CrossMark

Dynamic Multi-hop Clustering in a Wireless Sensor Network: Performance Improvement

Mohamed Elhoseny[1] · Ahmed Farouk[2] · Nanrun Zhou[3] ·
Ming-Ming Wang[4] · Soliman Abdalla[5] · Josep Batle[6]

Abstract A cluster-based model is preferable in wireless sensor network due to its ability to reduce energy consumption. However, managing the nodes inside the cluster in a dynamic environment is an open challenge. Selecting the cluster heads (CHs) is a cumbersome process that greatly affects the network performance. Although there are several studies that propose CH selection methods, most of them are not appropriate for a dynamic clustering environment. To avoid this problem, several methods were proposed based on intelligent algorithms such as fuzzy logic, genetic algorithm (GA), and neural networks. However, these algorithms work better within a single-hop clustering model framework, and the network lifetime constitutes a big issue in case of multi-hop clustering environments. This paper introduces a new CH selection method based on GA for both single-hop and the multi-hop cluster models. The proposed method is designed to meet the requirements of dynamic environments by electing the CH based on six main features, namely, (1) the remaining energy, (2) the consumed energy, (3) the number of nearby neighbors, (4) the energy aware distance, (5) the node vulnerability, and (6) the degree of mobility. We shall see how the corresponding results show that the proposed algorithm greatly extends the network lifetime.

Keywords Cluster head · Wireless sensor network · Genetic algorithm · Multi-hop model

✉ Mohamed Elhoseny
   mohamed_elhoseny@mans.edu.eg

[1] Faculty of Computers and Information Sciences, Mansoura University, Mansoura, Egypt

[2] Department of Information Technology, Alzahra Collage for Women, Muscat, Oman

[3] Department of Electronic Information Engineering, Nanchang University, Nanchang 330031, China

[4] Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science and Technology, Nanjing 210044, China

[5] Department of Physics, King Abdulaziz University, Jeddah, Saudi Arabia

[6] Department of Physics, Universitat de les Illes Balears, 07122 Palmade Mallorca, Balearic Islands, Spain

1 Introduction

Wireless sensor network (WSN) has been widely used in many fields including health care, transportation management, military surveillance, etc. The battery-operated sensors are placed in open fields without human attendance and they acquire data contiguously over time. Through the wireless network, this information is transmitted and aggregated at a server located far away from the deployed sensors. Sensors are usually energy constrained and recharging battery is often costly or economically unfavorable. Hence, the management of sensors is an important factor for extending the network availability. Besides minimizing the energy consumption of data acquisition by advancing the sensing technology, it is preferable to balance the energy level of sensor nodes so that the entire network achieves an optimum availability.

In order to enhance the network availability and scalability, several network topologies and management methods have been devised, among which grouping sensors into clusters has demonstrated many advantages such as enabling smaller routing table, conserving communication bandwidth, avoiding redundant exchange of messages, and reducing maintenance overhead [1–3]. In a clustering WSN, each cluster usually consists of at least one surrogate node, often referred to as the CH. The CH may be dynamically decided or preassigned by the network designer. The communication between the cluster and the base station is facilitated by this CH. The problem of choosing the CH in multi-hop clustering model is more complected than the single-hop model. This is because in a multi-hop model, the CH may need to choose another node as a parent node to connect it with the base station, whereas in a single-hop model we only need to determine the CH of each
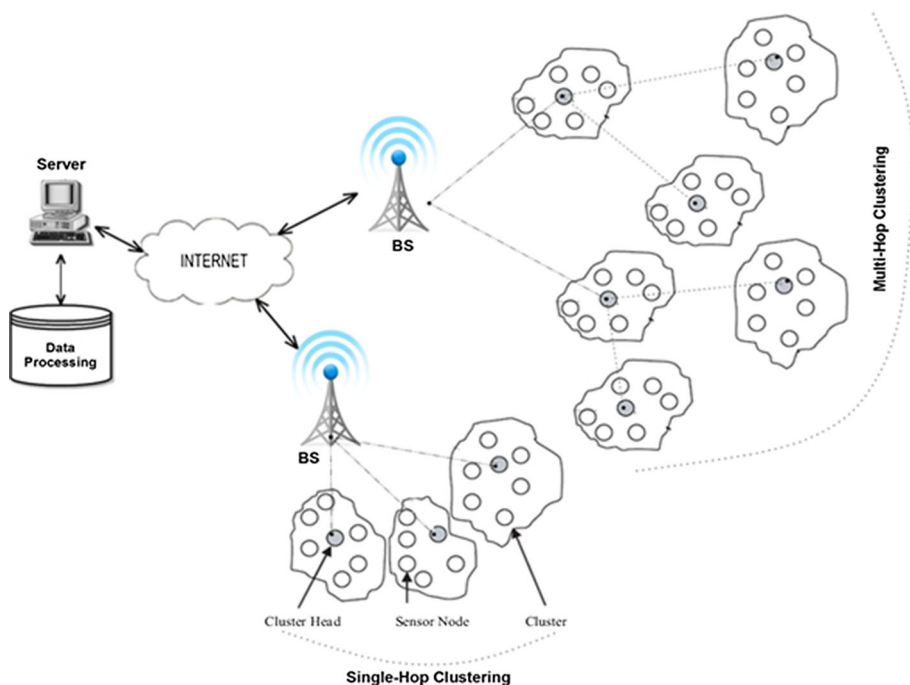

Fig. 1 Clustering models in WSN

cluster that will communicate directly with the base station. Figure 1 shows the difference between the two models. To balance the energy level of all nodes, selecting appropriate CHs becomes a key step. There exist many methods in the literature that address this issue [4–13]. In addition, a set of studies about energy efficiency of WSN have been provided [4, 5, 7, 8, 14], and several protocols and algorithms have been proposed in order to get an optimal power conservation in cluster-based WSN [9–13, 15]. However, most of them are not appropriate to work with WSN in dynamic environments, where sensors are unattended and it is very difficult to recharge their battery. In addition, they are not working well in the large sensor field in which the multi-hop clustering model is preferable. Low energy adaptive clustering hierarchy (LEACH) is one of the most promising cluster-based routing protocols [16]. LEACH is a distributed single-hop clustering algorithm. There, it is assumed that all nodes can perform long-distance transmission to the basestation (BS). Thus, selecting the CH depends on the energy used by all clusters. HEED is another cluster-based algorithm [17] based on the node's remaining energy, while a secondary parameter is the number of node neighbors. In this paper, we propose a new CH selection method based on GA, called DCH-GA, for both single-hop and the multi-hop cluster models. DCH-GA is designed to match the requirements of the dynamic environments by electing the CH based on six main features. These features are the remaining energy, the consumed energy, the number of nearby neighbors, EAD, the node vulnerability, and the degree of mobility. DCH-GA assigns each factor a degree of priority as explained in the Methodology section. To extend the network longevity, the count of clusters is determined automatically. Also, to avoid energy consumption at sensor nodes, DCH-GA runs the GA operations at the BS after each transmission round. This paper assumes that N sensor nodes are randomly distributed in a two-dimensional square field and that the sensor network has the following properties:

- All nodes are able to change their location from time to time.
- All nodes are able to send the data to the BS.
- The initial energy of all the nodes is the same.
- All nodes are energy-constrained.
- The network lifetime depends on the round first node die.
- The algorithm is working on multi-hop clustering model.
- Links are not symmetric. So, the transmission power level may differ between all pairs of nodes.
- All clusters are homogeneous, but they may have different sizes.
- There is only one BS located outside the network sensing field.

The remaining of this paper is organized as follows: Sect. 2.1 describes the GA and our hypotheses to elect the CH using it. Section 3 describes the proposed DCH-GA algorithm, and the parameters of the CH selection process. Finally, some conclusions are drawn in Sect. 4.

2 Related Work

Intelligent algorithms provide adaptive mechanisms that exhibit intelligent behavior in complex and dynamic environments like WSNs [18]. Various authors [18–22] discussed the routing protocols in cluster-base WSN based on intelligent algorithms such as RL, ACO, FL, GA, and NNs. On the other hand, many CH selection methods based on different techniques have been proposed. Local negotiated clustering algorithm (LNCA) presents a

novel clustering algorithm, which employs the similarity of nodes readings as an important criterion in cluster formation. ACE clusters the sensor network in a constant number of iterations using the node degree as the main parameter. In GA-WCA, the load balanced factor is considered as one of the weights along with a sum of distances from all neighbor nodes to CHs. LA2D-GA takes only the distance as a parameter to calculate the fitness function; also the representation of a chromosome in a two-dimensional grid represents valid statistics of a WSN [23].

Low energy adaptive clustering hierarchy (LEACH) [24] was proposed to create clusters in dynamic environment. However, the number of clusters must be determined in advance. Moreover, it can not be used in a multi-hop clustering environment. M-GEAR [25] is a clustering method in which the sensor nodes are divided into four logical regions depending on their locations in the network field. It needs an external gateway to be located at the center of the field while the BS is located out of the sensing field. M-GEAR selects CHs probabilistically in each region. Stable election protocol forms the network clusters by selecting sensor nodes as CHs using a weighted probability [26]. Lindsey and Raghavendra extended the LEACH method and proposed the PEGASIS algorithm [27], in which each node communicates only with the closest neighbor and takes turns when transmitting to the base station. The threshold sensitive stable election protocol (TSEP) [28] is another method in which there are three different levels of energies for the nodes, and the CHs selection is threshold-based. However, the TSEP method is designed for heterogeneous networks. Elbhiri et al. [29] developed a distributed energy-efficient clustering procedure that balances the CH selection by taking residual energy into consideration.

CHSCDP [30] is designed as an effective algorithm to select the optimal CHs among all nodes fairly, and tried to reduce the energy consumption to extend the lifetime of network. However, this method did not provide additional details about its ability to work in long area field. In addition, it assumes that the working environment is static and all its components are pre-known. In [31], a GA-based CH selection method is proposed for centralized clustering WSN. It aimed to create load balanced network than the traditional clustering algorithm. However, the GA running time still represents a big challenge for this method. LEACH-MAC [32] is proposed to form the optimum structure of WSN by forming its clusters. Although, LEACH-MAC achieved better network performance in terms of the required time to form the network clusters, it determines the count of the clusters in advance. To best of our knowledge, the count of the clusters can only be determined depending on the current environmental status of all sensor nodes. Thus, LEACH-MAC can not be used for the applications that need dynamic environments.

In Ref. [33], the Degree of connectivity is the main factor in choosing a CH. The degree of connectivity of a node, i.e. the number of its most immediate neighbors, is also a criterion that seems interesting to study. Intuitively, the more neighbors a sensor has, the more it seems to become an appropriate candidate as a CH, since a sensor with a low degree of connectivity might have little information retreival from its neighborhood to aggregate and to forward to the BS. In the initial phase, each sensor is involved into the neighborhood information exchanges (hello protocol), which allows it to determine its degree of connectivity and the location of BS. In EEUC [34], the distance between the node and the BS represents the main parameter for selecting the CH. The EEUC resulted in a network that is partitioned into clusters of unequal sizes, where clusters closer to BS have smaller sizes than those farther away from the BS.

2.1 The Proposed CH Selection Factors

Given the coordinates of the sensor nodes, the distance between two of them can be approximated with EAD. The value of EAD is computed depending of the Euclidean distance and take into account the environmental hindrances between the node and all other nodes in its cluster, such as trees or buildings. To calculate the EAD for the node v we assume that the distance between the BS and the node v is $ED_{v,BS}$. Also, we assume that the Euclidean distance is used to calculate the distance ED_{vc} between v and the center of its cluster. Each hindrance between v and any other nodes in the cluster is H and the total number of hindrances inside the cluster is nh. The number of nodes in the cluster is n. Each hindrance has a weight W which represents it's ability to consume the node energy. Let $W = \{w_1, w_2, w3, \ldots, w_{nh}\}$ where each H has an ID that is represented by the index i of w_i. So, the EAD can be represented as Eq. 1.

$$EAD_v = \frac{EAD_{vc} + \sum_{v=1}^{n} \sum_{j=1,v \neq j}^{n} \sum_{k=0}^{nh}(H_{vj}) * w + EAD_{v,BS}}{2} \qquad (1)$$

The consumed energy E, represents the energy needed to transfer the aggregated message from the cluster to the sink. According to the first order radio model [22], the consumed energy required to send a message with k bits length along a d distance is computed through Eq. 2 for a cluster with k member nodes, the cluster transfer energy is defined as follows:

$$E = \sum_{j=1}^{k} E_{T_{jh}} + kE_R + E_{T_{hs}} \qquad (2)$$

The first part of Eq. 2 shows the energy consumed to transmit messages from k member nodes to the CH. The second part shows the energy consumed by the CH to receive k messages from the member nodes. Finally, the third part represents the energy needed to transmit from the CH to the BS. The energy required ($E_{T_{ij}}$) to transmit a message of length l bits from a node i to a node j and the energy consumed in receiving the l-bit message (E_R) are given as follows:

$$E_{T_{ij}} = lE_e + l_{\epsilon_l}d_{ij}^4 \qquad (3)$$

$$E_{T_{ij}} = lE_e + l_{\epsilon_s}d_{ij}^2 \qquad (4)$$

$$E_R = lE_e + lE_{BF}, \qquad (5)$$

where d is the distance between nodes. However, for a long range transmission such as from a CH to the BS, the energy consumed is proportional to d^4. The E_{BF} represents the cost of beam forming approach to reduce the energy consumption. Also, the remaining battery power of a node v (donated by $RP(v)$) can be calculated as

$$E_{RP_v} = E_s - \sum_{t=0}^{c}(E_{Tv}(t) + E_{Rv}(t)), \qquad (6)$$

where Es is the initial energy of the node and c is the current time. The number of neighbors of the node v (donated by $N(v)$) is the count of nodes whose distance d from v

less than or equal to β. The β value differs from one application to another according to the coverage space of the WSN.

$$N(v) = |n_i| \Leftrightarrow d_{vn_i} \leq \beta : (v \neq n_i \& \beta > 0) \tag{7}$$

Computing the running average of the speed for every node till current time T gives a measure of its mobility [20]:

$$M_v = \frac{1}{T} \sum_{}^{T} \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}, \tag{8}$$

where (X_t, Y_t) and (X_{t-1}, Y_{t-1}) are the coordinates of the node v at times t and $t - 1$, respectively. Finally, and based on [35], a node that has high value of vulnerability means that failure of such node may lead to the disconnection of the entire network. Thus, we should avoid this node to be CHs. Vulnerability index VI of the node v can be calculated as Eq. 9.

$$VI_v = \frac{N_{before}^v}{N_{after}^v} \times \frac{L_{before}^v + 1}{L_{after}^v + 1} \tag{9}$$

Here, N_{before}^v is the number of nodes before removing v node, N_{after}^v is the number of nodes after removing v node. Likewise L_{before}^v is the number of levels before removing v node, and L_{after}^v is the number of levels after removing v node.

3 Methodology

3.1 Data Representation

Each factor is represented by a set of bits depending on its priority as shown in Table 1. These priorities represent one of the hypothesis we need to prove during this contribution using GA. At Table 1, the symbols H, L, and M are used to indicate to the priority degree whether it is *high*, *low*, or *medium* priority respectively. These priorities will affect the fitness value of the node. F_1, F_2, and F_5 are represented with only one bit because the domain value of each one has only two values. But F_3, F_4, and F_6 are represented by three bits each to get more flexibility because the domain contains four values. Also, this representation (of F_3, F_4, and F_6) produce logical results during the mutation process. An example of a chromosome that represents all features in direct binary representation (DBR) is shown at Fig. 2. The fitness function $f(x)$ evaluates the fitness value of each sensor node. Its value depends on the six features as shown in Eq. 10

$$f(x) = f(F_1, F_2, F_3, F_4, F_5, F_6) \tag{10}$$

So, the nodes with $F_1 = (0)$, $F_2 = (0)$, $F_3 = (000)$, $F_4 = (111)$, $F_5 = (1)$ and $F_6 = (000)$ have high fitness values and can be declared as CHs. And the node with the lowest fitness value can be ignored from the list. The fitness function of our problem can be set as in Eq. 11.

Table 1 Sensor node features

| Feature	Priority	No. of bits	Category	DBR	Value
F_1: Energy aware distance	L	1	Low	0	0
			High	1	1
F_2: Vulnerability index	L	1	Low	0	0
			High	1	1
F_3: Degree of mobility	H	3	Very low	000	0
			Low	001	1
			High	011	3
			Very high	111	7
F_4: Remaining battery power	H	3	Very low	000	0
			Low	001	1
			High	011	3
			Very high	111	7
F_5: Number of nearby heighbors	M	1	Low	0	0
			High	1	1
F_6: Consumed energy	H	3	Very low	000	0
			Low	001	1
			High	011	3
			Very high	111	7
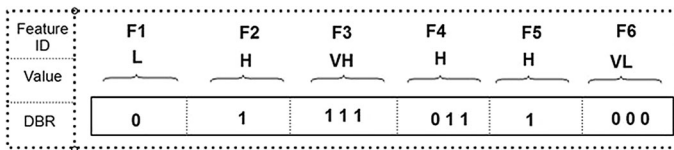
Fig. 2 Chromosome representation

$$f(x) = \frac{1 + F_4^2 + 2F_5}{1 + F_6^2 + F_3^2 + F_2 + F_1}, \tag{11}$$

where the remaining energy (F_4) has the highest priority, and the number of neighbors (F_5) has a medium priority. In addition to the fitness function, we shall calculate the probabilities of selecting $P(S_i)$, the expected count of select $(\overrightarrow{EC})$, and the actual count of select for each chromosome as shown in Eqs. 12 and 13:

$$p(s_i) = \frac{f(i)}{\sum_{x=1}^{n} f(x)} \tag{12}$$

$$\overrightarrow{EC} = \frac{f(i)}{\frac{1}{n}\sum_{x=1}^{n} f(x)}. \tag{13}$$

Recall that n is the number of the chromosomes in the population. For example, let us begin with a random initial population of size 4 (Table 2). A generation of the GA begins with reproduction. We select the mating pool of the next generation by spinning the weighted roulette wheel six times. So, the best node representation get more copies, the

Table 2 Example of initial population

Node index	Node content	F_1	F_2	F_3	F_4	F_5	F_6
1	010001111001	0	1	0	7	1	1
2	110110111000	1	1	3	3	1	0
3	101110011001	1	0	7	1	1	1
4	110011111011	1	1	1	7	1	3

average stay even, and the worst die off and will be excluded in further processing. To do that, we calculate $\overrightarrow{EC}$ and $P(S_i)$ as shown in Eqs. 12 and 13. The actual count value is the result of Roulette Wheel work. The Mate and Crossover values are randomly selected. Table 3 contains the next generation that resulSt from the GA work acting on our initial population. It shows that CHs have improved in the new population. Nodes that do not exist in the new generation will be ignored from the competition process in the future. In case a new node is not present in the current cluster, then it will be excluded also (Table 4).

3.2 Phase I: CH Selection Using GA

GAs are adaptive search methods that simulate some natural processes: selection, information, inheritance, random mutation and population dynamics [36]. A GA starts with a population of strings and thereafter generates successive populations of strings. A simple GA consists of three operators [19]: Crossover, Reproduction and Mutation. The chromosome of the GA contains all the building blocks to the solution of the problem at hand in a form that is suitable for the genetic operators and the fitness function. In our problem, each individual sensor node is represented by 6 features. Based on the CH selection factors that have been described previously, the CH selection algorithm is listed at Algorithm 1. In addition, its corresponding flowchart is shown at Fig. 3.

Algorithm 1 The Proposed DCH-GA Algorithm for CH Selection.

1: Initialize N, P_c, P_R, P_M, θ, and set $T = 1$.
2: population= create_population $(N,12)$ where 12 is the DBR of the 6 factors
3: **for** $i = 1, 2, \ldots, N$ **do**
4:     calculate $F_1, F_2, F_3, F_4, F_5, F_6$
5:     $f(x_i) = \frac{1+F_4^2+2F_5}{1+F_6^2+F_3^2+F_2+F_1}$
6:     $P(S_i) = \frac{f(i)}{\sum_{x=1}^{n} f(x)}$
7:     $count = $ roulette_wheel $(i, f(x))$
8:     $fx[i] = f(x_i)$
9:     $feature[i] = F_1, F_2, F_3, F_4, F_5, F_6$
10: **end for**
11: $f(x_t) = \sum_{i=0}^{N} f(x_i)$
12: **if** $f(x_t) - f(x_{t-1}) \leq \theta$ **then**
13:     CH $= return_{CH} (fx[], feature, population)$
14: **else**
15:     $T = T + 1$
16:     new_population $\Leftarrow$ next_generation($population$, P_c, P_R, P_M)
17:     population $\Leftarrow$ new_population
18:     Go To: 3
19: **end if**

Table 3 Example of using GA for CH selection

Node number(i)	$F(i)$	$P(S_i)$	EC	Actual count	Mating pool	Mate	Crossover point
1	17.334	0.773	3.093	2	0 1 0 0 0 1 1 1 1 0 0 1	1	5
2	1	0.044	0.178	1	1 1 0 1 1 0 1 1 1 0 0 0	4	3
3	0.077	0.003	0.013	0	0 1 0 0 0 1 1 1 1 0 0 1	1	5
4	4	0.178	0.713	1	1 1 0 0 1 1 1 1 1 0 1 1	2	3
Sum	22.4	1	4				
Avg	5.6	0.25	1				

Table 4 Example of using GA for CH selection

Node number	Node representation	F_1	F_2	F_3	F_4	F_5	F_6	$F(i)$
1	0 1 0 0 0 1 1 1 1 0 0 1	0	1	0	7	1	1	17.334
2	1 1 0 0 1 1 1 1 1 0 1 1	1	1	1	7	1	3	4
3	0 1 0 0 0 1 1 1 1 0 0 1	0	1	0	7	1	1	17.334
4	1 1 0 1 1 0 1 1 1 0 0 0	1	1	3	3	1	0	1
Sum								39.66
Avg								9.92

The creation of a population generates n number of nodes and represent each one in 12 bits. Returning the CH takes three arguments and searches for the best CH according to the fitness value. The next generation algorithm performs the crossover, mutation, and reproduction processes and returns a new generation of nodes. N, PC , PR, PM, θ, and T denote the population size, crossover probability, reproduction Probability, mutation probability, threshold, and time interval, respectively.

Figure 3 shows the main parts of the flowchart. It starts with generating the initial population of chromosomes of size n and calculates the values of the six features for each chromosome. Then it calculates the fitness function for each chromosome and applies the roulette wheel to compute the actual count of selecting the chromosome. After that the processes of mutation, reproduction, and crossover will performed according to their ratio. These processes will be repeated until the termination condition will be achieved. After that the CH will be selected according to the fitness value.

3.2.1 Phase II: Multi-hop Model Construction

The multi-hop clustering model is a special case of the clustering model in which a CH can not transmit the data directly to the BS for whatever reason. In this case, the CH tries to find another CH closer to the BS to work as a gateway between them. Now, assume that we have a WSN that contains a set of clusters that are formed by DCH-GA. Each cluster has a specific CH for a fixed round of data transmission. After each round, the BS selects a CH for each cluster depending on the available information regarding all nodes such as the remaining energy, location, and neighbors. Also, let us assume that some of CHs cannot transmit data directly to the BS if the distance between them is more than a certain threshold α. Each of them must transmit its data to an intermediate CH between it and the
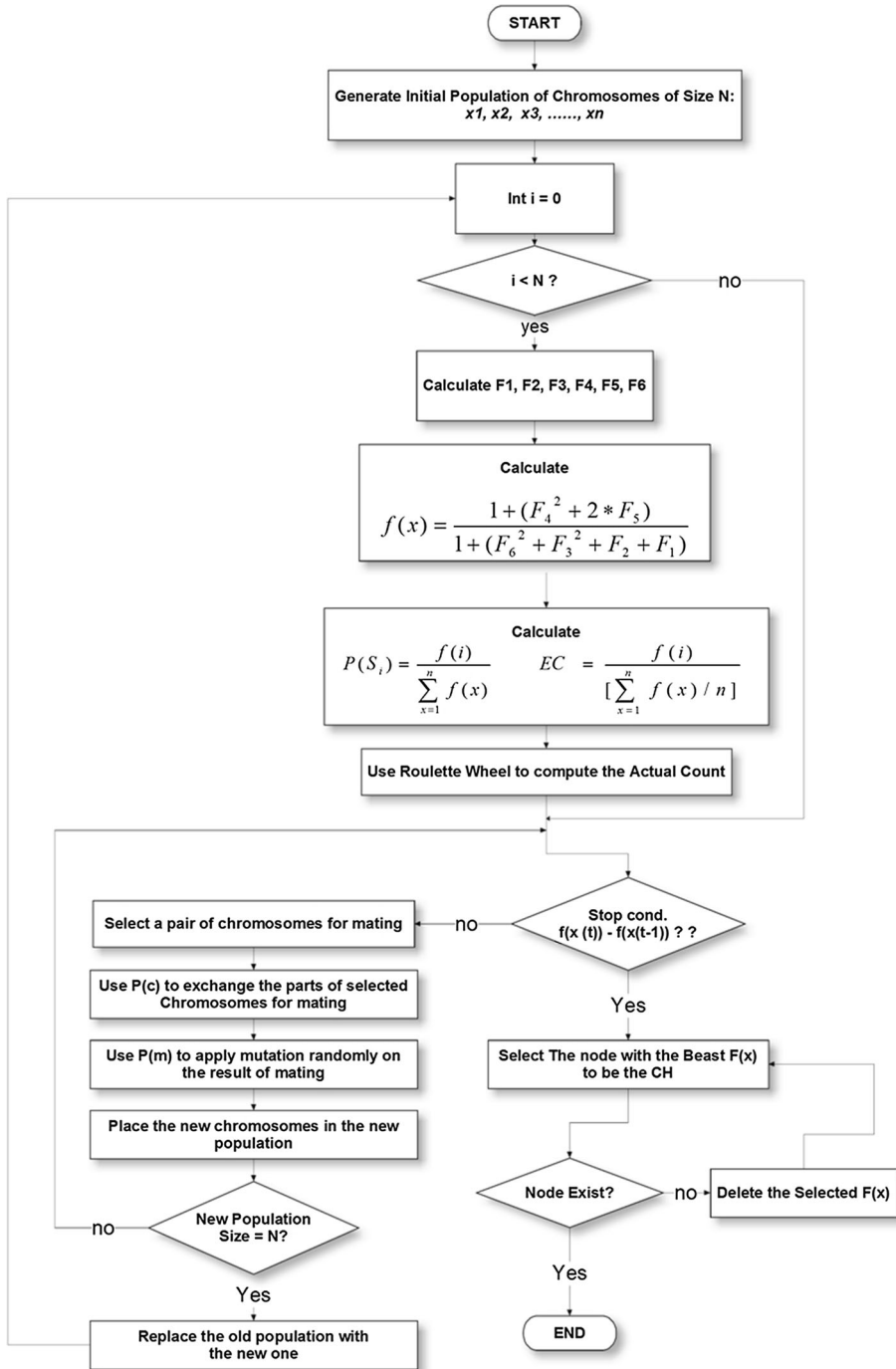
Fig. 3 Cluster model of WSN

BS. So, we propose a simple algorithm for constructing the multi-hop model after creating the clusters and determining the CH for each one. The goal of this algorithm is to extend the network lifetime by choosing the path with minimum energy consumption. We assume that this algorithm will run at the BS after each round of sending data to it from all CHs to reconstruct the network hierarchy. The BS will run it for the CH that is far away from it more than a threshold value α. It evaluates a CH n by combining $g(n)$, the cost to reach the CH, and $h(n)$, the cost to get from the CH to the BS. The algorithm selects the CH with minimum $f(n)$. Let us suppose that the CH that we need to determine the best path from it to the BS is given by i, the remaining energy of CH n is E_n, and the count of hindrances between i and n is $|H_i, n|$. The cost function can be calculated as Eqs. (14), (15), and (16).

$$f(n) = g(n) + h(n) \tag{14}$$

$$g(n) = ED_{i,n} + E_n + |H_i, n| \tag{15}$$

$$h(n) = ED_{n,BS} + |SN_n|, \tag{16}$$

where $|SN_n|$ refers to the counting of nodes in the cluster. This $|SN_n|$ is included in the calculation to indicate the traffic loads of the CH. The ED represents the expected distance between two nodes and is calculated based on the EAD. The ED can be calculated as the blow example at Fig. 4. After several iterations, the algorithm selects the path $\varrho[x, BS]$ from the CH x to the BS, which is the minimum energy path. The protocol assumes that each CH can reach any other within its environment as shown in Algorithm 2 where N is the count of the CHs. Figure 4 shows an example of a multi-hop intelligent algorithm. The example contains three clusters and one BS. For each cluster, one CH is defined along with its
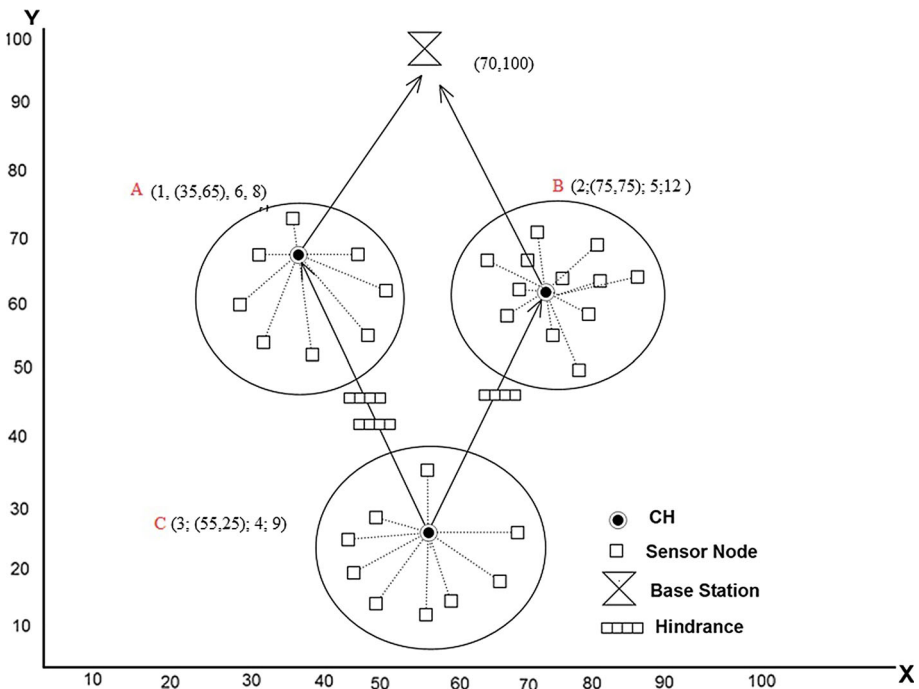

Fig. 4 An example of the proposed multi-hop intelligent algorithm

features. The features of the CH are ID, coordinates, Remaining Energy, and Number of Nodes in its cluster, respectively, as shown in the figure. Assume that the CH C cannot interact directly with the BS. It needs another CH to works as a mediator between them. Depending on its signal strength, it can select between A and B. Applying our proposed DCH-GA algorithm, C will calculate the cost function $f(A)$ and $f(B)$ for A and B, respectively as shown at Table 5. From Table 5, the node C will select the CH B to be its parent because it has the minimum cost value.

Algorithm 2 Multi-hop Clustering Algorithm

1: Initialize α;
2: **for** $x = 0, 1, \ldots, CH_s$ **do**
3:     **if** $ED_{x,BS} \leq \alpha$ **then**
4:         $\varrho[x, BS] = \Phi$, Where Φ represents the direct link between x and BS
5:     **else**
6:         **for** $i = 0, 2, \ldots, N$ **do**
7:             $h(i) = ED_{i,BS} + |SN_i|$
8:             $g(i) = ED_{i,x} + E_i + |H_i, x|$
9:             $f(i) = g(i) + h(i)$
10:        **end for**
11:        $\varrho[x, BS] = \text{Min}(f.\text{items})$, that is, the minimum distance between CH and all other CHs
12:    **end if**
13: **end for**
14: Return the Multi-hop Structure;

Given the coordinates of sensor nodes, ED between two of them can be approximated taking into account their remaining energy. The values of ED are computed by combining the Euclidean distance between the nodes and the environmental hindrances. The hindrances between the node and all other nodes in its cluster such as trees or buildings are taken into account. To calculate the ED between the node V and the node B, we assume that the Euclidean distance between B and V is D_{VB}. Also, we shall assume that the Euclidean distance is used to calculate the distance D_{VC} between V and its CH. Each hindrance between V and any other nodes in the cluster is H and the total number of hindrance inside the cluster is nh. The number of nodes in the cluster is n. As mentioned before, each hindrance has a weight W. By modified EAD, the ED can be represented as Eq. (17)

$$ED_V = D_{VC} + \sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} \sum_{i=0}^{nh} [H_{ij} w_i] + D_{VB} \tag{17}$$

In case of retrieving the distance between a CH i and the BS, ED can be calculated as a simple Euclidean distance ($ED_{i,BS} = Di, BS$). Algorithm 3 shows the main steps of the proposed ED algorithm.

Table 5 Cost Function for CHs. Where ED_{BS} is the EAD from the CH to the BS while ED_C is the EAD between it and the CH C

CH	E	|SN|	ED_{BS}	ED_C	|H|	g()	h()	f()
A	6	8	49.49	44.72	2	62.72	57.49	120.21
B	5	12	25.49	53.85	1	59.85	37.49	97.34

Algorithm 3 ED Algorithm

1: Get the location coordinates for: $B, V, C,$
2: Set $e = 0$;
3: Calculate $DVB = \sqrt{(x_B - x_V)^2 + (y_B - y_V)^2}$
4: Calculate $DVC = \sqrt{(x_C - x_V)^2 + (y_C - y_V)^2}$
5: **for** $i = 1, 2, \ldots, N$ **do**
6:   **for** $j = 1, 2, \ldots, N$ **do**
7:     **for** $k = 1, 2, \ldots, nh$ **do**
8:       $e = e + W[k] * H_{ij}$
9:     **end for**
10:   **end for**
11: **end for**
12: $ED = (e + ED_{V,BS} + ED_{V,C})$
13: return ED

Table 6 Network properties

| Network parameters	Value
Number of sensors	100
Initial energy of each sensor	0.25/0.5/1.0 J
Idle state energy consumption	50 nJ/bit
Data aggregation energy consumption	5 nJ/bit/report
Amplification energy (CH to BS) $d \geq d_0$	10 pJ/bit/m^2
Amplification energy (CH to BS) $d < d_0$	0.0013 pJ/bit/m^2
Amplification energy (Node to CH) $d \geq d_1$	$E_{fs}/10 = E_{fs1}$
Amplification energy (Node to CH) $d < d_1$	$E_{mp}/10 = E_{mp1}$
Packet size	400 bit
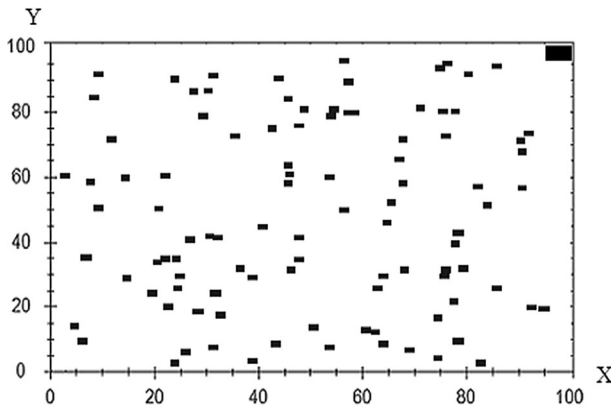
Fig. 5 An example of sensor and base station placement

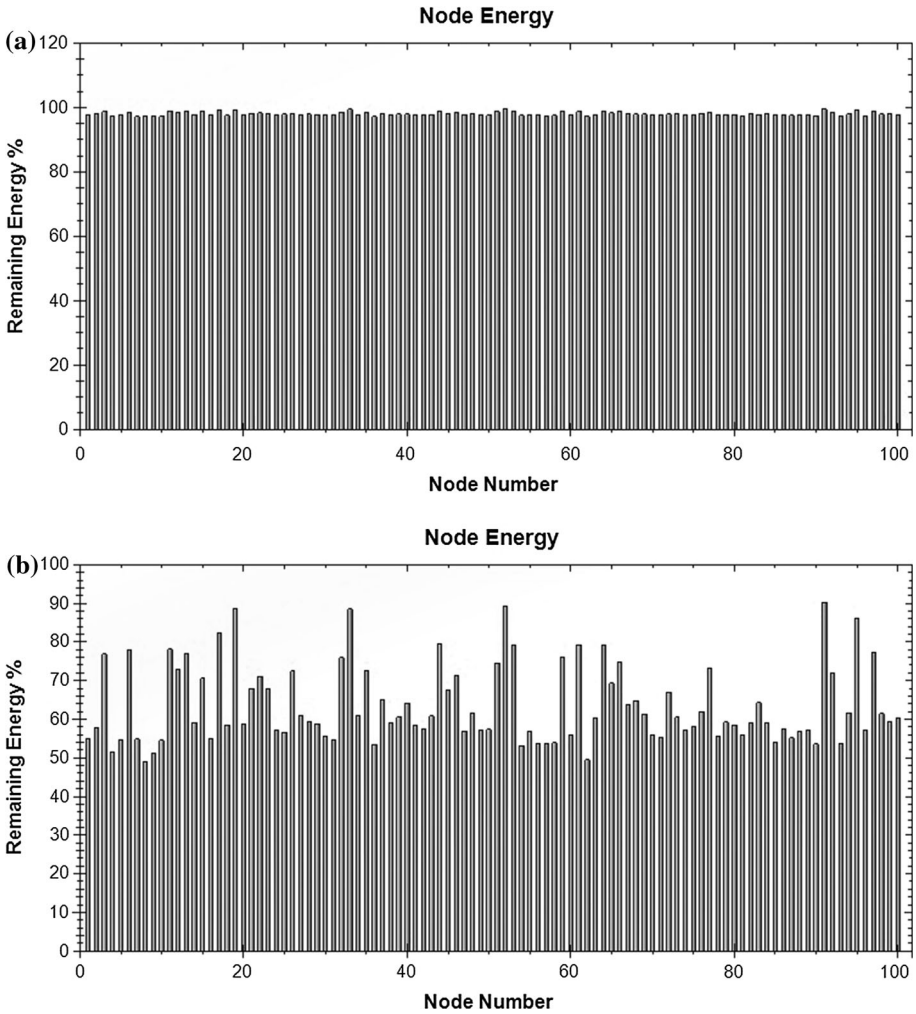**Node Energy**

(a)

Node Energy

(b)

Fig. 6 The remaining energy of all sensor nodes **a** at network transmission round 100 and **b** at network transmission round 1000

4 Experimental Results and Discussion

To evaluate DCH-GA, we repeated ten times the simulations and reported the average performance. Comparison studies were conducted with some other methods reported in the literature. Table 6 lists the network parameters used in our experiments and two scenarios are created to study the routine method performance in small and large networks.

Sensors are randomly placed within a 100 meters by 100 meters area and the base station is placed at a corner of the field, i.e., at (100, 100). Figure 5 illustrates an example of node distributions. Sensors are depicted with small rectangular icons and the base station is marked with a larger rectangular icon.

Figure 6 illustrates the percentage of the remaining energy of sensor nodes at transmission rounds of 100 and 1000. Nevertheless, it is clear that the remaining energy is fairly

Table 7 Network life time

Starting Energy (J/node)	Method	FND	LND
0.25	LEACH	394	665
	GA-WCA	651	971
	M-GEAR	490	893
	LA2D-GA	594	880
	TSEP	611	789
	DCH-GA	708	1415
0.50	LEACH	540	948
	GA-WCA	878	1415
	M-GEAR	700	1020
	LA2D-GA	914	1280
	TSEP	905	1180
	DCH-GA	1100	1702
1.0	LEACH	915	1950
	GA-WCA	1250	2800
	M-GEAR	1400	2770
	LA2D-GA	1746	3100
	TSEP	1814	2977
	DCH-GA	2260	3644

FND and LND represent the round in which the first and the last node die respectively

equal with some fluctuations. That is, as a consequence of DCH-GA, the variance among remaining energy is quite low, which implies that the sensor nodes shared the burden of relaying messages and, hence, elongated the overall network life. The simulation is repeated to check the network life time. The following GA parameters are used for simulation: population size $= 20$, generation size $= 120$, crossover rate $= 0.60$, and mutation rate $= 0.001$. The comparison is shown in Table 7. Figure 7 depicts the change of the percentage of live sensor nodes throughout the entire network life.

The forming of clusters was greatly influenced by the spatial location of sensor nodes. It is interesting to see that the low-initial-energy nodes that served as cluster heads are usually far away from the high-initial-energy ones, which justifies their role as CH.

Table 8 lists the average remaining energy of sensors and its standard deviation at various transmission rounds. Due to uneven distances to the CH of the member nodes, the energy expenditure for sensors varied, and is thus inevitable that STDs continued to increase. However, the small STDs values indicate a balanced energy consumption among sensors.

To evaluate the efficiency of DCH-GA, Table 9 lists the average time used to structure clusters in each transmission round. The time reported is before the first node became unavailable due to energy exhaustion. The number within parenthesis is the standard deviation. The first column represents the initial energy of the sensor nodes.

5 Conclusion

In a clustering WSN, each cluster usually consists of at least one surrogate node, often referred to as the CH. The CH may be dynamically chosen or preassigned by the network designer. Communication between the cluster and the base station is facilitated by this CH.
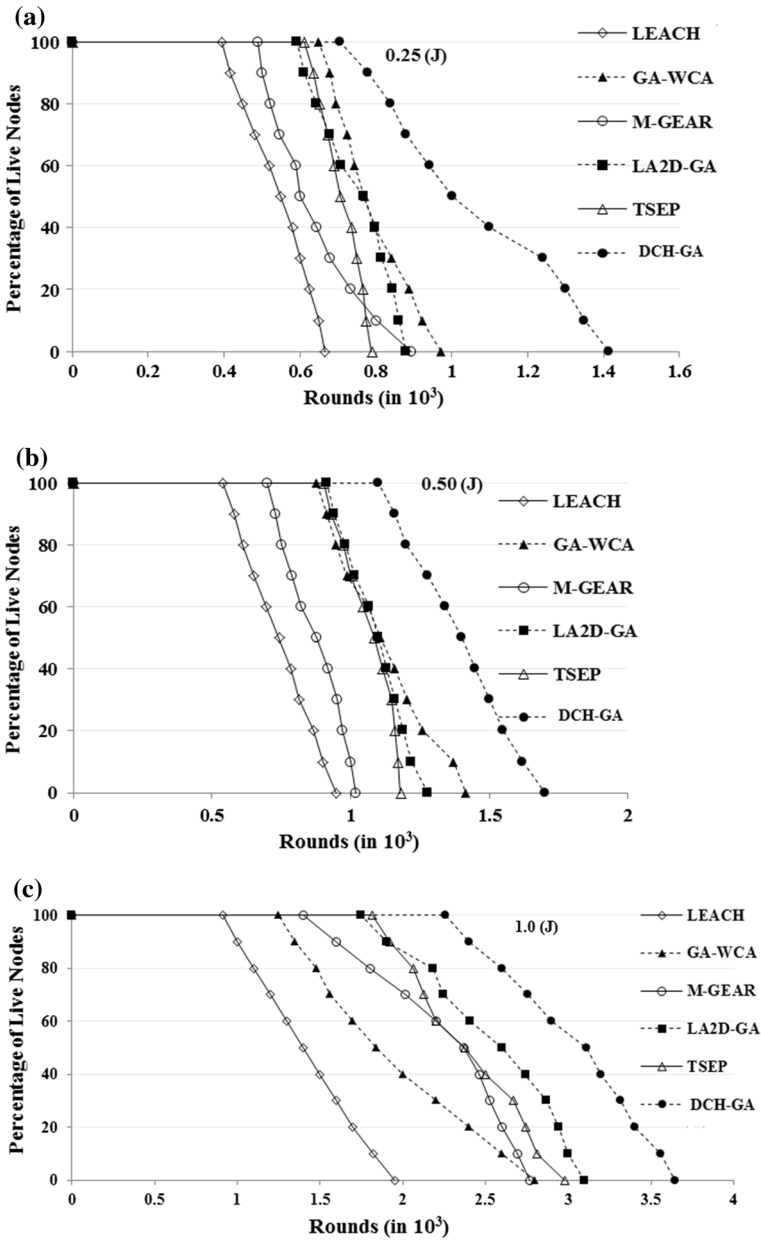
Fig. 7 Network lifetime with **a** initial energy 0.25 J/node, **b** initial energy 0.50 J/node, and c initial energy 1.0 J/node

The problem of choosing the CH in a multi-hop clustering model is more complex than in a single-hop model. The multi-hop clustering model is a special case of the clustering model in which a CH cannot transmit the data directly to the BS.

Table 8 Remaining energy (J) variance of sensor nodes

Rounds	500	1000	1500	2000	2500	3000
Mean	0.531	0.457	0.382	0.298	0.201	0.150
STD	0.031	0.042	0.033	0.037	0.048	0.056

Table 9 Average time (in seconds) for network structuring

Methods	LEACH	GA-WCA	M-GEAR	LA2D-GA	TSEP	DCH-GA
0.25	0.38	0.53	0.72	0.33	0.42	0.44
(J)	(0.02)	(0.15)	(0.07)	(0.06)	(0.08)	(0.02)
0.5	0.62	0.65	0.72	0.49	0.71	0.53
(J)	(0.12)	(0.27)	(0.12)	(0.19)	(0.28)	(0.11)
1.0	0.55	0.80	0.64	0.49	0.67	0.55
(J)	(0.14)	(0.44)	(0.21)	(0.23)	(0.19)	(0.20)

There are six main significant factors for selecting a CH node in a multi-hop cluster model in WSNs. These factors are: the distance from cluster center, the vulnerability index, the degree of mobility, the degree of mobility, the remaining battery power, the number of nearby neighbors, and the consumed energy. All these factors are related when choosing CH and ignoring one of them will affect the network lifetime. The degree of priority differs for each factor. In the present contribution, we have proposed a new CH selection method based on GA, called DCH-GA, for both single-hop and the multi-hop cluster models. The procedure introduced here is designed to match the requirements of the dynamic environments by electing the CH based on six main features.

During examination, we repeated simulations and reported the average performance. Comparison studies were conducted with some methods reported in the literature. In the future work, we shall focus on secure data transfer between each node and the CH node taking into consideration the dynamic environment inside the cluster.

References

1. Younis, M., & Akkaya, K. (2008). Strategies and techniques for node placement in wireless sensor networks: A survey. *Ad Hoc Networks*, *6*, 621–655.
2. Elhoseny, M., Yuan, X., Yu, Z., Mao, C., ElMinir, H. K., & Riad, A. M. (2014). Balancing energy consumption in heterogeneous wireless sensor networks using genetic algorithm. *IEEE Communications Letters*, *99*, 1–4.
3. Yuan, X., Elhoseny, M., ElMinir, H., & Riad, A. (2016). A genetic algorithm-based, dynamic clustering method towards improved wsn longevity. *Journal of Network and Systems Management*, *1–26*, 2016.
4. Rahman, A., Anwar, S., Pramanik, I., & Rahman, F. (2013). A survey on energy efficient routing techniques in wireless sensor network. In *International conference of Advanced communication Technology* (pp. 200–205).
5. Ali, J., Kumar, G., & Rai, M. (2013). Major energy efficient routing schemes in wireless sensor networks. *International Journal of Computers and Technology*, *4*(2), 261–266.
6. Elhoseny, M., Elminir, H., Riad, A., & Yuan, X. (2014). Recent advances of secure clustering protocols in wireless sensor networks. *International Journal of Computer Networks and Communications Security*, *2*(11), 400–413.

7. Pantazis, N., Nikolidakis, S., & Vergados, D. (2013). Energy-efficient routing protocols in wireless sensor networks: A survey. *Communications Surveys and Tutorials, 15*(2), 551–591.
8. Bhattacharjee, A., Bhallamudi, B., & Maqbool, Z. (2013). Energy-efficient hierarchical cluster based routing algorithm in WSN: A survey. *International Journal of Engineering Research and Technology, 2*(5), 302–311.
9. Du, T., Qu, S., Liu, F., & Wang, Q. (2015). An energy efficiency semi-static routing algorithm for WSNs based on HAC clustering method. *Information Fusion, 21*(1):18–29.
10. Tyagia, S., & Kumarb, N. (2013). A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks. *Journal of Network and Computer Applications, 36*(2), 623–645.
11. Attea, B. A., & Khalil, E. A. (2012). A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks. *Applied Soft Computing, 12*(7), 1950–1957.
12. Ruan, F., Yin, C., Chen, J., Wang, J., & Xue, S. (2013). A distance clustering routing algorithm considering energy for wireless sensor networks. *International Journal of Future Generation Communication and Networking, 6*(5), 73–80.
13. Iqbal, A., Akbar, M., Javaid, N., Bouk, S., Ilahi, M., & Khan, R. (2013). Advanced LEACH: A static clustering-based heterogeneous routing protocol for WSNs. *Journal of Basic and Applied Scientific Research, 3*(5), 864–872.
14. Elhoseny, M., Yuan, X., ElMinir, H., & Riad, A. (2014). Extending self-organizing network availability using genetic algorithm. In *International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE,* doi:10.1109/ICCCNT.2014.6963059.
15. Elhoseny, M., Elleithy, K., Elminir, H., Yuan, X., & Riad, A. (2015). Dynamic clustering of heterogeneous wireless sensor networks using a genetic algorithm, towards balancing energy exhaustion. *International Journal of Scientific and Engineering Research, 6*(8), 1243–1252.
16. Kang, S., & Nguyen, T. (2012). Distance based thresholds for cluster head selection in wireless sensor networks. *IEEE Communications Letters, 16*(9), 1396–1399.
17. Younis, O., & Fahmy, S. (2004). HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing, 3*(4), 366–379.
18. Guo, W., & Zhang, W. (2014). A survey on intelligent routing protocols in wireless sensor networks. *Journal of Network and Computer Applications, 38*, 185–201.
19. Ahmed, G., Khan, N., & Ramer, R. (2008). Cluster head selection using evolutionary computing in wireless sensor networks. In *Progress in electromagnetics research symposium* (pp. 883–886).
20. Bhaskar, N., Subhabrata, B., & Soumen, P. (2010). Genetic algorithm based optimization of clustering in ad-hoc networks. *International Journal of Computer Science and Information Security, 7*(1), 165–169.
21. Asim, M., & Mathur, V. (2013). Genetic algorithm based dynamic approach for routing protocols in mobile ad hoc networks. *Journal of Academia and Industrial Research, 2*(7), 437–441.
22. Karimi, A., Abedini, S., Zarafshan, F., & Al-Haddad, S. (2013). Cluster head selection using fuzzy logic and chaotic based genetic algorithm in wireless sensor network. *Journal of Basic and Applied Scientific Research, 3*(4), 694–703.
23. Rana, K., & Zaveri, M. (2013). Synthesized cluster head selection and routing for two tier wireless sensor network. *Journal of Computer Networks and Communications, 13*(3). doi:10.1155/2013/578241.
24. Heinzelman, W., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. In *The Hawaii International Conference on System Sciences, Maui, Hawaii.*
25. Nadeem, Q., Rasheed, M., Javaid1, N., Khan, Z., Maqsood, Y., & Din, A. (2013). M-GEAR gateway-based energy-aware multi-hop routing protocol for WSNs. In *Eighth international conference on broadband and wireless computing and communication and applications* (pp. 164–169).
26. Li, Q., & Qingxin, Z. (2006). Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks. *Computer Communications, 29*(12), 2230–2237.
27. Lindsey, S., & Raghavendra, C. (2002). Pegasis power-efficient gathering in sensor information systems. *IEEE Aerospace Conference Proceedings, 3*, 1125–1130.
28. Kashaf, A., Javaid, N., Khan, Z., & Khan, I. (2012). TSEP: Threshold-sensitive stable election protocol for WSNs. In *Conference on Frontiers of information technology* (pp. 164–168)
29. Elbhiri, B., Rachid, S., & Elfkihi, S. (2010). Developed distributed energy-effecient clustering (DDEEC) for heterogeneous wireless sensor. In *Communications and Mobile Network* (pp. 1–4). Rabat.
30. Qiang, Y., Pei, Bo., Wei, W., & Li, Y. (2015). An efficient cluster head selection approach for collaborative data processing in wireless sensor networks. *International Journal of Distributed Sensor Networks, 2015*. doi:10.1155/2015/794518.

31. Pala, V., Yogita, Y., Singh, G., & Yadav, P. (2015). Cluster head selection optimization based on genetic algorithm to prolong lifetime of wireless sensor networks. In *Third international conference on recent trends in computing* (pp. 1417–1423). Elsevier.
32. Batra, P., & Kant, K. (2016). Leach-mac a new cluster head selection algorithm for wireless sensor networks. *Wireless Networks, 22*(1), 49–60.
33. Diallo, C., Marot, M., & Becker, M. (2010). Single-node cluster reduction in WSN and energy-efficiency during cluster formation. In *The 9th IFIP annual mediterranean ad hoc networking conference, France*.
34. Chengfa, L., Mao, Y., Guihai, C., & Lie, W. (2005). An energy-efficient unequal clustering mechanism for wireless sensor networks. In *IEEE international conference on mobile adhoc and sensor systems, Washington, DC*.
35. Ahmed, G., Khan, N., & Khalid, Z. (2014). Cluster head selection using decision trees for wireless sensor networks. In *Sensor Networks and Information Processing* (pp. 173–178).
36 Tian, J., Gao, M., & Ge, G. (2016). Wireless sensor network node optimal coverage based on improved genetic algorithm and binary ant colony algorithm. EURASIP *Journal on Wireless Communications and Networking, 2016*, 104. doi:10.1186/s13638-016-0605-5.

Mohamed Elhoseny is a lecturer at the Faculty of Computers and Information at Mansoura University, Egypt. He got his PhD 2016 in Computer and Information Sciences after two years research scholar at Department of Computer Science and Engineering, University of North Texas, USA. His research interests include artificial wireless sensor network, data security, intelligence techniques, and big data.

Ahmed Farouk is a lecturer at Department of Information Technology, Alzahra Collage for Women, Oman. His research interest includes evolutionary computing, image processing, and quantum computing. He has several Publications in reputed and high impact journals like Nature Scientific Reports, QINP, and Security and Communications.

Nanrun Zhou is a Professor at Department of Electronic Information Engineering, Nanchang University, China. His research interest includes Wireless Communication, evolutionary computing, Information Security, and Quantum Computing.

Ming-Ming Wang was born in Shaanxi, China, on May 12, 1982. He received the Ph.D. degree from the School of Computer Science, Beijing University of Posts and Telecommunications in 2013. His current research interests include quantum cryptography, quantum information and quantum communication.
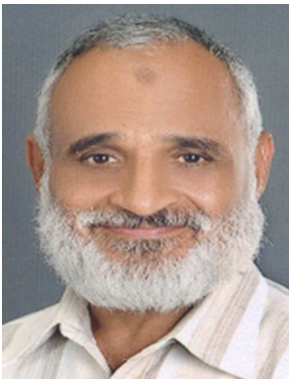
Prof. Dr. Soliman Abdalla works in the Department of Physics, at the King Abdulaziz University, in Jeddah, Saudi Arabia. He is interested in the physics of materials, as well as classical and quantum physics.

Josep Batle is a Lecturer at Department of Physics, Universitat de les Illes Balears, 07122 Palmade Mallorca, Balearic Islands, Spain. His research interest includes Quantum Information, Signal Processing, and Secure Data Communication.