

Dynamic Clustering of Heterogeneous Wireless Sensor Networks using a Genetic Algorithm, Towards Balancing Energy Exhaustion

Mohamed Elhoseny¹, Khaled Elleithy², Hamdi Elminir³, Xiaohui Yuan⁴, and Alaa Riad⁵

Abstract— Placing few heterogeneous nodes in Wireless Sensor network (WSN), such as nodes with more computing powers, is an effective way to increase network availability in terms of lifetime. Despite the success of various clustering strategies of heterogeneous WSN, the numerous possible sensor clusters make searching for an optimal network structure an open challenge. In this paper, we propose a heterogeneous sensor node clustering method using a Genetic Algorithm to optimize the energy exhaustion namely Dynamic Clustering of Heterogeneous WSNs using Genetic Algorithm 'DCHGA'. In DCHGA, the structure of the network is dynamically decided after each message transmission round. Compared with state-of-the-art methods, DCHGA greatly extended the network life and the average improvement with respect to the second best performance (using stable nodes) based on the first-node-die and the last-node-die were 33.8% and 13%, respectively. While in case of mobility heterogeneity of sensors, the improvement was between 12.6% and 9.8%. The balanced energy consumption greatly improved the network lifetime and allowed the sensor's energy to evenly deplete. The computational efficiency of DCHGA is comparable to the others and the overall average time across all experiments was 0.6 seconds with a standard deviation of 0.06.

Index Terms— Dynamic Clustering, Wireless Sensor Network, Heterogeneous Sensors

1 INTRODUCTION

The heterogeneous clustering model has been used in Wireless Sensor Networks (WSNs) to improve its performance in terms of network availability [1]. Although there are great works in the process of forming clusters, the dynamic nature of WSN and numerous possible cluster configurations make searching for an optimal network structure a complicated defy [2]. The heterogeneous model is an adapted model of homogeneous clustering model, i.e., LEACH [3]. This modification can be achieved by placing few heterogeneous nodes in network [4]–[6], such as nodes with more computing power. In a heterogeneous WSN, in addition to the network structuring factors, e.g., distance to the base-station, and distance among nodes, factors such as initial energy, data processing capability, ability to serve as a cluster head, and node mobility greatly influence the network lifespan [7]–[9]. Moreover, the lifetime of the network is maximized when the remaining energy of nodes in the network remains the same during the network lifetime. This is, however, difficult to achieve in a real-world WSN due to different roles of sensor nodes and various signal transmission distance. The nodes serving as cluster head consume additional energy to fulfill tasks such as

receiving messages from member nodes and relaying the aggregated messages to the base station. Balancing node energy consumption and extending the overall network lifespan are non-trivial given many factors that could affect the energy expenditure of each node [10], [11], [32].

To extend the network lifetime in a heterogeneous network, several methods have been proposed that account for one or more of these factors. Stable Election Protocol (SEP) [4] used weighted probabilities to elect cluster heads depending on the remaining energy in the sensor nodes. In addition, Developed Distributed Energy-Efficient Clustering (DDEEC) [5] method improved upon SEP by categorizing sensor nodes based on their energy level. The nodes with higher energy were the "advanced nodes" and the cluster head were selected from these group of nodes. Threshold Sensitive Stable Election Protocol (TSEP) [6] extended SEP method by grouping sensor nodes into three energy levels, and the cluster heads were selected based on thresholds. Similarly, Energy Efficient Heterogeneous Clustered scheme (EEHC) [12] and Efficient Three Level Energy algorithm (ETLE) [13] selected cluster heads based on the probability proportional to the residual energy. In Hybrid Energy Efficient Reactive protocol (HEER) [14], the cluster head selection is based on the ratio of residual energy of nodes and the average energy of the network. Both of Energy efficient heterogeneous clustered scheme (EEHC) [12] and Efficient Three Level Energy algorithm (ETLE) [13] assume three levels of heterogeneity and nodes are randomly distributed and are stationary. In EEHC, the cluster heads were selected based on weighted election probabilities of each node according to the residual energy. While in ETLE, each node

- ^{1,4} Department of Computer Science and Engineering, University of North Texas, Denton, TX, U.S.A.; E-Mails: Mohamed.elhoseny@unt.edu, xiaohui.yuan@unt.edu
- ² Department of Computer Science and Engineering, University of Bridgeport, CT, U.S.A.; E-Mail: elleithy@bridgeport.edu
- ⁴ Department of Electrical Engineering, Kafr El-Sheikh University, Kafr El-Sheikh, Egypt; E-mail: hamdy elminir@eng.kfs.edu.eg
- ⁵ Department of Information Systems, Mansoura University, Mansoura, Egypt; E-mail: amriad_2000@mans.edu.eg

chose a random number between 0 and 1. If the value of this random number was less than a threshold value, i.e., T , the node will be selected to serve as a cluster head. In Hybrid Energy Efficient Reactive protocol (HEER) [14], the CH selection is based on the ratio of residual energy of nodes and average energy of the network. All of these methods were proposed for WSN with initial energy as the only heterogeneity factor.

Although most of current research concentrated on energy as the only heterogeneity factor, many types of heterogeneous resources, e.g., communication capability, data processing power, and efficiency, were introduced to WSN for improved performance. Providing sensor node with more processing capabilities aims to prevent it from exhausting its energy quickly in case of acting as a cluster head. Allowing mobile sensors in heterogeneous model increases the number of WSN applications compared with stationary sensors, i.e., tracking animal movements applications [15]. On the other hand, denying some nodes to serve as a cluster head, e.g., nodes with low energy, increases its chance to stay alive. Searching for a balance among many factors is an involved and complex process. Heuristic optimization methods, such as Genetic Algorithms (GA), have been employed in the routing protocols of WSN [16]–[19]. When GA is used, a key objective is to define an appropriate fitness function that encodes the network structure. However most of GA-based work reported in literature was developed for homogeneous model, i.e., HCR [16], [20], while the remaining was concerned with heterogeneous WSN in which the difference between sensors in the initial energy is the dominant factor of heterogeneity. The Evolutionary Based Clustered Routing Protocol (ERP) [18] overcomes the limitations of clustering-algorithm-based GAs by uniting the clustering aspects of cohesion and separation error, and proposed a new fitness function based on these two aspects.

In this article, we propose a sensor clustering method for dynamically organizing heterogeneous WSN using GA called 'DCHGA'. DCHGA provides a framework to integrate multiple heterogeneity and clustering factors, which employs remaining energy, expected energy expenditure, network locality, and distance to the base station in search for an optimal, dynamic network structure for heterogeneous WSN. Heterogeneity factors are integrated as constraints to chromosomes and validation is performed to ensure network integrity. To avoid high energy consumption of sensor nodes, the base station will run the GA after each round to dynamically forming the structure of the network based on the new characteristics of the sensors, i.e., remaining energy. Genetic algorithm uses random search to suggest the best appropriate design. We use this algorithm in order to obtain the most efficient clustering structure. The reason for choosing GA is its convergence and its flexibility in solving multi-objective optimization problems like dynamic clustering of WSN [16]

The contribution of this work includes: First, a GA-based method is proposed to provide a dynamic clustering method for WSN. In addition, it provides flexibility of optimizing multiple factors concurrently. The GA chromosomes encode the selection of cluster heads and the dynamic clusters are formed accordingly. The separation of cluster head selection and network structure makes this method versatile for integrating

diverse factors. Second, the expected energy expenditure is derived, together with other energy and spatial metrics, to achieve balanced energy consumption across all nodes and improve the network's longevity.

In the reminder of this article, section 2 presents the related work of constructing heterogeneous WSN to extend its lifetime. Then, section 3 describes our proposed method for heterogeneous WSNs construction. Section 4 discusses our experimental results including a comparison study with state-of-the-art methods and analysis of energy consumption. Section 5 provides conclusions of this paper.

2 RELATED WORK

To extend the network lifetime in a heterogeneous network, Stable Election Protocol (SEP) [4] used weighted probabilities to elect cluster heads depending on the remaining energy in sensor nodes. In addition, Developed Distributed Energy-Efficient Clustering (DDEEC) [5] method improved upon SEP by categorizing sensor nodes based on their energy level. The nodes with higher energy were the "advanced nodes" and the cluster head were selected from these group of nodes. Threshold Sensitive Stable Election Protocol (TSEP) [6] extended SEP method by grouping sensor nodes into three energy levels, and the cluster heads were selected based on thresholds.

Similarly, Energy Efficient Heterogeneous Clustered scheme (EEHC) [12] and Efficient Three Level Energy algorithm (ETLE) [13] select cluster heads based on the probability proportional to the residual energy. In Hybrid Energy Efficient Reactive protocol (HEER) [14], the cluster head selection is based on the ratio of residual energy of nodes and the average energy of the network. Both of Energy efficient heterogeneous clustered scheme (EEHC) [12] and Efficient Three Level Energy algorithm (ETLE) [13] assume three levels of heterogeneity and nodes are randomly distributed and are stationary. In EEHC, the cluster heads are selected based on weighted election probabilities of each node according to the residual energy. While in ETLE, each node chooses a random number between 0 and 1. If the value of this random number was less than a threshold value, i.e., T , the node will be selected to serve as a cluster head. In Hybrid Energy Efficient Reactive protocol (HEER) [14], the CH selection is based on the ratio of residual energy of nodes and average energy of the network. All of these methods are proposed for WSN with initial energy as the heterogeneity factor.

In [21], the Degree of connectivity is the main factor of selecting a CH. The degree of connectivity of a node, i.e. the number of its neighbors, is also a criterion that seems interesting to study. Intuitively, the more neighbors a sensor has, the more it seems to be an appropriate candidate as cluster head, since a sensor with a low degree of connectivity might have little information, from its neighborhood, to aggregate and to forward to the BS. In the initial phase, each sensor is involved in the neighborhood information exchanges (hello protocol), which allows it to determine its degree of connectivity and the location of BS. In EEUC [22], the distance between the node and the BS is the main parameter for selecting the CH. The EEUC

resulted in a network that is partitioned into clusters of unequal size, and that the clusters closer to BS have smaller sizes than those farther from the BS.

Many intelligent algorithms which provide adaptive methods that present intelligent behavior in complex and dynamic environments like WSNs are exist [23]. Various works reported in literature [23]–[27] debated the routing protocols in Cluster-based WSN based on intelligent algorithms as reinforcement learning, ant colony optimization, fuzzy logic, genetic algorithm, and neural networks. Furthermore, a lot of clustering mechanisms have been proposed. For example, Local Negotiated Clustering Algorithm presents a novel clustering method, which uses the similarity of nodes readings as an important feature during the process of creating cluster. ACE constructs the WSN clusters in a fixed number of iterations using the node degree as the main factor. In GA-WCA, load balanced factor with a sum of distance from all neighbor nodes to CHs represents the main factor in network construction. on the other hand, LA2D-GA depends only on the distance as the main parameter to calculate fitness function that is used to evaluate the chance of the node to be a CH. LA2D-GA represent the chromosome in a two-dimensional grid which elucidates valid statistics of a WSN [28].

In [29], a two level fuzzy logic approach is used to Cluster Head (CH) election based on four parameters namely- number of neighbor nodes, remaining energy, energy dispersion and distance from the base station. The authors supported their idea for number of neighbors by stating that The number of neighbor nodes has been considered to be one determining parameter because CH must be chosen from an area where sufficient neighbor nodes are available LELE [30] protocol selects CH on basis of remaining energy and the distance between a node and its neighbors, and the node with maximum energy and suitable position is chosen as the CH. LELE is proposed to improve load balancing in LEACH protocol Leader Election with Load balancing Energy . So when the network is operating, the probability of the nodes becoming leader decreases or increases depending on the difference of the energy level of one node and neighbors, the distance of the node from neighbors, as well as the number of neighbors, and the probability of the nodes' to become a leader.

In [21], the Degree of connectivity is the main factor of selecting a CH. The degree of connectivity of a node, i.e. the number of its neighbors, is also a criterion that seems interesting to study. Intuitively, the more neighbors a sensor has, the more it seems to be an appropriate candidate as a cluster head, since a sensor with a low degree of connectivity might have little information, from its neighborhood, to aggregate and to forward to the BS. In the initial phase, each sensor is involved in the neighborhood information exchanges (hello protocol), which allows it to determine its degree of connectivity and the location of BS. In EEUC [22], the distance between the node and the BS is the main parameter for selecting the CH. The EEUC resulted in a network that is partitioned into clusters of unequal size, and that the clusters closer to BS have smaller sizes than those farther from the BS.

Searching for an optimal balance among many factors is non-trivial. Genetic Algorithm (GA) has been applied in the routing

protocol of WSN [17]–[19]. When GA is used, a key objective is to define an appropriate fitness function that encodes the network structure and its goodness. However, most of GA-based work was developed for homogeneous model, i.e., HCR [20], while the remaining was concerned with static heterogeneous WSN. There are no additional efforts reported for the mobile heterogeneous model. The Evolutionary Based Clustered Routing Protocol (ERP) [18] overcomes the limitations of clustering-algorithm-based genetic algorithms by uniting the clustering aspects of cohesion and separation error, and proposed a new fitness function based on these two aspects.

3 HETEROGENEOUS WSN CLUSTERING USING GENEIC ALGORITHM

3.1 Energy Model and Clustering Factors

As we deal with two levels of heterogeneity, our model has two types of sensors: normal and advanced sensor nodes. The advanced sensor has additional initial energy and lower energy consumption for data processing, i.e., receiving and transmitting messages. Based on that, we adopt the first order radio model to describe the sensor's energy [31] as shown in Figure 1. The consumed energy E of a normal sensor node s is the summation of energy used to:

- 1) Acquire l bits of data ($E_s^A(l)$)
- 2) Receive l' bits of data ($E_s^R(l')$)
- 3) Process l'' bits of data ($E_s^P(l'')$)
- 4) Transmit l'' bits of data over distance d ($E_s^T(l'', d)$),
and
- 5) Move from location x to location y .

$$E_s = E_s^A(l) + E_s^R(l') + E_s^P(l'') + E_s^T(l'', d) + E_s^M(x, y) \quad (1)$$

where $E_s^R = E_i + l' E^*$ and E_i is the idle energy expenditure. $E_s^T = E_i + l'' d^n$, and $n = 4$ for long distance transmission and $n = 2$ for short distance transmission, and E^* represents the cost of beam forming approach for energy reduction. The long and short transmission distance is determined by the distance threshold as we will explain later.

To compute the expected consumed energy \hat{E} of a non-CH sensor node s' and a CH sensor node s , assume l bits of data are collected by each sensor node in a round. Given N_s sensors in a cluster, the expected consumed energy \hat{E} are computed as follows:

$$\hat{E}_s = E + ID^2(s', s), \quad (2)$$

$$\hat{E}_s = E + N_s I E^* + (N_s + 1) I D^4(s', B), \quad (3)$$

where E is the constant energy consumption including the energy of data acquisition, processing, idle and moving. Functions $D(s, B)$ and $D(s, s')$ use Euclidean distance to give the distance between sensor nodes inside the cluster and from the cluster head to the base station, respectively. In addition, the local sensor density is proportional to the number of sensors within the δ -vicinity as follows:

$$G_s(\delta) \propto ||S_s||, \text{ and } S_s = \{s_i; D(s, s_i) \leq \delta\} \quad (4)$$

where S_s is the set of sensor nodes in the δ -vicinity of s and function $|| \cdot ||$ gives the set size.

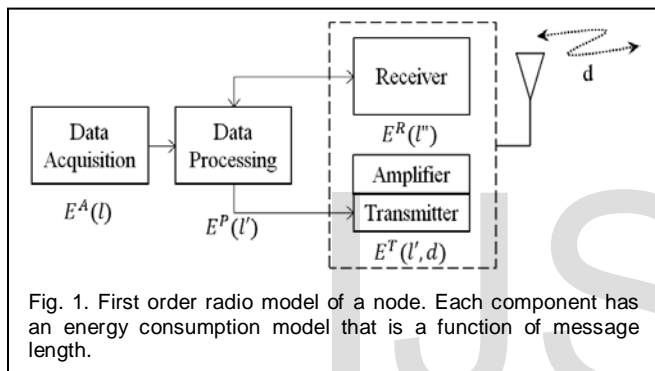


Fig. 1. First order radio model of a node. Each component has an energy consumption model that is a function of message length.

3.2 Network Structure Building using Genetic Algorithm

In our proposed framework, a binary chromosome is used to specify the CHs in the network, in which a one represents a CH and a zero represents a member node to a cluster as shown at Figure 2. When a sensor becomes inactive, i.e., out of power, its corresponding gene value is set to -1, which exempts this sensor from further GA operations.

The mapping to sensor clusters from a chromosome is to minimize the network communication distance D as follows:

$$\hat{D} = \sum_{i=1}^C \sum_{j=1}^{N_{s_i}} D(s_i, s_j) \quad (5)$$

where C is the number of clusters in a network and N_{s_i} is the number of member nodes in a cluster headed by node s_i . In practice, minimizing D is equivalent to assigning sensor nodes to clusters following the nearest neighbor rule.

The fitness function integrates energy factors, spatial distances, and the local sensor density:

$$f = \sum_s \frac{E_s(t)}{E_s(0)} + \frac{\tilde{E}}{\hat{E}} + \frac{1}{\hat{D}} + \frac{1}{N} \sum_{s'} G_{s'}(\delta), \quad (6)$$

where $E_s(t)$ is the remaining energy of sensor node s at round t and $E_s(0)$ is the initial energy of sensor node s. \tilde{E} is the total energy cost if the messages are transmitted directly from all sensor nodes to the BS. \hat{D} is the total distance between the CHs and the BS:

$$\hat{D} = \sum_{i=1}^C D(s_i, BS) \quad (7)$$

where each s_i is a sensor node that serves as a CH. Including sensor density favors the choice of CHs with more close neighbors. In cases where it is clear one or more factors play more vital role, uneven weights can be employed in the fitness function.

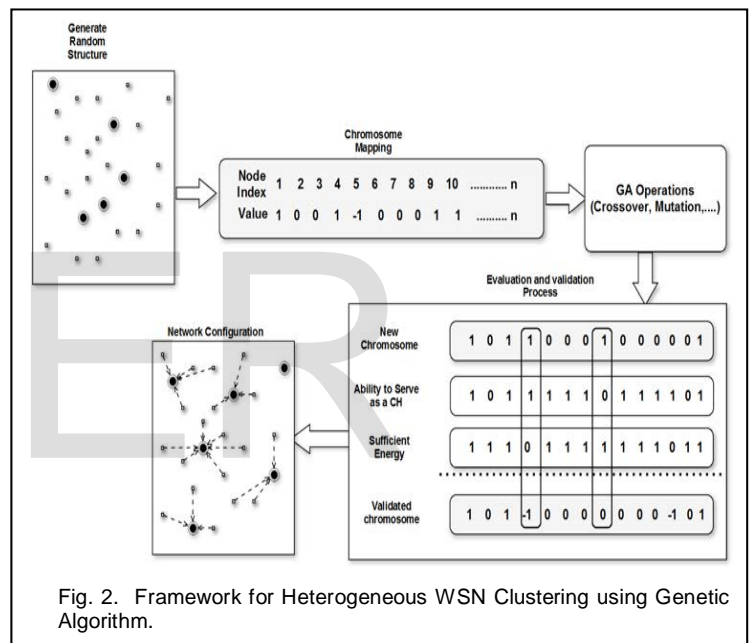


Fig. 2. Framework for Heterogeneous WSN Clustering using Genetic Algorithm.

3.3 Network Structure Validation and Evaluation

In a heterogeneous WSN, functions and capabilities of sensors significantly vary. Some sensors are unable to serve as a cluster head and some are preferred to take the role due to their superior processing power and available energy. However, classical optimization method such as GA provides no integrated mechanism for ensuring alignment of different roles of the sensors. In addition, the random initialization and GA operations could introduce chromosomes that completely violate the current sensor properties. In DCHGA, heterogeneity is presented as constraints and hence a validation process is needed before evaluating chromosomes' fitness to ensure network integrity.

Figure 2 illustrates also the validation process that leverages static and dynamic sensor properties. In the process of GA optimization, a new chromosome represents the proposed structure for the WSN. Each gene in the chromosome defines

the expected role of the corresponding sensor node, i.e., whether it serves as a cluster head or a member node. The process consults 'the ability to serve as a CH', and 'the Sufficient Energy' tables. The role of 'The ability to serve as a CH' table is to determine whether the node can serve as a cluster head (one represents serving as cluster head; otherwise, member node). While, the 'Sufficient Energy' table is used to show the current energy status of the node, i.e., zero for disabled node and one for available node. The validation process determines if a chromosome complies with the constraints and hence retained in the offspring pool; otherwise, the chromosome is abandoned.

GA generates new chromosomes through crossover and mutation operations and evaluates their fitness. The crossover operation is performed with two randomly selected chromosomes determined by a crossover probability to regulate the operation. When crossover is excluded, the parent chromosomes are duplicated to the offspring without change. Varying the crossover probability alters the evolution speed of the search process. In practice, the value of crossover is close to 1. The mutation operation involves altering the value at a randomly selected gene within the chromosome. Similarly, a mutation probability is used to regulate the performance of mutation. Different from the crossover probability, the mutation probability is usually fairly small. Essentially mutation operation could create completely new species, i.e., an arbitrary locus in the fitness landscape. Hence, it is a means to get out of a local optimum. Recall that when a sensor node becomes inactive, its corresponding gene is set to -1 to exempt it from mutation operations.

After the validation process is executed, Eq. (6) is used to evaluate the fitness of chromosomes. An intermediate pool of chromosomes is created to hold the individuals created in a generation, and depending on the needs, the user can specify any intermediate population size that is greater than the initial population size.

The evolution terminates when one of the following criteria is satisfied: 1) the maximum number of generations is reached; or 2) the fitness function converges. Upon completion of the GA evolution, the chromosome that gives the best fitness value is used to restructuring the nodes.

3.4 Heterogeneous WSN clustering Algorithm

Algorithm 1 presents DCHGA method. In this algorithm, $q \in [1, Q]$ denotes the number of generations, and the population size is P . The pool of chromosome, denoted by U , is initialized with randomly generated individuals.

In crossover operation, two chromosomes are randomly selected from U and, according to the crossover probability α ,

two new chromosomes are created by switching consecutive genes. In mutation operations, the value of a randomly picked gene is altered between 0 and 1 according to the mutation

Algorithm 1 Dynamic clustering of Heterogeneous WSN using Genetic Algorithm

- 1: Randomly generate a pool of S Network structures $X = \{x_1, x_2, \dots, x_S\}$.
 - 2: Represent each structure in S by one chromosome to get a pool P of chromosomes $U = \{u_1, u_2, \dots, u_P\}$.
 - 3: $\forall u_i \in U$, Forming the WSN clusters by minimizing Eq. (5)
 - 4: Evaluate the fitness of each $u_i \in U$ following Eqs. (6).
 - 5: for $q = 1, 2, \dots, Q$ do
 - 6: $\tilde{U} \leftarrow \emptyset$
 - 7: for $p = 1, 2, \dots, P$ do
 - 8: Randomly select $u_a, u_b \in U$ ($a \neq b$) based on the normalized fitness $\tilde{f}(u)$:

$$\tilde{f}(u) = \frac{f(u)}{\sum_c f(u)}$$
 - 9: Cross over u_a and u_b according to α

$$\mathcal{C}(u_a, u_b | \alpha) \Rightarrow u'_a, u'_b$$
 - 10: Perform mutation on u'_a and u'_b according to β

$$\mathcal{M}(u'_a | \beta) \Rightarrow \tilde{u}_a, \quad \mathcal{M}(u'_b | \beta) \Rightarrow \tilde{u}_b$$
 - 11: Evaluate $f(\tilde{u}_a)$ and $f(\tilde{u}_b)$.
 - 12: $\tilde{U} \leftarrow \tilde{U} \cup \{\tilde{u}_a, \tilde{u}_b\}$
 - 13: end for
 - 14: $U \leftarrow \{u_i; u_i \in \tilde{U} \text{ and } f(u_i)\}$
 - 15: end for
 - 16: Find the chromosome u^* that satisfies

$$u^* = \arg \max_u f(u), u \in U$$
 - 17: Return the network structure x^* by mapping u^* back
-

probability β .

4 RESULTS AND DISCUSSION

In our evaluation, we assumed that each sensor node can directly reach the base station if it is provided with sufficient energy. The simulated sensor network was in an area of 100 meters by 100 meters (m) with 50 sensors randomly placed in the field and the data packet size was 400 bits. The network parameters used in our experiments are listed in Table I. The heterogeneity includes different initial power, data processing efficiency, capability of serving as cluster head, and node mobility. For the sensors with greater data processing efficiency, the energy used is 50% of that used by a regular sensor. 10% of sensor nodes possessed greater initial energy and data processing efficiency and 10% of sensor nodes were unable to serve as cluster heads. The heterogeneous sensors were chosen randomly in each experiment. Regarding GA running parameters, we used the population size of 30 for 30 generations. The crossover probability and mutation probability are 0.8 and 0.006, respectively. As we mentioned before, the mutation process should be close to 1 while the mutation should be close to zero. The neighborhood distance δ was 20 meters (m) throughout our experiments.

To evaluate DCHGA in different environments, we created two cases with low and high sensors density, i.e., 50 node and 100 node, respectively. For each case, two scenarios of heterogeneity are designed: 1) sensors may differ in their initial energy, and 2) sensors may differ in initial energy, data pro-

cessing capability, and the ability to serve as a cluster head. Comparison studies were conducted with five state-of-the-art methods including HEER [14], TSEP [6], DDEEC [5], ETLE [13], and ERP [18].

heterogeneity was in the ranges of 20.8% to 38.4%. This means the more heterogeneity capabilities assigned to the advanced nodes, the more network lifetime. It is clear that our proposed heterogeneity factors greatly extended the network lifespan.

TABLE I. NETWORK PARAMETERS.

Parameters	Values	
Field Area	100m × 100m	
Base-station location	Center of the field	
Energy of regular sensor	0.5J	
Energy of advanced sensor	1.0J	
Idle state energy	50nJ/bit	
Data aggregation energy	5nJ/bit	
Amplification energy (cluster head to base-station)	$d \geq d_0$ $d < d_0$	10pJ/bit/m^2 0.0013pJ/bit/m^2
Amplification energy (sensor to cluster head)	$d \geq d_1$ $d < d_1$	$E_{ts}/10 = E_{s1}$ $E_{mp}/10 = E_{mp1}$
Packet size	400 bits	
Percentage of advance sensors	0.1	

Scenario 1: This scenario aims to evaluate the impact of heterogeneity in terms of initial energy of the sensor nodes. Table VI compares network life of DCHGA with five state-of-the-art methods, which include HEER [14], TSEP [6], DDEEC [5], ETLE [13], and ERP [18]. The average number of rounds when the first node died (FND) and last node died (LND) are reported; and 10 experiments were conducted for the analysis. DCHGA exhibited the longest average network life. The average improvement with respect to the second best performance based on FND and LND are 33.8% and 13%, respectively. Fig. 3 depicts the number of live nodes throughout the network life, which presents a progressive view. The dash line with solid dot shows the results of DCHGA. The balanced energy consumption greatly improved the network life and allowed the sensor energy to deplete evenly. This means the stability [18] of DCHGA is the best one compared with the five other methods. It is clear that DCHGA greatly extended the network life.

Fig. 3 depicts the change of the percentage of live sensor nodes throughout the entire network life. It is evident that the improvement of DCHGA method is significant.

Scenario 2: The purpose of this scenario is to evaluate the impact of heterogeneity in terms of initial energy, processing capability, and the ability of the sensor to act as a cluster head. Table III presents the percentage of live sensor nodes throughout the life span of the WSN using the two cases for each scenario. The number of round is the average of 10 experiments with random sensor node placement. Compared with initial energy heterogeneity scenario, using the mentioned three factors of heterogeneity together yielded the largest number of rounds when the first sensor node dies. Depending on the sensor density, the improvement of using these three factors of

TABLE II. NETWORK LIFE SPAN WITH DIFFERENT METHODS IN THE TWO PROPOSED CASES. FND: ROUND AT WHICH FIRST NODE DIE. LND: ROUND AT WHICH LAST NODE DIE.

Methods	100 Sensor Nodes		50 Sensor Nodes	
	FND	LND	FND	LND
ETLE	2040	8200	1514	6904
ERP	2756	10370	2010	9200
HEER	2340	7400	1789	6150
DDEEC	1879	10000	1100	8900
TSEP	2613	8000	1986	7640
DCHGA	3510	12250	2690	10400
Improvement(DCHGA)	27.3 %	18.1%	33.8%	13%

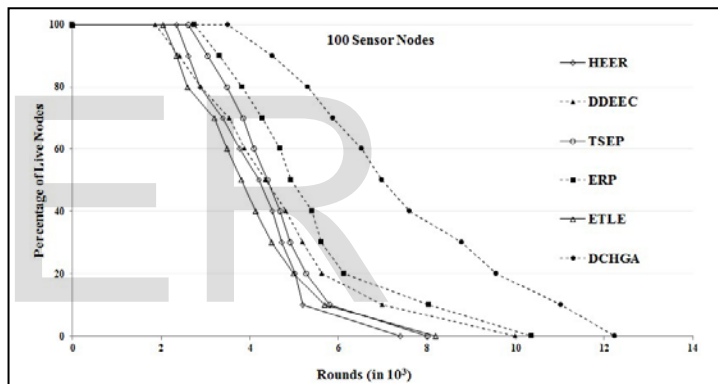


Fig. 3 (a). Network lifetime for high density field with 100 sensors

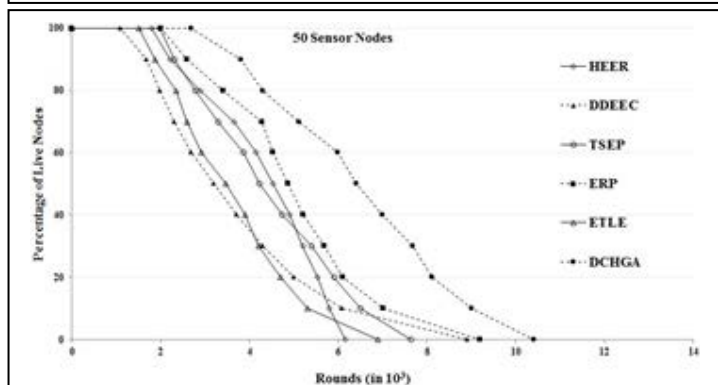


Fig. 3 (b). Network lifetime for low density field with 50 sensors

TABLE III. NETWORK LIFE SPAN WITH DIFFERENT HETEROGENEITY FACTORS IN THE TWO PROPOSED CASES. S-CH: SOME NODES CAN NOT SERVE AS A CLUSTERS HEAD.

Heterogeneity Factors	100 Senso Nodes		50 Senso Nodes	
	FND	LND	FND	LND
Energy	3510	12250	2690	10400
Energy, Processing, and S-CH	4860	19045	3250	14680

Figure 4 depicts the change of the percentage of live sensor nodes throughout the entire network life. It is evident that the improvement of using different heterogeneity factors is significant. Our experiments also showed that the average number of clusters before the first node die was 7% and 8% in high and low density fields respectively.

Figure 5 illustrates an example of the remaining energy at various transmission rounds of all sensors. At round 0, i.e., the initialization, 5 nodes (highlighted with green bars) were fueled with greater energy at 1J. The red bars mark sensors unable to serve as cluster head. As transmission continued, the remaining energy of sensors gradually reduced mostly evenly.

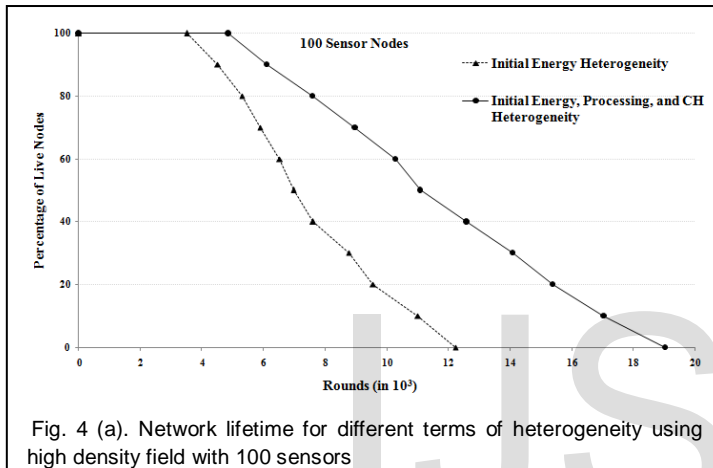


Fig. 4 (a). Network lifetime for different terms of heterogeneity using high density field with 100 sensors

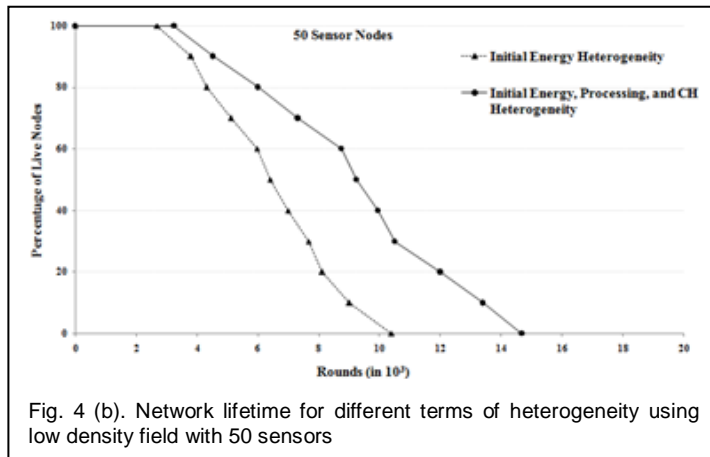


Fig. 4 (b). Network lifetime for different terms of heterogeneity using low density field with 50 sensors

Table IV lists the average remaining energy of the low-initial-energy sensors and its standard deviation at various transmission rounds. Due to unequal distances to the cluster head of the member nodes, energy expenditure for sensors varied, and it is inevitable that STDs continued to increase. However, the small STDs indicate balanced energy consumption among sensors.

Figure 6 illustrates the spatial and frequency view of sensor nodes serving as cluster heads throughout the life of the network. The size of sphere is proportional to the number of

times a sensor served as cluster head. It is clear that the ones with higher initial energy served as cluster head most times. The placement of higher energy sensors is randomized, which is unfortunately uneven in the field. Despite that the high-initial-energy sensors dominated the choice of cluster head, their spatial disadvantage, i.e., closely located with each other, made some low-initial-energy sensors to act as cluster head to serve nearby sensors. The average number of clusters in all rounds of our 10 experiments is 6, among which 97% of times high-initial-energy nodes served as cluster head. 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 head are usually far away from the high-initial-energy ones, which justifies their role as cluster head.

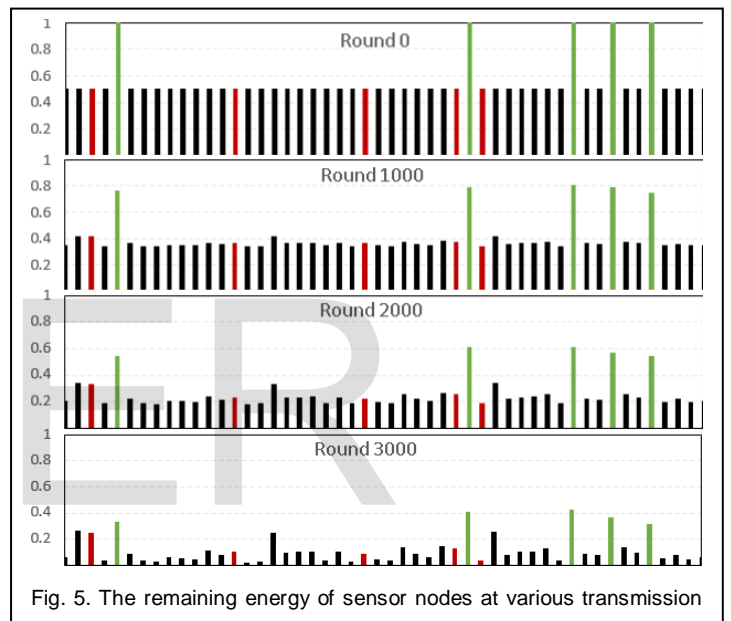


Fig. 5. The remaining energy of sensor nodes at various transmission

TABLE IV. REMAINING ENERGY (J) VARIANCE OF SENSOR NODES.

Rounds	500	1000	1500	2000	2500	3000
Mean	0.431	0.363	0.295	0.226	0.158	0.090
STD	0.010	0.020	0.031	0.042	0.052	0.062

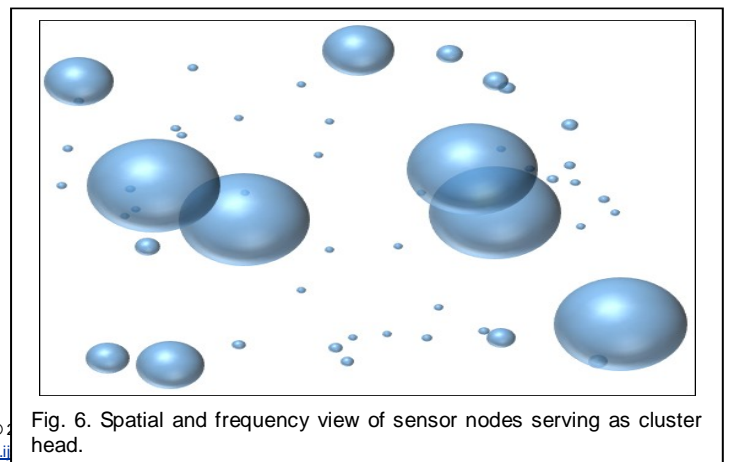


Fig. 6. Spatial and frequency view of sensor nodes serving as cluster head.

Efficiency is an important factor in real-world applications. Our experiments were conducted in a computer with Intel core i5 2.6GHz CPU, 4GB memory, and Windows 7 operating system. The algorithms were implemented with C# programming language. Table VII 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. In addition to 50 sensors in the field, we also experimented with 100 randomly placed sensors with the other parameters remain the same. The average time used by GAHN was comparable to the other methods.

TABLE V. AVERAGE TIME (IN SECONDS) FOR NETWORK STRUCTURING.

Methods	ETLE	ERP	HEER	DDEEC	TSEP	DCHGA
50	0.42	0.60	0.43	0.39	0.45	0.54
sensor	(0.03)	(0.12)	(0.02)	(0.04)	(0.06)	(0.06)
100	0.53	0.71	0.51	0.55	0.61	0.63
sensor	(0.10)	(0.34)	(0.21)	(0.11)	(0.17)	(0.27)

To evaluate the proposed method in terms of mobility heterogeneity, we will use M-DCHGA to indicate to our method with node mobility. Table VI presents the percentage of live sensor nodes throughout the life span of the WSN using different cases, i.e, static and mobile sensors. The number of round is the average of 10 experiments with random sensor node placement. Depending on the sensor density, the improvement was in the ranges of 27.3% to 33.44% using static sensors. While in case of sensors mobility, it was between 12.6% and 9.8%. This means the stability [18] of DCHGA is the best one compared with the five other methods. It is clear that DCHGA method greatly extended the network life.

Figure 7 depicts the change of the percentage of live sensor nodes throughout the entire network life. The experiments showed that the average number of clusters before the first node die was 8%, and 6% for high density scenario in case of static and mobile sensors respectively. While it was 9%, and 8% for low density scenario in case of static and mobile sensors respectively.

MOBILE NODES HETEROGENEITY. FND: ROUND AT WHICH FIRST NODE DIE. LND: ROUND AT WHICH LAST NODE DIE.

Methods	100 Sensor Nodes		50 Sensor Nodes	
	FND	LND	FND	LND
DCHGA	3510	12250	2690	10400
M-DCHGA	3105	10700	2208	9565
Improvement(DCHGA)	27.3 %	18.1%	33.8%	13%
Improvement(M-DCHGA)	12.6 %	3.2%	9.8%	3.9%

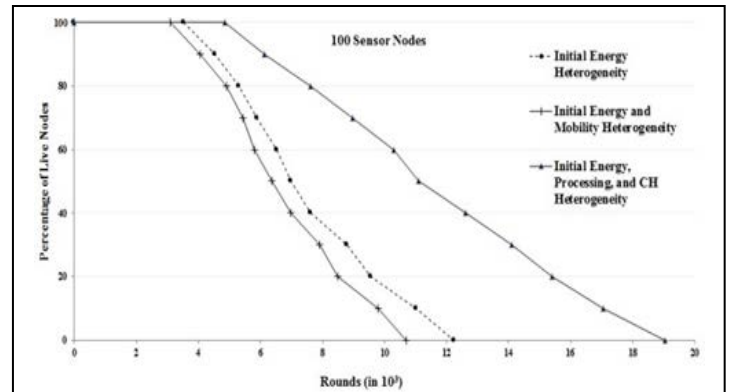


Fig. 7 (a). Network lifetime for high density scenario

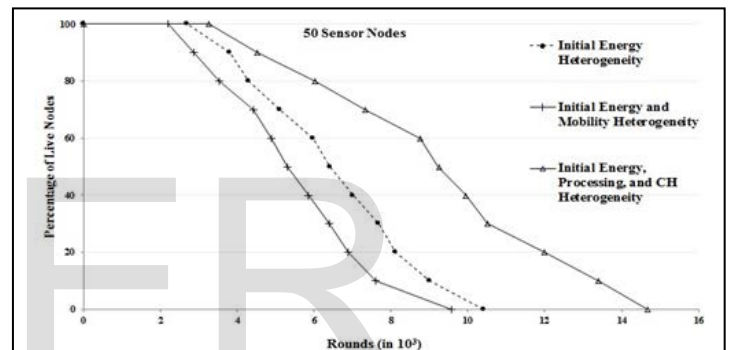


Fig. 7 (b). Network lifetime for low density scenario

To evaluate M-DCHGA efficiency, table VII lists the average time (in seconds) and standard deviation used to form clusters in each transmission round. The time reported is before the first node became unavailable due to energy exhaustion. Despite the standard deviation increased when the number of sensor nodes was doubled, the average time was very close for all cases. It is evident that the efficiency of DCHGA is mostly independent from sensor mobility and number of sensors. The overall average time across all experiments is 0.6 seconds with a standard deviation of 0.06. The efficiency of M-DCHGA method is also satisfactory.

TABLE VII. AVERAGE TIME (IN SECONDS) USED TO IDENTIFY OPTIMAL NETWORK STRUCTURE IN EACH ROUND USING DCHGA AND M-DCHGA COMPARED WITH THE FIVE STATE-OF-ART METHODS.

Methods	100 Sensor Nodes		50 Sensor Nodes	
	Average		Average	
DCHGA	0.63	0.27	0.54	0.06
M-DCHGA	0.67	0.31	0.56	0.05

5 CONCLUSIONS

In homogeneous WSN, clustering protocols assumed that all the sensor nodes are supplied with the same characteristics i.e., initial energy. However, placing few heterogeneous nodes in WSN, such as nodes with more computing powers, is an effective way to increase network lifetime and reliability. In a heterogeneous WSN, in addition to the network structuring factors, e.g., distance to the base-station, and distance among nodes, factors such as initial energy, data processing capability, ability to serve as cluster head, node mobility greatly influence the network lifespan. In addition, the lifetime of the network is maximized when the remaining energy of nodes in the network remains the same during the network lifetime. This is, however, difficult to achieve in a real-world WSN due to different roles of sensor nodes and various signal transmission distance.

In this paper, we propose a heterogeneous sensor node clustering method based on Genetic Algorithm called DCHGA to optimize the energy exhaustion. In DCHGA, the structure of the network is dynamically decided after each message transmission round. In addition, it provides a framework to integrate multiple heterogeneity factors, i.e., initial energy, data processing capability, ability to serve as cluster head, node mobility. Compared with state-of-the-art methods, DCHGA greatly extended the network life and the average improvement with respect to the second best performance based on the first-node-die and the last-node-die were 33.8% and 13%, respectively. While in case of sensors mobility, the improvement was between 12.6% and 9.8%. The computational efficiency of DCHGA is comparable to other algorithms and the overall average time across all experiments was 0.6 seconds with a standard deviation of 0.06.

REFERENCES

- [1] D. Xie, Q. Zhou, X. You, B. Li, and X. Yuan. A novel energy-efficient cluster formation strategy: From the perspective of cluster members. *IEEE Communications Letters*, 17(11):2044-2047, NOVEMBER 2013.
- [2] M. Elhoseny, X. Yuan, H. ElMinir, and A. Riad. Extending self-organizing network availability using genetic algorithm. In *ICCCNT*, Hefei, China, July 2014.
- [3] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *The Hawaii International Conference on System Sciences*, Maui, Hawaii, Jan 2000.
- [4] G. Smaragdakis, I. Matta, and A. Bestavros. SEP: a stable election protocol for clustered heterogeneous wireless sensor network. In *Second International Workshop on Sensor and Actor Network Protocols and Applications*, 2004.
- [5] B. Elbhiri, S. Rachid, and S. Elfkhi. Developed distributed energy-efficient clustering (DDEEC) for heterogeneous wireless sensor. In *Communications and Mobile Network*, pages 1-4, Rabat, 2010.
- [6] A. Kashaf, N. Javaid, Z. Khan, and I. Khan. TSEP: Threshold-sensitive stable election protocol for WSNs. In *Conference on Frontiers of Information Technology*, pages 164-168, 2012.
- [7] A. Iqbal, M. Akbar, N. Javaid, S. Bouk, M. Ilahi, and R. Khan. Advanced LEACH: A static clustering-based heterogeneous routing protocol for WSNs. *Journal of Basic and Applied Scientific Research*, 3(5):864-872, 2013.
- [8] M. Elhoseny, X. Yuan, Z. Yu, C. Mao, H. El-Minir, and A. Riad. Balancing energy consumption in heterogeneous wireless sensor networks using genetic algorithm. *IEEE Communications Letters*, pp(99):1, 2014.
- [9] T. Sudeep, N. Kumar, and J. Niu. EEMHR: energy-efficient multi-level heterogeneous routing protocol for wireless sensor networks. *International Journal of Communication Systems*, 27(9):1289-1318, 2014.
- [10] K. Tripathi, N. Singh, and K. Verma. Two-tiered wireless sensor networks - base station optimal positioning case study. *IET Wireless Sensor Systems*, 2(4):351-360, December 2012.
- [11] L. Wang, C. Wang, and C. Liu. Optimal number of clusters in dense wireless sensor networks: A cross-layer approach. *IEEE Transactions On Vehicular Technology*, 58(2):966-976, February 2009.
- [12] D. Kumar, T. Aseri, and R. Patel. EEHC: energy efficient heterogeneous clustered scheme for wireless sensor network. *Computer Communications*, 32(4):662-667, 2009.
- [13] N. Tuah, M. Ismail, and K. Jumari. Energy efficient algorithm for heterogeneous wireless sensor network. In *IEEE International Conference on Control System and Computing and Engineering*, pages 92-96, Penang, November 2011.
- [14] N. Javaid, N. Mohammad, K. Latif, U. Qasim, A. Khan, and M. Khan. HEER: hybrid energy efficient reactive protocol for wireless sensor networks. In *Saudi International Electronics and Communications and Photonics Conference*, pages 1-4, Riyadh, April 2013.
- [15] S. Ehsan, K. Bradford, M. Brugger, B. Hamdaoui, Y. Kovchegov, D. Johnson, and M. Louhaichi. Design and analysis of delay-tolerant sensor networks for monitoring and tracking free-roaming animals. *IEEE Transactions on Wireless Communications*, 11(3):1220-1227, 2012.
- [16] P. Ali, H. Mashhadi, and S. Javadi. An optimal energy efficient clustering method in wireless sensor networks using multiobjective genetic algorithm. *International Journal of Communication Systems*, 26(1):114-126, 2013.
- [17] S. Bayrakl and S. Erdogan. Genetic algorithm based energy efficient clusters in wireless sensor networks. *Procedia Computer Science*, 10:247-254, 2012.
- [18] B. A. Attea and E. A. Khalil. A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks. *Applied Soft Computing*, 12(7):1950-1957, 2012.
- [19] Y. Wu and W. Liu. Routing protocol based on genetic algorithm for energy harvesting-wireless sensor networks. *IET Wireless Sensor Systems*, 3(2):112-118, June 2013.
- [20] S. Hussain, A. Matin, and O. Islam. Genetic algorithm for energy efficient clusters in wireless sensor networks. In *The 4th International Conference on Information Technology ITNG*, pages 147-154. IEEE, April 2007.
- [21] C. Diallo, M. Marot, and M. Becker. Single-node cluster reduction in WSN and energy-efficiency during cluster formation. In *the 9th IFIP Annual Mediterranean Ad Hoc Networking Conference*, France, 2010.
- [22] L. Chengfa, Y. Mao, C. Guihai, and W. Lie. An energy-efficient unequal clustering mechanism for wireless sensor networks. In *IEEE International Conference on Mobile Adhoc and Sensor Systems*, Washington, DC, Nov 2005.
- [23] W. Guo and W. Zhang. A survey on intelligent routing protocols in wireless sensor networks. *Journal of Network and Computer Applications*, 38:185-201, 2014.

- [24] G. Ahmed, N. Khan, and R. Ramer. Cluster head selection using evolutionary computing in wireless sensor networks. In Progress in Electromagnetics Research Symposium, pages 883–886, 2008.
- [25] N. Bhaskar, B. Subhabrata, and P. Soumen. Genetic algorithm based optimization of clustering in ad-hoc networks. International Journal of Computer Science and Information Security, 7(1):165–169, 2010.
- [26] M. Asim and V. Mathur. Genetic algorithm based dynamic approach for routing protocols in mobile ad hoc networks. Journal of Academia and Industrial Research, 2(7):437–441, 2013.
- [27] A. Karimi, S. Abedini, F. Zarafshan, and S. Al-Haddad. 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, 2013.
- [28] K. Rana and M. Zaveri. Synthesized cluster head selection and routing for two tier wireless sensor network. Journal of Computer Networks and Communications, 13(3), 2013.
- [29] K. Kannammal, T. Purusothaman, and M. Manjusha. An efficient cluster based routing in wireless sensor networks. Journal of Theoretical and Applied Information Technology, 59(3), 2014.
- [30] M. Shirmohammadi, K. Faez, and M. Chhardoli. LELE: leader election with load balancing energy. In International Conference on Communications and Mobile Computing, pages 106–110, 2009.
- [31] M. Elhoseny, H. ElMinir, A. Riad and X. Yuan. Recent Advances of Secure Clustering Protocols in Wireless Sensor Networks. International journal of Computer networks and Communication Security. 2(11):400-413, 2014.
- [32] M. Elhoseny, H. ElMinir and A. Riad. Secure Routing in Wireless Sensor Networks: A State of the Art. International journal of Computer Applications. 67(7):7-12, 2013.

IJSER