for increasing the lifetime of sensor networks by using meta-heuristic bee algorithms," Int. J. Wirel Sensors. Commun. Control, vol. 10, no. 1, pp. 25–36, 2020.

[20]M. Sedighimanesh * and H. Z. H. and A. Sedighimanesh, "Routing Algorithm based on Clustering for Increasing the Lifetime of Sensor Networks by Using Meta-Heuristic Bee Algorithms," International Journal of Sensors, Wireless Communications and Control, vol. 10, no. 1. pp. 25–36, 2020, doi: http://dx.doi.org/10.2174/2210327909666190129154802.

[21]M. Sedighimanesh, H. Zandhesami, and A. Sedighimanesh, "Presenting the Hybrid Algorithm of Honeybee-Harmony in Clustering and Routing of Wireless Sensor Networks," Int. J. Wirel Sensors. Commun. Control, vol. 9, no. 3, pp. 357–371, 2019.