Energy-Aware Clustering in the Internet of Things by Using the Genetic Algorithm

Mohammad Esmaeili[1], Shahram Jamali[2] and Hamed Shahbazi Fard[3]
[1]*Department of Computer Engineering, Science and Research Branch of Islamic Azad University, Ardabil, Iran.*
[2]*Department of Computer Engineering, University of Mohaghegh Ardabili, Ardabil, Iran.*
[3]*Computer Networks Research Lab, Electrical Engineering Technologies Research Center, Sahand University of Technology, Sahand New Town, Tabriz, Iran.*
H_Shahbazi@sut.ac.ir

Abstract*— Internet of things (IoT) uses a lot of key technologies to collect different types of data around the world to make an intelligent and integrated whole. This concept can be as simple as a connection between a smartphone and a smart TV, or can be complex communications between the urban infrastructure and traffic monitoring systems. One of the most challenging issues in the IoT environment is how to make it scalable and energy-efficient with regard to its growing dimensions. Object clustering is a mechanism that increases scalability and provides energy efficiency by minimizing communication energy consumption. Since IoT is a large scale dynamic environment, clustering of its objects is a NP-Complete problem. This paper formulates energy-aware clustering of things as an optimization problem targeting an optimum point in which, the total consumed energy and communication cost are minimal. Then. it employs the Genetic Algorithm (GA) to solve this optimization problem by extracting the optimal number of clusters as well as the members of each cluster. In this paper, a multi objective GA for clustering that has not premature convergence problem is used. In addition, for fast GA execution multiple implementation, considerations has been measured. Moreover, the consumed energy for received and sent data, node to node and node to BS distance have been considered as effective parameters in energy consumption formulation. Numerical simulation results show the efficiency of this method in terms of the consumed energy, network lifetime, the number of dead nodes and load balancing.

***Index Terms*— Clustering; Energy-Aware; Genetic Algorithm; Internet of Things.**

I. INTRODUCTION

The phrase of Internet of Things (IoT) was initially used by Kevin Ashton in 1999 to describe a world in which everything including human, animal and inanimate objects would have digital identity for themselves and are capable to deliver data via communication networks like internet or intranet. Moreover, objects could be controlled and managed by smart devices like smartphones, Tablets, and computers [1] (see Fig. 1). Presumably, the transition towards the next generation networks [2] has led to the emergence of new concepts and novel demands. In relation to this, the appearance of IoT is one of the thousand results of the internet expansion and development in wireless technologies and micro-electromechanical systems. One of the most important features of IoT is facilitating the connection to the internet for all kind of electrical objects. This implies that different home appliances (such as watches, electric lamps, refrigerators, etc.) could be remotely controlled, turned on and off through the Internet. The process of data transition in IoT environment no longer require the interaction of "human by human" or "human by computer" [3]: the data is transmitted automatically based on the default configuration and in certain time periods (permanently or momentarily).

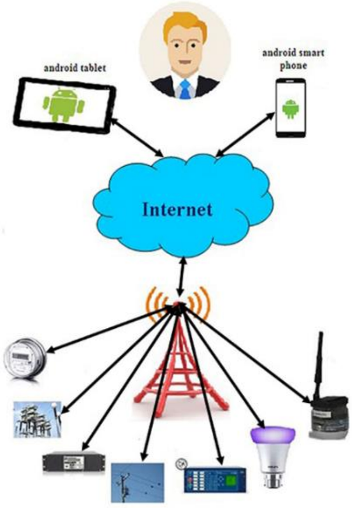


Figure 1: Control and management of things in the IoT

Numerous published literatures and scientific articles have investigated diverse aspects and features of IoT e.g. [1] enabling technologies for main communication, wired and wireless networks as well as elements of wireless sensor networks. Enabling technologies for IoT, especially by using RFID and its potential has been evaluated in [4]. Enabling technologies for IoT such as ubiquitous computing, embedded devices, sensor networks, internet protocols, and many others have led to conversion of IoT from conventional method to smart one [5]. IoT can be used in health section too, for example the use of IoT for wireless devices in hospitals that employs 6LoWPAN/IEEE 802.15.4, Bluetooth and NFC for mHealth and eHealth programs [6, 7]. Various business models and architectures are proposed for IoT and its potential in economic is massive [8-10]. Architectures of IoT and challenges for application development in IoT domain have been considered in [11].

Power saving is crucial to increase the nodes' life in IoT. To achieve optimum power consumption and consequently better results in term of network life, power saving procedures should be developed [12]. With regards to the nature of the objects in IoT, the power source (battery) of the objects in IoT is limited [13]. On the other hand, high scale data request from adjacent devices may lead to receiving thousands of conflicting messages [14]. Therefore, controlling the scalability, communication among the objects, complexity and increase of the energy consumption in communications need to be addressed in an appropriate way [15, 16]. An efficient mechanism should consider some factors like load balancing, reliability, quality of service, high stability and provision of algorithm with low complexity [17-19].

Presumably, IoT will be the largest engineered system ever created by mankind. Accordingly, among all the above-mentioned features, scalability and minimized energy consumption play more important roles than the others [13, 14, 16]. Hence, this paper proposes a clustering algorithm to overcome these challenges. It has been shown that clustering of such environments is a NP-Complete problem. Hence, in this paper, the clustering algorithm is redefined as an optimization problem and then the Genetic Algorithm (GA) is employed to solve this problem leading to the clustering of IoT nodes and enhancing the scalability and the load balancing in the IoT environment.

In comparison with the other methods of optimization, the reasons for the use of genetic algorithms for clustering and the GA have some important benefits [20]:

- Parallel processing is one of the most important superiority of the GA. This means that instead of a variable, we can grow a whole of population and moving towards achieving the optimal point simultaneously. Thus, the speed of convergence is very high.
- This method can be used to optimize the problems that are not well-behaved to their parameters (For example, problems with high oscillations or functions, which are highly not linear)
- This method is ideal for optimization of discrete quantity problems.
- In this method, it doesn't matter whether the function should be differentiable or not, whereas in most of the other methods, the optimization is based on the different derivatives of the function.

Other algorithms, such as the Partial Swarm Optimization (PSO), K-Means and Ant Colony Algorithm do not have these advantages.

- Ant Colony Algorithm: According to random selection of the objects by the ants and the multiplicity of repetitions, the result may be placed in the optimal local livestock, which may result in stagnation [21].
- Partial Swarm Optimization: There is no evolutionary operator and may converge prematurely. Further, the GA has the advantage of saving new solutions (A subset of the best ones) [22].
- K-Means: Dependents are on the initial values of cluster centers. Thus, a poor selection of them may result in falling in the trap of the local optimum. On the other hand, increasing the volume of data may require the algorithm to consume more time to find a local optimum [23].

The GA, which is used in this method is a multi-purpose algorithm for clustering that does not have premature convergence problem. Further, the fact that the IoT communication is bidirectional (objects are capable of sending sensed events and receiving executed commands) both of consumed energy for received and sent data in the energy consumption formula has been considered. In addition, node to node and node to CH distance has been taken into account as an effective factor in the energy consumption.

The rest of this paper is organized as follows: in Section 2, related works are presented. In Section 3 and 4, the problem formulation and the proposed method have been described, respectively. Section 5 shows the simulation results and assesses the proposed mechanism. Finally, Section 6 concludes the paper.

II. Related Works

Clustering divides the network nodes into multiple groups, wherein the nodes in each group are geographically close to each other. Each cluster has a head of cluster that is responsible for controlling all activities of the group including the transmission, data aggregation, management and maintenance of structure [24]. Energy consumption, network life and scalability of the network could be improved by clustering [25]. To increase the network life time, the GA classify the nodes in independent cluster sets, leading to the elimination of unnecessary communications among nodes. Therefore, network life rises [26, 27]. In [28], an intelligent cluster [29] is presented in wireless sensors network. In this method, a GA is used to minimize communication gaps using a binary representation that is a representative of one sensor node. According to [28], the limitation of this method is that it does not consider the binary method, and the sensor node may substitute between the active and inactive state frequently. An efficient plan of cluster-energy based on network optimization by the GA is proposed in [30]. This plan supposes the network area as a virtual network similar to cell packaging and consider each cell as a cluster. The GA has been used to divide the nodes among grids equally to guarantee load balancing, hence resulting in the increase of the network life. When one of the groups moves from the source to the Base Station (BS), the change of pattern in the energy consumption and nodes, which are closer to BS would have more chances of data transmission. Hence, the energy consumption of network will rise: However, this condition is considered in this model. [31] presented an algorithm that uses chaos logic according to GA in a way that each node has calculated its chance to be a Cluster Head (CH) based on energy, density, and centrality. Nodes that have high energy inform the BS to give a suggestion to nominate a potential CH. BS chooses a cluster head by using GA based on a cluttered and chaotic reasoning. Although this method uses information, such as remaining energy, node density and centrality which guarantee the network life, it suffers from the problem of the increase communication between the nodes and base station as another form of energy loss. In [32, 33], an energy efficient algorithm using GA is provided, which is based on clustering (GABEEC). However, this approach tries to maximize network life by minimizing communication distance, but encounter an increase in the overhead data for delivering information about the remaining energy of nodes to the BS. The increase of signalling leads to energy depletion and results in the reduction of network life. A new clustering algorithm with cluster members (NCACM) for avoiding dissipation of energy is proposed in [34] to reduce the energy consumption and to extend the network lifetime. The authors determined a confidence value for any sensing node that wants to be a CH with some parameters, such as the remaining energy of nodes, distances between the nodes, and distances between the CHs in each round. The critical problem of these kind of approaches is that the cluster size is not uniform and some clusters consist of a huge number of sensor nodes in a large area. Thus, the network load is unbalanced and some sensor nodes have to transfer data through longer distances.

The above-mentioned clustering methods have mostly considered one aspect of theorem (nodes have been considered fixed or dynamic) and the problem of scalability remains as an unsolved challenge. Since IoT includes home appliances, such as refrigerator, which has a permanent power supply, both of the battery sensor nodes and non-battery sensor nodes (power plugged) should be considered. Hence, in clustering these objects, a balance must be established between homogeneous clusters that have the same objects and heterogeneous clusters that have different objects in terms of power supply type. If a cluster had only non-battery objects, then optimization of energy consumption would be vain [35]. Thus, clustering of IoT objects with huge diversity (being fixed or mobile, battery less or self-powered nodes) and controlling the scalability and their energy consumption are important and challenging issues for research. To address these issues, a clustering algorithm that considers scalability and communication cost among objects is suggested.

III. Problem Formulation

Consider the typical architecture of the IoT environment, given in Figure 2. As shown in this figure, there are some geographically placed BSs to cover all the nodes of the IoT ecosystem. This is due to the limited amount of energy of things, although the energy is one of the most important challenges of IoT. Obviously, the direct communication between nodes and their corresponding BS leads to the acceleration of energy depletion in the nodes [36]. Hence, clustering-based communication is an efficient solution for this purpose. To design a formula for this energy-efficient clustering, the following factors are considered: the distance of each node from the BS, the distance of each node from the other nodes inside the same cluster, the distance of the CH from the BS and the distance between CHs.

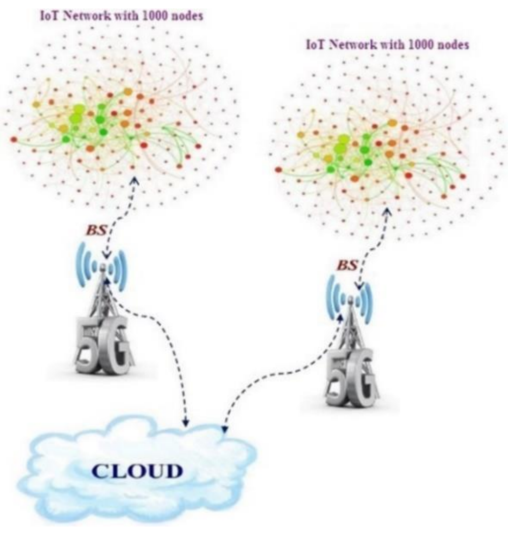


Figure 2: Proposed architecture for the IoT

The definition of notations used in this section is as follows:

- N: the set of sensor nodes
- d_{ij}: the distance from sensor node $i \in N$ to sensor node $j \in N$ (m),
- f_i: the distance from sensor node $i \in N$ to the BS (m),
- b_i: the battery level of sensor node $i \in N$ (J),
- l: the data size sent by a sensor node (bit),
- E: the coefficient for the radio dissipate to run the transmitter or receiver circuitry (J/bit),
- EDA: the coefficient for data aggregation (J/bit),
- n: the number of sensor nodes which have positive battery level,
- α: the parameter to determine CH candidates ($0 < \alpha \leq 1$),
- S_i: 0 if sensor node i has a positive battery level and 1 otherwise,
- d_{is}: the distance from the node i to the BS node s,
- d_{ih}: the distance from node i to the CH h,
- d_{hs}: the distance from the CH h to the BS node s.

The amount of energy used for data transmission is defined by two models, depending on the distance between the sensor nodes. If the distance is less than the threshold distance (i.e. d_0), the free space model is used [36, 37]. In all other conditions, the multi-path model is used [36, 37]. In the former model, energy consumption is proportional to the squared distance, and in the later model, energy consumption is proportional to biquadrate distance. The amount of energy used for data transmission from sensor node i to sensor node j is given by,

$$D_{ij} = \begin{cases} E + \mathcal{E}_{fs} d_{ij}^2 (\, d_{ij} < d_0) \\ E + \mathcal{E}_{mp} d_{ij}^4 (\, d_{ij} \geq \, d_0) \end{cases} \tag{1}$$

$$F_i = \begin{cases} E + \mathcal{E}_{fs} f_i^2 (\, f_i < d_0) \\ E + \mathcal{E}_{mp} f_i^4 (\, f_i \geq \, d_0) \end{cases} \tag{2}$$

where: ε_{fs} = Coefficient of free space model (pJ/bit/m2)
ε_{mp}= Coefficient of multi-path model (pJ/bit/m4)

Amplifier energy used for data reception from a sensor node is one E. Note that E is a fixed energy consumption. Since every sensor node, including those serving as CHs sends one bit of data, it consumes ld_{ij} or lf_i joules of energy. Consequently, decision variables are introduced:

- x_i: binary variable such that x_i=1, if sensor node $i \in N$ is selected as a CH, and otherwise x_i=0.
- y_{ij}: binary variable such that y_{ij}=1, if sensor node $i \in N$ belongs to the cluster, where sensor node $j \in N$ is a CH, and otherwise y_{ij}=0

To improve the network efficiency, a new formulation for the clustering problem of sensor networks is proposed. The clustering problem is formulated as the following integer programming problem by maximizing Equation (3) subject to Equations (4) to (6):

$$\frac{\sum_{i\in N}(b_i-(l\sum_{j\in N}D_{ij}y_{ij}+lF_ix_i)-lE\sum_{j\in N}y_{ji}-lE_{DA}\sum_{j\in N}y_{ji})}{\sum_{i=1}^{m}D_{is}+\sum_{i=1}^{k}d_{ih}+d_{hs}} \tag{3}$$

$$x_i + \sum_{j\in N} y_{ij} + s_i = 1 \;\;, i \in N \tag{4}$$

$$\left(b_i - \frac{a}{n}\sum b_k\right) x_i \geq 0 \;, i \in N \tag{5}$$

$$y_{ij} = x_j \,,\; i, j \in N,\; x_i \in \{0, 1\},\, i \in N,\, y_{ij} \in \{0, 1\},\; i, j \in N \tag{6}$$

The objective of Equation (3) is to maximize the sum of sensor node battery levels after each iteration. Each term in brackets of objective Equation (3) is described as follows:

- The second term enclosed within parentheses represents the total energy consumption of sensor node i used for data transmission.
- The third and the fourth terms represent the energy consumptions of sensor node i used for data reception and for data aggregation, respectively.
- Below the fraction line, the first term represents the total direct distance from all nodes to the BS. The second term represents the total distance of the nodes to CH node and CH to BS.

From Equation (4), each sensor node either plays the role of a CH or sends data to the nearest CH as long as its battery level is positive. Equation (5) ensures that each sensor node which has at least α times as much as the average battery level of all live sensor nodes will be a candidate to be a CH. Equation (6) states that all of the nodes can receive data.

IV. INTELLIGENT CLUSTERING ALGORITHM

This section presents an IoT clustering scheme using the GA. It is a nature-inspired heuristic approach for generating high-quality solutions for optimization and search problems inspired by the process of natural evolution using operators such as mutation, crossover and selection [30]. GA starts with a population of candidate solutions and then employs some techniques such as crossover, mutation, selection and inheritance to reach to an optimal solution (see Figure 3).

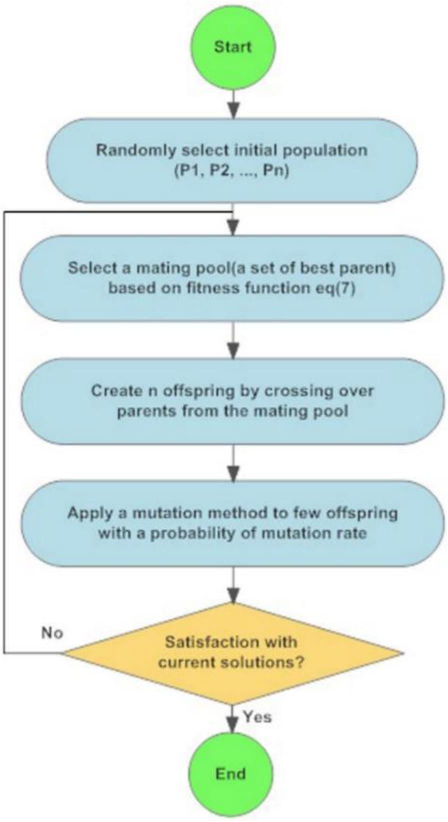


Figure 3: Flowchart of genetic algorithm [16]

To formulate the problem as a GA-based optimization problem, the following issues must be addressed:

A. Chromosome Representation

As shown in Figure 4, the binary representation is used for the chromosome. Each chromosome indicates a cluster, in which each bit corresponds to one thing or node. In this representation, 1 means that the corresponding node is a CH and 0 means that it is an ordinary node.

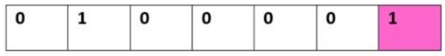


Figure 4: A sample chromosome representation

B. Initial Population

The initial population consists of 100 randomly generated individuals (chromosomes). The initial population is generated randomly so that the ordinary nodes and the CHs are distributed with a reasonable ratio in each chromosome.

C. Fitness Function

To add the fitness in this design, three factors are considered. The first factor is the total energy consumption for sending information from the network to the sink, denoted by V in Equation (7). Obviously this factor is desired to be reduced. The second factor is the total sum of distance between all nodes and the sink. Minimizing this distance leads to a reduced amount of energy consumption. Finally, the number of cluster heads, which is favorable to be as fewer as possible. Hence, the fitness function can be designed by maximizing Equation (7).

$$F = 100/V + ((DD - C))/DD + ((N - TCH))/N \tag{7}$$

where: DD = Total sum distance of all nodes to BS
C = Total sum distance of the nodes to the heads and the heads to the sink
N = Number of nodes
TCH = Number of the heads

Obviously, N and DD have fixed value, hence GA must search those values for V, TCH and C that maximize the F function.

D. Crossover

In this paper, one-point crossover is used. If an ordinary node becomes a cluster head after crossover, all other ordinary nodes should check if they are closer to this new CH. If so, they switch their membership to this new head. This new head is detached from its previous head. If a CH becomes an ordinary node, all of its members must find new CHs. Every node is either a CH or a member of a CH in the network. Figure 5 gives an illustrative example.

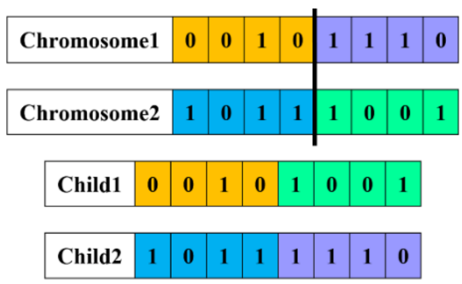


Figure 5: A Crossover over chromosomes 1and 2 to generate children 1 and 2

E. Mutation

The mutation operator is applied to each bit of an individual chromosome with a probability of mutation rate. When applied, a bit whose value is 0 is mutated into 1 and vice versa as shown in Figure 6.

Before Mutation	**0**	**0**	**0**	**1**	**0**	**1**
After Mutation	**0**	**1**	**0**	**0**	**1**	**0**

Figure 6: A sample chromosome mutation

F. Selection

The selection process determines which of the chromosomes from the current population should be selected for crossover to create new chromosomes. The combined population, consisting of the new chromosomes and the existing population will be the basis for the next selection. The chromosomes with better fitness values have better chances of selection. In Roulette-Wheel, [32, 33] which is used in this paper, chromosomes with highest fitness will be selected for making new offspring. Then, among these selected chromosomes, the ones with less fitness than others will be removed and new offspring would be replaced with the former ones.

Based on these principles and functions, the general description of the proposed method, which is called the Genetic Algorithm for Clustering of IoT (GAC-IoT) is given in Figure 7. As shown in this Pseudo Code, the algorithm is terminated after a specific number of iteration i.e. Max-Iteration which has been set to 500 in this paper or repeated values in three consequent iterations. In GA, a population consists of some chromosomes. The best chromosome is used to produce the next generation. Based on the fitness function, next generation is produced. Initially, each parameter of fitness function has a default value, which is updated after selecting the best chromosome and producing the next generation. It is the result of detecting suitable clusters and their status for energy efficiency.

Pseudo code for GAC-IoT

```
1.  Input
    Number of Nodes, Initial Energy, Population size, Mutation rate, Max-Iteration.
2.  Output
    Best solution x
3.  Begin
    i= 0;
    Create initial population Randomly.
4.  While (i < Max-Iteration) do
     i = i+1;
     evaluate fitness for each individual. (Eq.(7))
     Select mating pool based on fitness function.
     Apply crossover and mutation and produce Offspring(i).
     If (Offspring(i)== Offspring(i-1)== Offspring(i-2)) then break.
5.  End while.
    Best soulution = Offspring(i)
    Return best solution x
6.  End.
```

Figure 7: Pseudo code for GAC-IoT

V. IMPLEMENTATION RESULTS EVALUATION

To implement the proposed algorithm, the MATLAB software package is used. In this numerical simulation, the setup given in Table 1 will be used.

In our simulation, the number of the nodes is set to 1000, initial energy of each node is 2J and max iteration number is 500. Note that these parameters have been chosen in a way similar to the parameters used in the analogous methods [38, 39].

Table 1
Simulation Parameters

Parameter	Value
Workspace area	(100*100) m^2
Node Numbers	1000
Initial Energy	2J
Pop size	100
Crossover type	One-Point
Mutation rate	0.4
Selection type	Roulette-Wheel
Max-Iteration	500

Defining a big size population for the initial population is beneficial for problem solving as it allows the GA to search for bigger space, leading to a better solution. However, it may result in the increase of the number of calculations needed for producing each generation, the time complexity, and the memory consumption. Hence, defining a suitable population size is an important factor for efficiency. With regard to numerous experiments, the population size of 100 is selected for the simulation.

The mutation operator is used to add diversity to the population and to extend the search space of the GA. Additionally, it prevents from premature convergence. Lower mutation rate can lead to low diversity of generations, while higher mutation may cause a huge distance among them. Therefore, in the GAC-IoT, the mutation rate is defined as 0.4. In the selection operator, which is the Roulette-Wheel type, chromosomes are selected based on their fitness and the better chromosomes have more chances to be chosen for reproduction.

On the other hand, this simulation adopts the clustering characteristics given in Table 2.

Table 2
The Clustering Algorithm Characteristics

Parameter	Value
Cluster size	Unequal
Cluster Count	Variable
Intra-Cluster Communication	Single-hop
Inter-Cluster Communication	Multi-hop
CH Mobility	Mobile
Mobility model	Time Variant Community Mobility Model (TVC)[40]
CH Node type	Heterogeneous
CH Role	Aggregation/fusion
Clustering Method	Distributed
CH Election	Fitness-based
Algorithm Complexity	Variable

Best to the authors' knowledge, a number of clusters in the previous works that propose clustering with GA were fixed and predefined, but in the GAC-IoT, it is determined dynamically. In each cluster, communications among nodes are single-hop and direct. On the other hand, they are multi-hop among different cluster nodes. Recognizing that the objects in the IoT environment are varied, the type of CH node is considered heterogeneously. The major role of CH is gathering data from nodes inside the cluster, processing them and sending to the BS. With regard to the advantages of decentralization like high speed and better performance, distributed clustering method is used. Moreover, the selection of CH in each node is performed based on the fitness function.

In order to compare the performance of the proposed methodology from the viewpoint of energy consumption with other methods, it is implemented in the same predefined conditions with other recommended methods. Further, the output of each implementation is stored and compared.

The BS broadcasts require details of the network to all nodes, including the query execution schedule, the number of CHs, the members of each cluster and the number of transmission for this configuration. The cluster formation state is completed when all nodes receive broadcasted data from the BS, and clusters are produced based on these data. Energy level of the nodes in the network is reduced by the communication and data transmission among them. Additionally, the CHs consume energy to receive data from nodes in the cluster, process them and return the results to the BS, considering the communications among nodes in the network are bidirectional (send and receive).

One iteration in the simulation means one-time network data transmission to all nodes. In each iteration, clustering is performed on new CHs, considering the best chromosomes are selected and the quality of population is improved. When an object sends information to the network, the nodes close to it and the distance between them is specified. Moreover, in each method, the distance between the node and the cluster head is calculated. Obviously, communications between the nodes result in reduction of energy of each node.

Note that the GA related calculations are performed by the BS and it has access to permanent power plug. In addition, we have taken into account some measures for fast GA execution; First of all, for the implementation of a population, a variable length array has been used, as it has more flexibility and boosts sorting, replacement and insertion functions. Secondly, each chromosome has been evaluated just one time and repeated computations of fitness function for each chromosome has been prevented because of its cost and time complexity. Moreover, for the implementation of GA rather than the logical programming languages that consume so much time and memory for calculations, we have used an object oriented language that eases error detection, comparison and other GA related functions. With regard to GA execution, the reduction of energy consumption of whole network is witnessed, considering calculations related to the process of sensed information by each cluster node are performed in that cluster by the selected CH.

The simulation results are shown in Figure 8 to 13. As shown in these figures, the proposed algorithm GAC-IoT has been compared with the GABEEC and NCACM algorithms. These comparison methods is adopted due to their similarity in simulation aspects and parameters. The GABEEC is selected because it uses the GA and the Roulette-Wheel, while the NCACM is chosen because it is a distributed method and the unequal size of the cluster.

Figure 8 shows the total residual energy in the network in different iterations drawn from the three methods. This amount starts with 2000 J and iteration number 500 and it reaches to about 290, 470 and 830 J for the GABEEC, NCACM and GAC-IoT respectively.

As shown in Figure 8, the increase number of iterations is due to the increase communications among nodes as they consume more energy, which result in the reduction of energy of each node, and finally the whole network. Under the GABEEC and NCACM, the reduction of energy has more speed, leading to the GAC-IoT to consume less amount of energy in comparison with the other algorithms.

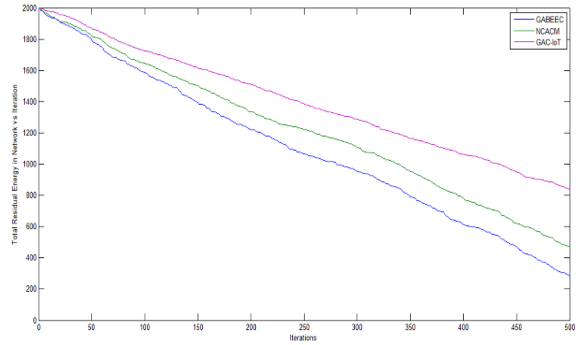
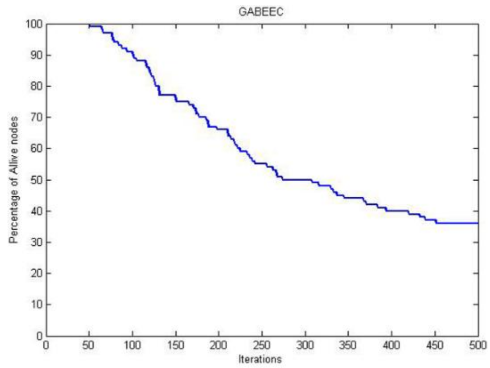
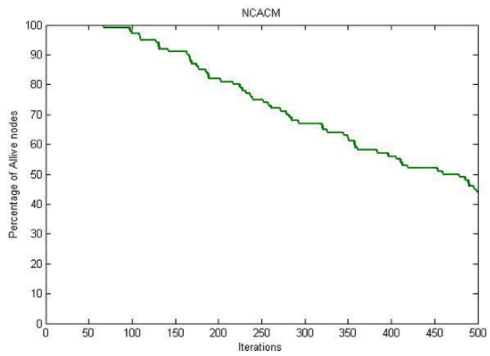


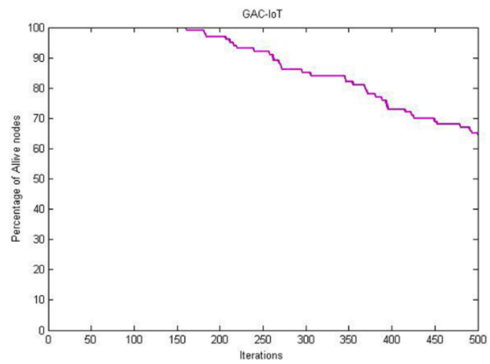
Figure 8: Comparison based on the total residual energy in the network



(a) GABEEC algorithm



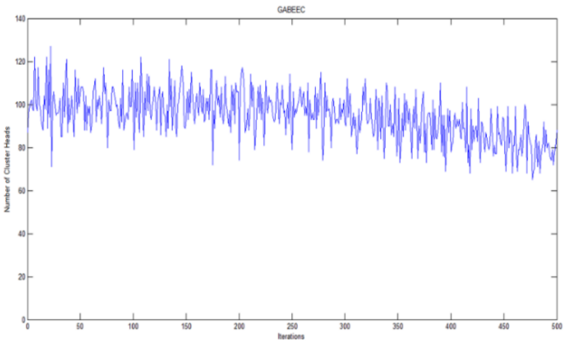
(b) NCACM algorithm



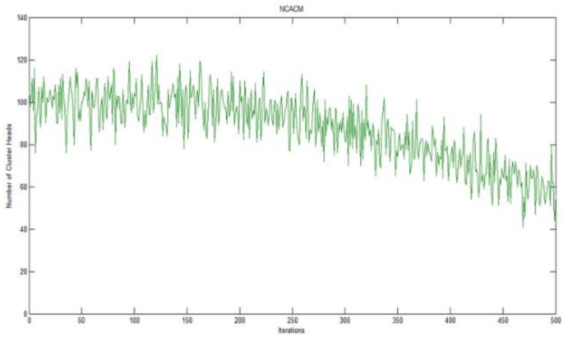
(c) GAC-IoT algorithm

Figure 9: Percentage of the alive nodes for 500 iterations

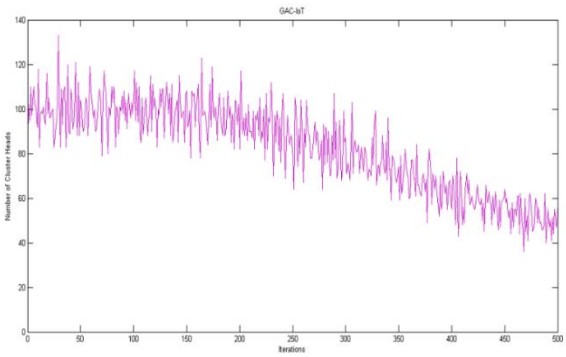
Figure 9 displays the percentage of alive nodes under the three methods for 500 iterations. The result of GABEEC method is shown in Figure 9(a). As shown in Figure 9(a), all nodes are alive until about iteration number 50. Then, it reduces and reaches to about 36% at iteration number 500. The result of the NCACM method is shown in Figure 9(b). It can be found that all nodes are alive until about iteration number 70. Then, it reduces and reaches to about 45% at iteration number 500. Finally, the result of GAC-IoT method is shown in Figure 9(c). In this method, all nodes are alive until about iteration number 160. It reduces and finally reaches to about 66% at iteration number 500. Therefore, there are more alive nodes after 500 iterations, when GAC-IoT manages the network.



(a) GABEEC algorithm



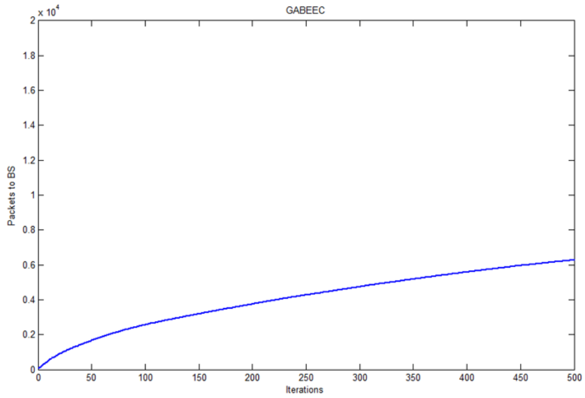
(b) NCACM algorithm



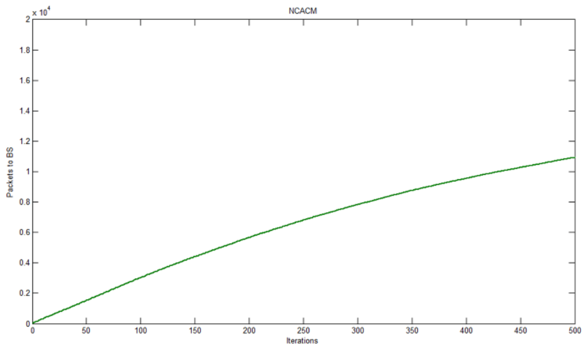
(c) GAC-IoT algorithm

Figure 10: Number of the cluster heads in different algorithms

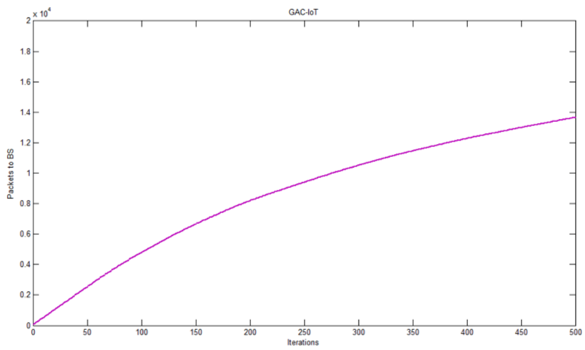
Figure 10 depicts the number of cluster heads under the three methods for 500 iterations. Overall, it can be seen that the number of cluster heads has a descending trend in all of the diagrams. The results of the GABEEC method is presented Figure 10(a). Initially, the number of cluster heads fluctuated between around 80 and 120, and then it experienced a marginal drop and finished around 80 at the end of simulation. The result of the NCACM method is shown in Figure 10(b). The trend of this diagram is similar to the GABEEC method, but its downward slope was sharper and at the last iterations, it fluctuated between 40 and 80. Finally, the result of the GAC-IoT method is shown in Figure 10(c). The analysis of this diagram reveals that the two previous methods have quite similar results; nevertheless, it declines gradually and eventually reaches to a number between 40 and 60.



(a) GABEEC algorithm



(b) NCACM algorithm

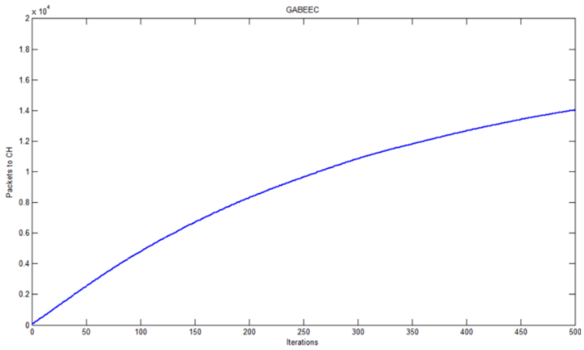


(c) GAC-IoT algorithm

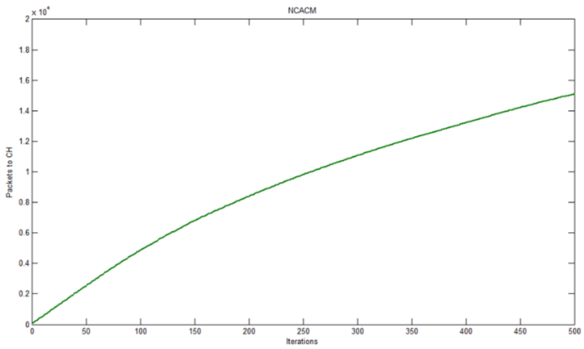
Figure 11: Number of packets that send to the BS in different algorithms

Figure 11 displays the number of packets sent to the BS based on the three methods for 500 iterations. shows the results of . The results of the GABEEC, NCACM, and GAC-IoT method are shown in Figure 11(a), Figure 11(b) and Figure 11(c). It is evident that in all of these three diagrams, the number of packets sent to the BS raised approximately from zero to 620, 1100 and for the GABEEC, NCACM and GAC-IoT methods, respectively.

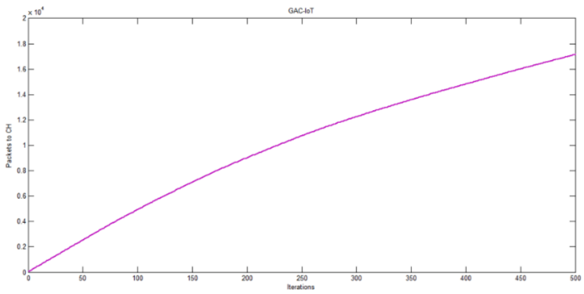
Figure 12 illustrates the number of packets sent to CH based on the three methods for 500 iterations. The results of the GABEEC, NCACM and GAC-IoT method are shown in Figure 12(a), Figure 12(b) and Figure 12(c), respectively. It is apparent from the diagrams that the number of packets sent to the CH increased from zero to around 1400, 1500 and 1700 for the GABEEC, NCACM and GAC-IoT methods, respectively.



(a) GABEEC algorithm



(b) NCACM algorithm



(c) GAC-IoT algorithm

Figure 12: Number of packets that send to the CH in different algorithms

In this section, the three methods are compared from the viewpoint of the ability to create the same load balancing in the entire network. Load balancing means that the establishment of balance and fairness between the nodes for the distribution of data traffic in the entire network, that is there are no condition that some nodes are unemployed while others are employed. The objective of the load balancing is finding efficient mapping of tasks among nodes in the network. On the other hand, each node has almost equal number of task to perform. Hence, the overall execution time reaches to the minimum.

Based on the simulation results in Figure 13, the GABEEC method has about 65%, while the NCACM has about 70%, and the GAC-IoT has about 90% of the network load that are balanced among clusters. Therefore, it can be found that the GAC-IoT offers better result in terms of load balancing in comparison with the other algorithms.

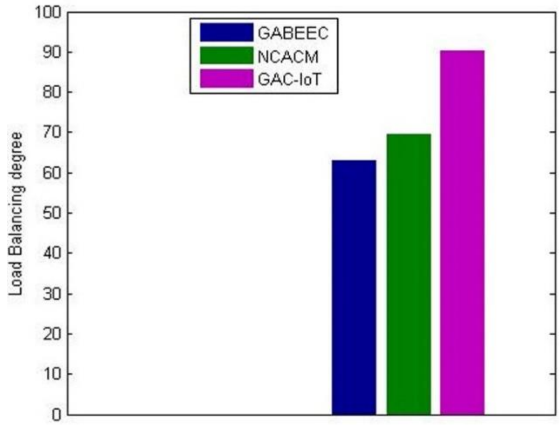


Figure 13: Load balancing factor in different algorithms

VI. CONCLUSION

In this paper, an energy-aware clustering algorithm for the Internet of Things has been proposed. The clustering problem of IoT has been redefined as an optimization problem and then the genetic algorithm has been used for to solve the problem. By using an intelligent multi-objective GA, the premature convergence problem has been prevented. Additionally, several extra factors have been considered for the formulation of the energy consumption: Both consumed energy for receiving and sending data, in which the important parameters are the distance from node to node and node to BS. Moreover, for increasing the speed of GA execution, multiple measures have been considered in the implementation phase. The simulation results showed that the proposed algorithm has better performance in comparison with previous algorithms. Specifically, the GABEEC and NCACM has better performance with respect to the consumed energy, network lifetime, number of dead nodes and load balancing. The reason is that, the GABEEC and NCACM need a vast amount of communications among nodes that lead to energy wastage. To address this issue, an energy-aware clustering method that reduces energy consumption of the IoT network has been proposed.

REFERENCES

[1] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," Computer networks, vol. 54, pp. 2787-2805, 2010.

[2] N. Maleki, H. S. Fard, M. Dadkhah, and M. Movassagh, "Scenarios for the Transition to NGN," IJCSNS, vol. 17, p. 115, 2017.

[3] J. Tang, Z. Zhou, J. Niu, and Q. Wang, "An energy efficient hierarchical clustering index tree for facilitating time-correlated region queries in the Internet of Things," Journal of Network and Computer Applications, vol. 40, pp. 1-11, 2014.

[4] D.-L. Yang, F. Liu, and Y.-D. Liang, "A survey of the internet of things," in Proceedings of the 1st International Conference on E-Business Intelligence (ICEBI2010), 2010.

[5] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," IEEE Communications Surveys & Tutorials, vol. 17, pp. 2347-2376, 2015.

[6] P. López, D. Fernández, A. J. Jara, and A. F. Skarmeta, "Survey of internet of things technologies for clinical environments," in Advanced

Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on, 2013, pp. 1349-1354.
[7] M. Bahrami and M. Esmaeili, "Convergence of IoT and WBSN for Smart Control of Patients Health," Networking and Communication Engineering, vol. 9, pp. 161-168, 2017.
[8] E. Fleisch, M. Weinberger, and F. Wortmann, "Business models and the internet of things," in Interoperability and Open-Source Solutions for the Internet of Things, ed: Springer, 2015, pp. 6-10.
[9] E. Bucherer and D. Uckelmann, "Business models for the internet of things," in Architecting the internet of things, ed: Springer, 2011, pp. 253-277.
[10] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, et al., "Internet of things in the 5G era: Enablers, architecture, and business models," IEEE Journal on Selected Areas in Communications, vol. 34, pp. 510-527, 2016.
[11] R. Khan, S. U. Khan, R. Zaheer, and S. Khan, "Future internet: the internet of things architecture, possible applications and key challenges," in Frontiers of Information Technology (FIT), 2012 10th International Conference on, 2012, pp. 257-260.
[12] J.-Y. Chang, "A distributed cluster computing energy-efficient routing scheme for internet of things systems," Wireless Personal Communications, vol. 82, pp. 757-776, 2015.
[13] S. Zhou, K.-J. Lin, and C.-S. Shih, "Device clustering for fault monitoring in Internet of Things systems," in Internet of Things (WF-IoT), 2015 IEEE 2nd World Forum on, 2015, pp. 228-233.
[14] Y. Sung, S. Lee, and M. Lee, "A Multi-Hop Clustering Mechanism for Scalable IoT Networks," Sensors, vol. 18, p. 961, 2018.
[15] M. Esmaeili and S. Jamali, "IoT based scheduling for energy saving in a wireless ecosystem," Wireless Communication, 2016.
[16] H. Lin, L. Wang, and R. Kong, "Energy efficient clustering protocol for large-scale sensor networks," IEEE Sensors Journal, vol. 15, pp. 7150-7160, 2015.
[17] D. Turgut, S. K. Das, R. Elmasri, and B. Turgut, "Optimizing clustering algorithm in mobile ad hoc networks using genetic algorithmic approach," in Global Telecommunications Conference, 2002. GLOBECOM'02. IEEE, 2002, pp. 62-66.
[18] H. S. Fard and A. G. Rahbar, "Physical constraint and load aware seamless handover for IPTV in wireless LANs," Computers & Electrical Engineering, vol. 56, pp. 222-242, 2016.
[19] T. Sánchez López, D. C. Ranasinghe, M. Harrison, and D. McFarlane, "Adding sense to the Internet of Things," Personal and Ubiquitous Computing, vol. 16, pp. 291-308, 2012.
[20] E.-G. Talbi, Metaheuristics: from design to implementation vol. 74: John Wiley & Sons, 2009.
[21] M. Dorigo and M. Birattari, "Ant colony optimization," in Encyclopedia of machine learning, ed: Springer, 2011, pp. 36-39.
[22] F. Marini and B. Walczak, "Particle swarm optimization (PSO). A tutorial," Chemometrics and Intelligent Laboratory Systems, vol. 149, pp. 153-165, 2015.
[23] M. E. Celebi, H. A. Kingravi, and P. A. Vela, "A comparative study of efficient initialization methods for the k-means clustering algorithm," Expert systems with applications, vol. 40, pp. 200-210, 2013.
[24] J.-L. Liu and C. V. Ravishankar, "LEACH-GA: Genetic algorithm-based energy-efficient adaptive clustering protocol for wireless sensor networks," International Journal of Machine Learning and Computing, vol. 1, p. 79, 2011.
[25] J. A. Stankovic, T. Abdelzaher, C. Lu, L. Sha, and J. C. Hou, "Real-time communication and coordination in embedded sensor networks," Proceedings of the IEEE, vol. 91, pp. 1002-1022, 2003.
[26] S. Hussain, A. W. Matin, and O. Islam, "Genetic algorithm for energy efficient clusters in wireless sensor networks," in Information Technology, 2007. ITNG'07. Fourth International Conference on, 2007, pp. 147-154.
[27] A. Norouzi and A. H. Zaim, "Genetic algorithm application in optimization of wireless sensor networks," The Scientific World Journal, vol. 2014, 2014.
[28] J. Jia, X. Wu, J. Chen, and X. Wang, "Exploiting sensor redistribution for eliminating the energy hole problem in mobile sensor networks," EURASIP Journal on Wireless Communications and Networking, vol. 2012, p. 68, 2012.
[29] S. Jamali and P. Jafarzadeh, "An intelligent intrusion detection system by using hierarchically structured learning automata," Neural Computing and Applications, vol. 28, pp. 1001-1008, 2017.
[30] L. Kong, J.-S. Pan, V. Snášel, P.-W. Tsai, and T.-W. Sung, "An energy-aware routing protocol for wireless sensor network based on genetic algorithm," Telecommunication Systems, vol. 67, pp. 451-463, 2018.
[31] W. Mardini, Y. Khamayseh, M. B. Yassein, and M. H. Khatatbeh, "Mining Internet of Things for intelligent objects using genetic algorithm," Computers & Electrical Engineering, vol. 66, pp. 423-434, 2018.
[32] S. Bayraklı and S. Z. Erdogan, "Genetic algorithm based energy efficient clusters (gabeec) in wireless sensor networks," Procedia Computer Science, vol. 10, pp. 247-254, 2012.
[33] M. Esmaeili and S. Jamali, "A Survey: Optimization of Energy Consumption by using the Genetic Algorithm in WSN based Internet of Things," CiiT International Journal of Wireless Communication, 2016.
[34] S. Babaie, A. K. Zadeh, and M. G. Amiri, "The new clustering algorithm with cluster members bounds for energy dissipation avoidance in wireless sensor network," in Computer Design and Applications (ICCDA), 2010 International Conference on, 2010, pp. V2-613-V2-617.
[35] T. Ducrocq, N. Mitton, and M. Hauspie, "Energy-based clustering for wireless sensor network lifetime optimization," in Wireless Communications and Networking Conference (WCNC), 2013 IEEE, 2013, pp. 968-973.
[36] S. Jamali, L. Rezaei, and S. J. Gudakahriz, "An energy-efficient routing protocol for MANETs: a particle swarm optimization approach," Journal of applied research and technology, vol. 11, pp. 803-812, 2013.
[37] D. Tse and P. Viswanath, Fundamentals of wireless communication: Cambridge university press, 2005.
[38] E. Heidari and A. Movaghar, "An efficient method based on genetic algorithms to solve sensor network optimization problem," arXiv preprint arXiv:1104.0355, 2011.
[39] V. Vashishth, A. Chhabra, A. Khanna, D. K. Sharma, and J. Singh, "An Energy Efficient Routing Protocol for Wireless Internet-of-Things Sensor Networks," arXiv preprint arXiv:1808.01039, 2018.
[40] W.-J. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, "Modeling spatial and temporal dependencies of user mobility in wireless mobile networks," IEEE/ACM Transactions on Networking (ToN), vol. 17, no. 5, pp. 1564-1577, 2009.