CrossMark

A Genetic Algorithm-Based, Dynamic Clustering Method Towards Improved WSN Longevity

Xiaohui Yuan[1,2] · Mohamed Elhoseny[2,3] ·
Hamdy K. El-Minir[4] · Alaa M. Riad[3]

Abstract The dynamic nature of wireless sensor networks (WSNs) and numerous possible cluster configurations make searching for an optimal network structure on-the-fly an open challenge. To address this problem, we propose a genetic algorithm-based, self-organizing network clustering (GASONeC) method that provides a framework to dynamically optimize wireless sensor node clusters. In GASONeC, the residual energy, the expected energy expenditure, the distance to the base station, and the number of nodes in the vicinity are employed in search for an optimal, dynamic network structure. Balancing these factors is the key of organizing nodes into appropriate clusters and designating a surrogate node as cluster head. Compared to the state-of-the-art methods, GASONeC greatly extends the network life and the improvement up to 43.44 %. The node density greatly affects the network longevity. Due to the increased distance between nodes, the network life is usually shortened. In addition, when the base station is placed far from the sensor field, it is preferred that more clusters are formed to conserve energy. The overall average time of GASONeC is 0.58 s with a standard deviation of 0.05.

✉ Xiaohui Yuan
xiaohui.yuan@unt.edu

1 College of Information and Engineering, China University of Geosciences, Wuhan 430074, China

2 Department of Computer Science and Engineering, University of North Texas, Denton, TX 76210, USA

3 Department of Information Systems, Mansoura University, Mansoura 35516, Egypt

4 Department of Electrical Engineering, Kafr El-Sheikh University, Kafr El-Sheikh 33511, Egypt

1 Introduction

Forming network clusters is an effective way to improve the scalability and longevity of wireless sensor networks (WSN) [1–4]. In a cluster-based sensor network, data from cluster members are relayed by a node, namely cluster head, to the base station, which is usually outside the sensor field and could be at a far distance. Ideally the lifetime of a homogeneous WSN is maximized when the residual energy of nodes in the network are about the same, i.e., no single node completely depletes its energy before the others. This is, however, difficult to achieve in a real-world cluster-based networks due to variations in energy consumption as different roles of sensor nodes and various signal transmission distances. The nodes serving as cluster heads consume additional energy to fulfill tasks such as receiving messages from the 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 [5, 6].

Energy consumption of a node is attributed to data acquisition, processing, and transmission. In a complex network, factors such as distance among nodes, distance to the base station, and data throughput greatly influence the residual energy of each node. Methods have been developed that account for one or more such factors to achieve extended network longevity. Low energy adaptive clustering hierarchy (LEACH) [7] is developed to allow different nodes to share the burden of relaying messages. Rotating the role as the cluster head among nodes in the cluster, balances the load and provides a means for fault-tolerance [8]. Energy-efficient unequal clustering (EEUC) [9] selects the cluster head based on the distances between the node and the base station. Leader election with load balancing energy (LELE) [10] takes the residual energy into consideration and nodes with more energy are more likely to serve as cluster head. In [11], the number of neighbors, residual energy, and distance to the base station are used in constructing clusters and selecting cluster heads. Without even looking into the characteristics of each single node, the network itself varies greatly in the number of sensor nodes, their placement, and the arrangement of the base station. A pre-determined communication structure or a randomized clustering scheme is far from fulfilling the critical needs of extending the network longevity. Despite the great efforts in automatically organizing sensor nodes, the dynamic nature of sensor networks and numerous possible cluster configurations make searching for an optimal network structure on-the-fly an open challenge.

In this article, we propose a genetic algorithm-based, self-organizing network clustering (GASONeC) method that provides a framework to integrate multiple factors and optimize dynamic node clustering. In GASONeC, the residual energy, the expected energy expenditure, the distance to the base station, and the number of nodes in the vicinity are employed to search for an optimal, dynamic network structure. Balancing these factors is the key for organizing nodes into appropriate clusters and designating a surrogate node as the cluster head. The factors are encoded into the fitness function of the genetic algorithm (GA). In the optimization

process, each GA chromosome represents a designation map of cluster heads. Given these cluster heads, the clusters are then formed following the nearest neighbor rule, and the fitness of a WSN is therefore determined by the evaluation of all clusters. In each transmission round, the network structure is updated dynamically to achieve a balance of residual energy of sensor nodes and hence, extend network longevity.

The contribution of this work is twofold: first, a GA-based clustering method is proposed that provides a network encoding scheme and flexible optimization approach that employs a fitness function to integrate energy and spatial factors. The GA chromosomes encode the selection of cluster heads and dynamic clusters are formed accordingly. The separation of the cluster head selection and network structure make this method versatile for taking various factors into consideration. Second, the expected energy expenditure is derived, together with other energy and spatial metrics, to achieve balanced energy consumption across all nodes and improve network longevity.

The rest of this article is organized as follows. Section 2 presents the related work of constructing clusters to extend network life. Section 3 describes our proposed GASONeC method in detail. To be self-contained, we include a brief review of GA. Section 4 discusses our experimental results in comparison to the state-of-the-art methods. Section 5 concludes this article.

2 Related Work

To balance the energy expenditure among nodes, the LEACH was proposed to create clusters of sensor nodes dynamically [7]. The LEACH method randomly selects nodes as cluster heads and forms clusters accordingly. The randomized process avoids election of certain nodes as the cluster head that could exhaust the energy of these nodes much earlier than the others could. Despite its effectiveness of diversifying cluster heads, the resulted clusters are far from optimal. Some early improvements utilized node distance property. Power-efficient gathering in sensor information systems (PEGASIS) [12] restricts each node to communicate only with a close neighbor to form a node chain, and the nodes within a cluster take turns to transmit data to the base station. In EEUC [9], the distance between sensor nodes and the base station is used to select the cluster head. The clusters closer to the base station are smaller in size than those farther from the base station. To expand the network coverage, Nadeem et al. [13] proposed to divide nodes into four logical regions based on their locations in the field. A gateway node is placed at the center of the area. If the distance of a node to the base station (or the gateway) is less than a threshold, this node uses a direct communication; otherwise, a cluster is formed and the cluster head is selected probabilistically. Nayak and Devulapalli [14] extended the LEACH method by introducing fuzzy logic in creating clusters.

Closely related to the node distance is the degree of node, i.e., the number of neighboring nodes. In [15], the degree of connectivity is the main factor of selecting a cluster head. Intuitively, a node with more neighbors is a good candidate to serve as the cluster head, since a node with a low degree of connectivity received a less amount of data from its neighbors to aggregate and to forward to the base station. In

the initial phase, each node is involved in the neighborhood information exchanges, i.e., hello protocol, that allows it to determine its degree of connectivity and the location of the base station.

Alternatively, measures of energy have been used in search for the optimal cluster structure. In a stable election protocol (SEP) [16], nodes are allowed to be initialized with different amounts of energy. It uses weighted probabilities to select cluster heads according to the residual energy. Developed distributed energy-efficient clustering (DDEEC) [17] improves upon SEP by categorizing nodes based on their energy level. The nodes with higher energy are named the "advanced nodes" and are those from which cluster heads are selected. Threshold sensitive stable election protocol (TSEP) [18] extends SEP by grouping nodes into three different energy levels, and the cluster head is selected based on a threshold. Modified LEACH (MODLEACH) [19] introduces a cluster head replacement scheme using a threshold. If the residual energy of a cluster head exceeds this threshold, it continues to serve as the cluster head for the next round. Arunraja et al. [20] proposed a distributed energy efficient clustering algorithm that selects cluster heads based on residual energy. Nodes with the least communication costs and high residual energy are included in a cluster.

More recently, combining energy and distance metrics has been extensively explored to identify the optimal network structure. WCA [21] aggregates the degrees of nodes, transmission power, mobility, and battery power with weights in every round. The nodes with the least weighted sum are chosen as the cluster heads. A hybrid energy efficient distributed (HEED) clustering method [22] combines the residual energy with the distance of nodes to their neighbors or the number of neighbors in the process of selecting cluster heads. LELE [10] method selects a cluster head based on the residual energy and the distance between a node and its neighbors. The node with the largest amount of energy and the number of neighbors is chosen as the cluster head. In [11], cluster heads for the subsequent rounds are selected based on the number of neighbors, residual energy, and distance to the base station. In [23], a two-level fuzzy method is used that includes factors of local and global levels. In the local level, the node's capability to serve as a cluster head is evaluated based on its energy and the number of neighbors; whereas, in the global level, three parameters have been considered: centrality, closeness to the base station, and the distance between cluster heads. In [24], a fuzzy logic approach is used for cluster head selection by considering the number of neighboring nodes, residual energy, energy dispersion and distance to the base station.

Searching for a balance among the aforementioned factors is non-trivial and optimization methods are usually used. A GA has been applied in the routing protocol of a WSN [25–28]. A key objective is to define an appropriate fitness function that encodes the network structure and its goodness. A genetic algorithm and weighed clustering algorithm (GA-WCA) [29] consist of two phases: first, a GA is used to determine the cluster heads based on their locations; second, a WCA algorithm forms the clusters by assigning each node to a cluster head, which depends on the load balancing factor and the total distance from all neighbor nodes to the cluster heads. Seo et al. [30] employed the distance in fitness function to select a cluster head. In the energy harvesting genetic-based unequal clustering-

**Table 1** Summary of the state-of-the-art methods

Methods	Strengths	Weakness
LEACH [8]	Dynamic CH selection	Fixed number of clusters and random CH selection
PEGASIS [14]	Applicable to multi-hope WSNs	Static CH selection
EEUC [5]	Applicable to single and multi-hop networks	Static CH selection and fixed number of clusters
M-GEAR [13]	Suitable for large WSN and works with heterogeneous WSN	Required network gateways
Diallo [6]	Balanced cluster size	Static CH selection
SEP [22], TSEP [18], DDEEC [7], MODLEACH [15]	Works with heterogeneous WSN	Required network gateways
WCA [31]	Allow node mobility	small WSN
HEED [27]	Dynamic CH selection	Fixed number of clusters and inefficient processing time
LELE [21]	Dynamic CH selection	Unbalanced load for CHs
GA-WCA [18]	Balanced nodes distribution among CHs. Allow node mobility	Inefficient processing time
LA2DGA [20]	Efficient CHs selection.	Fixed number of clusters and unbalanced load for CHs
EHGUC-OAPR [26]	Efficient cluster formation	Fixed number of clusters
GABEEC [2]	Dynamic CH selection	Fixed number of clusters

optimal adaptive performance routing (EHGUC-OAPR) method [28] the base station uses an EHGUC algorithm to form unequal size clusters followed by an OAPR algorithm to construct an optimal routing among cluster heads. In genetic algorithm based energy efficient clusters (GABEEC) [26] clusters are created in the set-up phase and remain the same without changes. The GA is used to select cluster heads in each round. The selection of new cluster heads is based on the residual energy of the current cluster heads and the member nodes (Table 1).

3 GASONeC for WSN Clustering

3.1 Sensor Energy Model

We adopt the first order radio model (as shown in Fig. 1) for describing energy expenditures in a wireless sensor. The energy expenditure E of a node s is the total energy used to acquire, process, transmit, and receive data:
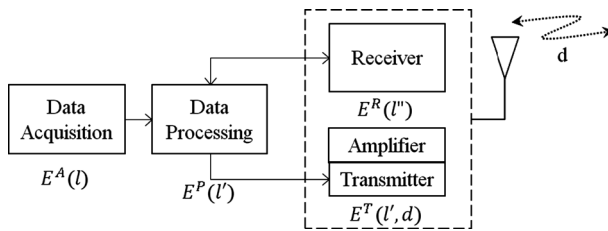
Fig. 1 First order radio model. The energy expenditure model is a function of message length

$$E_s = E_s^A(l) + E_s^P(l') + E_s^R(l'') + E_s^T(l', d), \tag{1}$$

where $E_s^A(l)$ denotes the energy used to acquire l bits of data, $E_s^P(l')$ denotes the energy used to process l' bits of data, $E_s^R(l'')$ denotes the energy used to receive l'' bits of data from nearby nodes if node s serves as a cluster head, and $E_s^T(l', d)$ denotes the energy of transmitting l' bits of data over a distance d. $E_s^R(l'')$ is non-zero only if node s serves as a cluster head. We assume the amount of energy used by each node in data acquisition is the same.

When computing the energy to transmit or receive a message of l bits, we adopt the following formulas for the transmitter energy consumption E_s^T and receiver energy consumption E_s^R:

$$E_s^T = E_i + ld^n \tag{2}$$

$$E_s^R = E_i + lE^* \tag{3}$$

where E_i is the idle energy expenditure and d is the distance between the transmitter and receiver as shown in Fig. 1. Depending on the distance between the transmitter and receiver, the transmission energy consumption E^T is proportional to a different order of the distance and can be modeled by varying the power term n. In this model, we use $n = 4$ for long distance transmission, i.e., transmitting messages from the cluster head to the base station, and $n = 2$ for short distance transmission, e.g., a node its cluster head. The communications within the cluster take the short distance transmission; the communications between the cluster head and base station require long distance transmission. E^* in Eq. (3) represents the cost of the beam forming approach to reduce the energy consumption.

Based on the consumed energy, the residual energy of a node s at time t is computed as follows:

$$\tilde{E}_s(t) = E_s(0) - \sum_{i=1}^{t} E_s(i), \tag{4}$$

where $E(0)$ is the initial energy of the node and t is the time in term of transmission rounds. The residual energy of every node is updated in each round.

Nodes are stationary and their geospatial locations are known to the base station. In each network transmission round, the network structure is computed based on the

current network status and is broadcast to the nodes. The nodes receive the cluster head assignments and communicate to the base station via the associated cluster heads.

3.2 Energy and Spatial Factors for Longevity

To extend the longevity of a WSN, the key is to avoid some nodes depleting energy before the others. Ideally, network life is maximized when nodes retain the same amount of energy throughout the entire network lifespan. A widely employed energy metric is the residual energy [22]. The nodes with a lower amount of residual energy are less preferred to serve as cluster heads due to the extra energy needed to relay messages from the member nodes. However, the residual energy presents the current status and does not characterize the energy level of a node after the next transmission round, which is a direct measure of the goodness of a network clustering structure. Unless the clustering structure remains the same for the life of the network, the dynamically formed local communication requires different amounts of energy contributions from the to-be cluster head. Hence, it is crucial to gauge the expected energy expenditure for each node in a possible network structure.

Assume l bits of data are collected by each node in a round. In a cluster that consists of N_s nodes, data are aggregated by the head node s. Following Eqs. (1) and (2), the expected energy consumption (denoted with $\hat{E}$) of a non-cluster head node s' and a head node s can be computed as follows:

$$\hat{E}_{s'} = E + lD^2(s', s), \tag{5}$$

$$\hat{E}_s = E + N_s lE^* + (N_s + 1)lD^4(s, BS) \tag{6}$$

where E is the constant energy expenditure that includes energy used for data acquisition, processing and being idle. Functions $D(s', s)$ and $D(s, B)$ give the distances between nodes s' and the cluster head s and from node s to the base station B, respectively, which are approximated with Euclidean distance.

In addition to the energy metrics, a network spatial feature is another influencer to the network life. The spatial characteristics such as the distance to the base station and local node density (LSD) provide additional dimensions for intelligent clustering. It becomes more important when the amount of the residual energy of nodes is very similar, in which case the best network structure is as good as random clustering if spatial factors are not considered.

To characterize the local node density, a neighborhood distance threshold δ is used. The density is proportional to the number of neighbors within the δ-vicinity as follows:

$$G_s(\delta) \propto \|S_s\|, \quad \text{and} \quad S_s = \{s_i; D(s, s_i) \le \delta\} \tag{7}$$

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

3.3 Dynamically Structuring Sensor Network Using GA

Structuring a sensor network requires grouping nearby nodes into clusters and designating a surrogate node as the cluster head to each cluster. It is essentially a multi-parameter optimization problem, in which energy and spatial factors need to be optimized concurrently to achieve overall network longevity.

In GASONeC, a binary chromosome is used to specify the cluster heads in the network, in which a one represents a cluster head and a zero represents a member node to a cluster. An example is shown in Fig. 2 and the cluster heads are highlighted with filled circles. The dash arrows depict the cluster membership. When a node becomes inactive, i.e., out of power, its corresponding gene value is set to -1, which exempts the node from further GA operations.

The mapping to node clusters from a chromosome is to minimize the network communication distance $\mathbb{D}$ as follows:

$$\mathbb{D} = \sum_{i=1}^{C} \sum_{j=1}^{N_{s_i}} D\big(s_i, s_j\big) \qquad (8)$$

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 $\mathbb{D}$ is equivalent to assigning nodes to clusters following the nearest neighbor rule.

3.3.1 Fitness Function

A key component of the GASONeC method is the fitness function. As discussed in Sect. 3.2, there are many factors that influence the network's life. Fitness function provides a means to optimize several factors concurrently. The following is a fitness function that integrates energy and spatial factors:

Fig. 2 GA chromosome and the mapping to network clusters
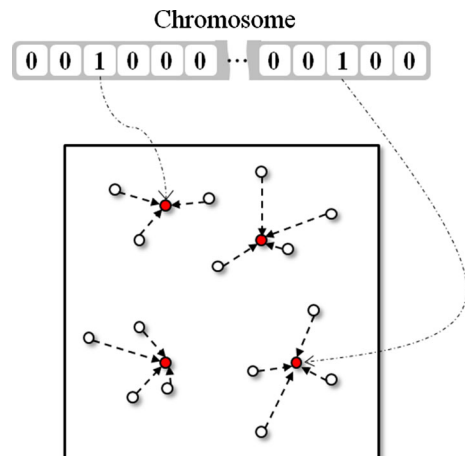
$$f_1 = \sum_s \frac{E_s(t)}{E_s(0)} + \frac{\tilde{E}}{\hat{E}} + \frac{1}{\hat{D}}, \qquad (9)$$

where $E_s(t)$ is the residual energy of node s at round t and $E_s(0)$ is the initial energy of node s. $\tilde{E}$ is the total energy cost if the messages are transmitted directly from all nodes to the base station and $\hat{E}$ is the expected energy expenditure. $\hat{D}$ is the total distance between the cluster heads and the base station:

$$\hat{D} = \sum_{i=1}^{C} D(s_i, BS) \qquad (10)$$

where each s_i is a node that serves as a cluster head. The terms are normalized to the range of (0 1). This objective function includes normalized residual energy, the ratio of expected energy expenditures, and the inverse of the total distance between cluster heads and the base station. When distance is considered, a node that is closer to the base station is usually preferred. That is, maximizing the inverse of the distance gives us a solution with minimum distance. Similarly, we aim to minimize the expected energy expenditures, and an inverse operation is used. As to the residual energy, the goal is to maximize it. Yet, we should avoid possible bias induced by its dynamic range. Hence, a normalization is performed. Without knowledge of the priority of the factors, we assume these three terms are equally important and therefore the fitness function takes even weights. However, in cases where it is clear one or more factors play a more vital role, uneven weights can be employed in the fitness function.

Figure 3 depicts a scenario in the middle of network life. Table 2 gives the states of the nodes in this exemplar scenario. The nodes marked with red served as cluster heads in the previous transmission rounds, and their remaining energy is relatively

Fig. 3 An exemplar scenario of network restructuring
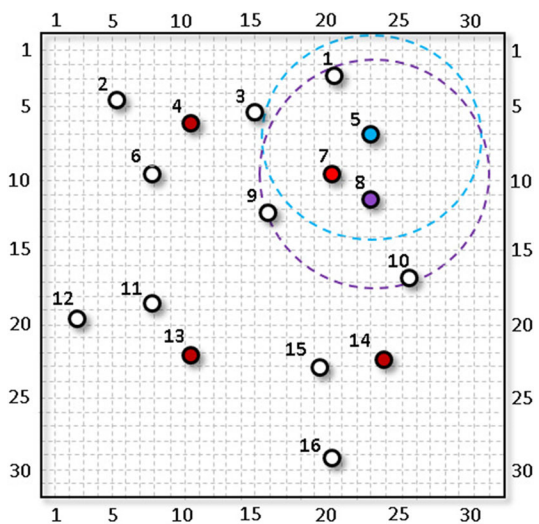
Table 2 States of the network nodes in the exemplar scenario in Fig. 3

Idx	Nodes loc.	Resiudal energy	Expected energy expenditure		CH distance to base station	
			5 as CH	8 as CH	5 as CH	8 as CH
1	3, 20	0.55	0.011	0.55	0	0
2	4, 5	0.6	0.011	0.011	0	0
3	5, 15	0.58	0.01	0.01	0	0
4	6, 11	0.11	0.55	0.05	30	30
5	7, 23	0.7	0.04	0.01	30	0
6	9, 8	0.56	0.01	0.01	0	0
7	9, 20	0.15	0.01	0.01	0	0
8	11, 23	0.65	0.01	0.05	0	33
9	12, 16	0.55	0.55	0.011	0	0
10	17, 25	0.58	0.011	0.011	0	0
11	18, 8	0.6	0.011	0.011	0	0
12	19, 2	0.5	0.55	0.55	0	0
13	22, 10	0.1	0.06	0.06	40	40
14	22, 24	0.11	0.08	0.06	40	40
15	23, 19	0.62	0.01	0.01	0	0
16	29, 20	0.6	0.55	0.55	0	0

low. In this new round, node 5 and node 8 are potential cluster heads to be selected. Based on the residual energy and distance to the base station, node 5 is at an advantage. Its residual energy is at 0.7 and distance to the base station is at 30. If node 5 serves as the cluster head, the cluster headed by node 5 includes nodes 1, 7, and 8, as depicted with the blue circle in Fig. 3. If node 8 serves as the cluster head, its cluster includes 1, 5, 7, 9, and 10. That is, node 9 is no longer a member of the cluster headed by node 4, and node 10 is no longer a member of the cluster headed by node 14. Hence, the expected energy for those nodes changes, as shown in Table 2. Let $\tilde{E}$ be 0.5. The fitness value for the chromosome that specifies nodes 4, 5, 13, and 14 as the cluster head is 1.51 and the fitness value for the chromosome that specifies nodes 4, 8, 13, and 14 as the cluster head is 1.57. Hence, node 8 is more likely to be selected as the cluster head. Since node 8 alleviates the burden for node 14, which is already at a fairly low energy level, the network structure is clearly a better choice. The inclusion of the expected energy expenditures allows a global balance among nodes.

Alternatively, we take into account the local node density as shown in Eq. 7 and have the following fitness function:

$$f_2 = \sum_s \frac{E_s(t)}{E_s(0)} + \frac{\tilde{E}}{\hat{E}} + \frac{1}{\hat{D}} + \frac{1}{N} \sum_{s'} G_{s'}(\delta), \tag{11}$$

where s' denotes nodes serving as cluster heads. Including node density favors the choice of cluster heads with more close neighbors. Note that the first term on the

residual energy is an aggregation of energy of all nodes including cluster heads and member nodes; whereas, the last term includes only the nodes serving as cluster heads.

Algorithm 1 Genetic Algorithm-based, Self-Organizing Network Clustering.

1: Generate a pool of P chromosomes $U = \{u_1, u_2, \ldots, u_P\}$.
2: $\forall u_i \in U$, structuring the WSN by minimizing Eq. (8)
3: Evaluate the fitness of each $u_i \in U$ following Eqs. (9) or (11).
4: **for** $q = 1, 2, \ldots, Q$ **do**
5: $\quad \tilde{U} \Leftarrow \emptyset$
6: $\quad$ **for** $p = 1, 2, \ldots, P$ **do**
7: $\quad\quad$ Randomly select $u_a, u_b \in U$ ($a \neq b$) based on the normalized fitness $\tilde{f}(u)$:

$$\tilde{f}(u) = \frac{f(u)}{\sum_c f(u)}.$$

8: $\quad\quad$ Cross over u_a and u_b according to α

$$\mathcal{C}(u_a, u_b | \alpha) \Rightarrow u_a', u_b'.$$

9: $\quad\quad$ Perform mutation on u_a' and u_b' according to β

$$\mathcal{M}(u_a' | \beta) \Rightarrow \tilde{u}_a, \quad \mathcal{M}(u_b' | \beta) \Rightarrow \tilde{u}_b.$$

10: $\quad\quad$ Evaluate $f(\tilde{u}_a)$ and $f(\tilde{u}_b)$.
11: $\quad\quad$ $\tilde{U} \Leftarrow \tilde{U} \cup \{\tilde{u}_a, \tilde{u}_b\}$
12: $\quad$ **end for**
13: $\quad$ $U \Leftarrow \{u_i; u_i \in \tilde{U} \text{ and } f(u_i)\}$
14: **end for**
15: Return the chromosome u^* that satisfies

$$u^* = \arg\max_u f(u), u \in U$$

3.3.2 Genetic Algorithm-Based, Self-Organizing Network Clustering

Algorithm 1 presents GASONeC method. In this algorithm, $q \in [1, Q]$ denotes the number of generations, and the population size is P. The pool of chromosomes, denoted by U, is initialized with randomly generated individuals. An intermediate pool of chromosomes, denoted by $\tilde{U}$, is used to hold the individuals created in a generation. In the 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 the mutation operations, the value of a randomly picked gene is altered between 0 and 1 according to the mutation probability β.

GASONeC starts with randomly initialized chromosomes and clusters are then formed for an instance of network structure based on each chromosome following Eq. (8). In addition, each chromosome is evaluated according to the objective

function in Eqs. (9) or (11). Such initialization provides multiple random network structures for the optimization process.

In the evolution process, an intermediate pool of chromosomes, denoted with U, are created. GA operations are performed to randomly select individuals. If a crossover probability allows, a pair of chromosomes are used to generate two new individuals by exchanging part of the genes in the chromosomes. If a mutation probability allows, mutation is performed over an individual by randomly changing the value of a gene, i.e., change a randomly selected bit from 1 to 0 or vice versa. The three GA operations continue until a specified number of chromosomes are generated for U. In most cases, the size of this intermediate pool of chromosomes is greater than the population size for GA so that better breeds are more likely to survive. Each newly generated chromosome is evaluated according to the objective function to get a fitness value. A random selection with respect to the fitness value is used to decide which chromosomes continue to the next round of evolution.

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

We assume the population size of the intermediate pool is R and the number of the generation is Q. Since the size of the intermediate pool is much greater than or equal to the population size P, the complexity of GA is $O(RQ)$. In addition, the evaluation of a chromosome is an independent process of the GA, which is a function of two sequential processes: node clustering and network evaluation. In the clustering process, each node needs to compute the distance to the cluster head. Hence, the complexity is $O(NC)$, where N is the number of nodes and C is the number of clusters. The cluster evaluation is of $O(C)$, and the overall complexity is $O(NCRQ)$. In practice, $R, Q \ll N, C$, we have $O(NC)$ as a reasonable complexity approximation.

4 Experimental Results and Discussion

4.1 Experimental Settings

In this evaluation, the WSN has the following properties:

- There is one base station that receives data from nodes;
- Nodes are stationary and their locations are known;
- Provided with sufficient energy, each node can directly reach the base station;
- The characteristics and initial energy of each node are the same.

Table 3 lists the network parameters used in these experiments. In running GA, we use the population size of 30 for 30 generations. The crossover probability and mutation probability are 0.8 and 0.006, respectively. Note that a large mutation rate may lead to loss of good solution; whereas a small crossover rate could hinge the

Table 3 Network properties

Properties		Values
Number of nodes		100
Initial node energy		0.5 J
Idle state energy		50 nJ/bit
Data aggregation energy		5 nJ/bit
Amplification energy	$d \geq d_0$	10 pJ/bit/m^2
(Cluster head to base station)	$d < d_0$	0.0013 pJ/bit/m^2
Amplification energy	$d \geq d_1$	$E_{fs}/10 = E_{fs1}$
(Node to cluster head)	$d < d_1$	$E_{mp}/10 = E_{mp1}$
Packet size		400 bits

convergence [32]. The neighborhood distance threshold δ is 20 m when LSD is calculated.

The average performance of ten repetitions is reported. In each experiment, nodes are randomly placed in the field and the base station is also randomly placed at a certain distance to the field center. Comparison studies are conducted with eight state-of-the-art methods including LEACH [7], MODLEACH [19], HEED [22], PEGASIS [12], SEP [16], TSEP [18], M-GEAR [13], DDEEC [17], GA-WCA [29], and GABEEC [26], among which GA-WCA and GABEEC are methods that employ GA.

4.2 Network Longevity

To maximize the network life, no node shall deplete its energy before any others. That is, between transmission rounds, the residual energy of all nodes is at the same level. In practice, however, this is infeasible, especially when the data volume transmitted within a round is large. Nevertheless, by restructuring the network, the discrepancy among nodes can be suppressed and the network life then improves.

Figure 4 illustrates the average residual energy (in terms of percentage with respect to the initial energy) of nodes in the field using the GASONeC method. In these experiments, 100 nodes were randomly placed in a field of 100 m by 100 m, and the base station was positioned in the center of the field (Fig. 4a) and at the field boundary (Fig. 4b). The fitness function used in this experiment included energy terms and the distance to the base station, i.e., the fitness function in Eq. (9). Ten experiments were conducted with randomly placed nodes, and the average energy levels of all nodes at the transmission rounds of 500, 1000, and 1500 were visualized.

It is evident that the residual energy of all nodes is mostly at the same level throughout the network's life .[1] Table 4 lists the variance of the residual energy of these two exemplar cases. The energy discrepancy among the nodes increases as the transmission continues. This is inevitable in practice due to the geospatial difference

---

[1] We use the number of rounds between the start of the network until the first node becomes unavailable as the network life.
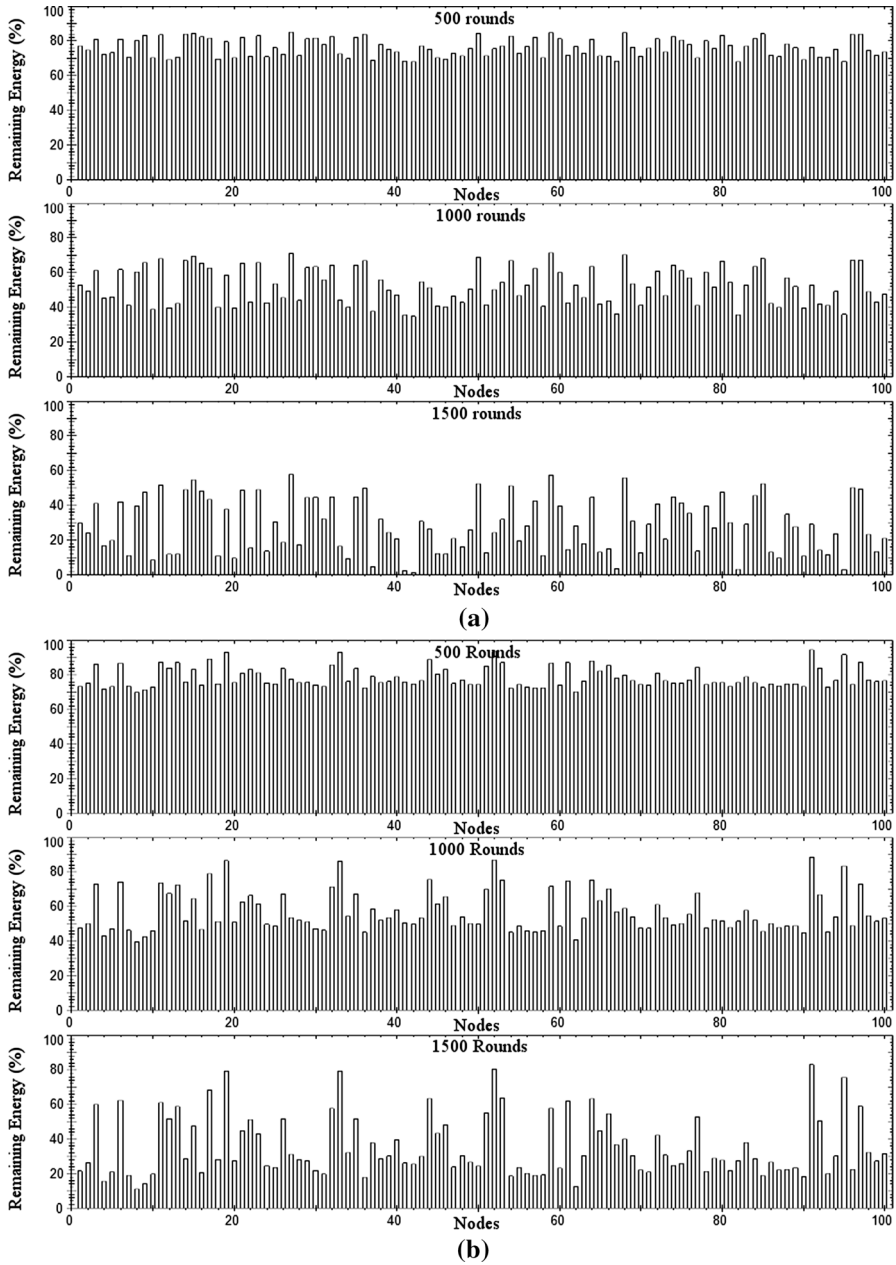
Fig. 4 The percentage of the residual energy of the nodes. **a** Network with base station placed in the center of the field, **b** network with base station placed in the boundary of the field

Table 4 Residual energy (J) variance of nodes

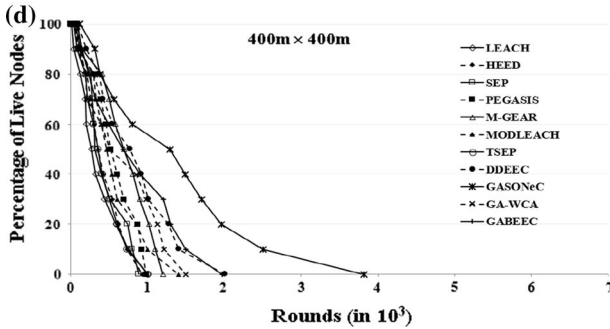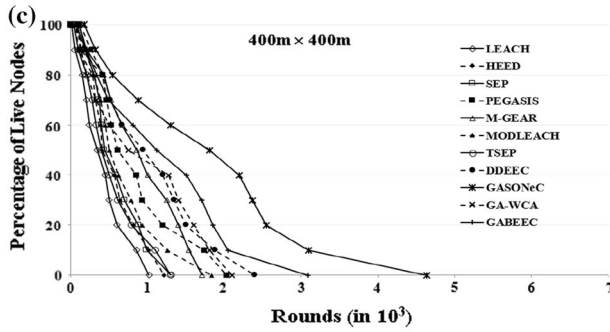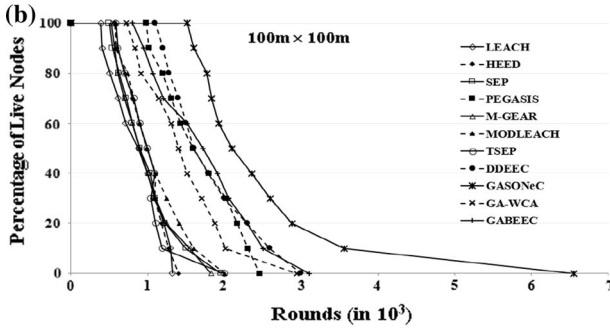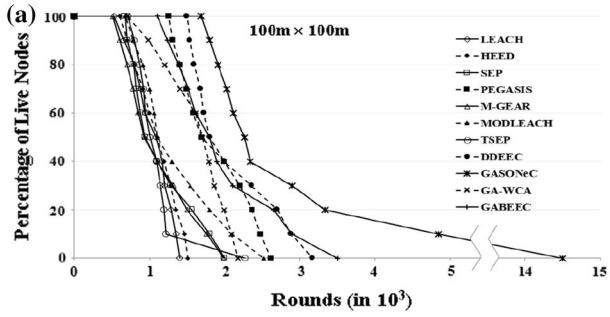| Rounds	Base-station at center			Base-station at boundary		
	500	1000	1500	500	1000	1500
Mean	0.393	0.285	0.179	0.381	0.261	0.141
SD	0.003	0.059	0.088	0.026	0.053	0.078

Note the initial energy is $0.5J$

among the nodes. On the other hand, if we allow more generations in GA, it is possible to identify a better network structure to minimize the variance of the residual energy. As a consequence, it takes longer time to come to a solution, which impacts the efficiency of the network.

Figure 5 depicts the average number of live nodes throughout the entire network lifespan using the state-of-the-art methods. In Fig. 5a, b, 100 nodes are randomly placed in a 100m by 100m field. Figure 5a illustrates results with the base station placed in the field center, and Fig. 5b illustrates the results with the base station placed on the field boundary. Figure 5c, d illustrate results of a sensor field with randomly placed nodes in an area of 400 m by 400 m. The base station locations for Fig. 5c, d are at the field center and on the field boundary, respectively. In these plots, the x-axis shows the number of network transmission rounds (in 1000s); whereas the y-axis shows the percentage of live nodes. Note that the number of nodes deployed in the field is the same in all cases (see Table 3). As the network transmission continues, the number of live nodes decreases as more nodes deplete their energy.

The node density greatly affects the network longevity. Figure 5a, c shows the same placement of the base station, but the node density in Fig. 5c, i.e., field 400 m by 400 m, is much lower. Due to the increased distance between nodes, the network life is shortened. A similar pattern can also be observed when the base station is placed on the boundary of the field, as shown in Fig. 5b, d. Essentially, the density is a transformed view of distance. As the distance between nodes increases, the advantage of forming node clusters becomes less significant in extending the network life.

Placement of the base station also has an impact on the network life, particularly to GASONeC method. Comparing Fig. 5a, b, the network life using GASONeC is much longer when the base station is placed in the center of the field. The other methods are not affected as much as GASONeC although the decrease in network life is observed with the base station placed on the field boundary.

Table 5 presents the average network transmission round of different methods when the first node became unavailable due to energy exhaustion. By comparing the results with respect to the base station placement, it is clear that as the distance between nodes and the base station increases the average network life is shortened. This is mostly due to the extra energy required to forward data to the base station. If we compare the results with respect to the field size, networks in a larger area (i.e., sparse sensor network) tend to have shorter lifespan. In these experiments, GASONeC facilitated the greatest number of network transmission rounds. In

◄**Fig. 5** Network lifetime in terms of network transmission rounds. **a** Base station in the center of a 100 m by 100 m field. For clear visualization and consistence among all figures, the x-axis in **a** is truncated. The number of rounds when the last node depletes all its energy for GASONeC is 14,500. **b** Base station on the boundary of a 100 m by 100 m field. **c** Base station in the center of a 400 m by 400 m field. **d** Base station on the boundary of a 400 m by 400 m field

Methods	100 m by 100 m		400 m by 400 m	
	Center	Boundary	Center	Boundary
M-GEAR	510	530	102	80
LEACH	528	390	10	7
HEED	603	570	65	50
SEP	672	500	70	54
MODLEACH	670	586	50	43
TSEP	710	585	35	33
PEGASIS	1255	980	91	78
DDEEC	1496	1100	122	92
GASONeC	**1691**	**1515**	**175**	**112**
Improvement	13 %	37.3 %	43.4 %	21.7 %

Table 5 The average number of network transmission rounds before the first node became unavailable due to energy exhaustion

The greatest average number of network transmission rounds are given in bold

contrast to the second best results, the maximum improvement rate of network life is 43.4 %.

4.3 Local Node Density

Communication to the cluster head in a dense sensor area usually costs less energy for the member nodes. Hence, selecting a node that has more close neighboring nodes to serve as a cluster head is advantageous to reduce energy expenditures. In the GASONeC method, we devise a fitness function that includes the energy terms as well as local node density to facilitate joint optimization. Figure 6 illustrates the percentage of live nodes using fitness functions with local node density, i.e., Eq. (11), and without local node density, i.e., Eq. (9).

Figure 6a shows the results of a sensing field of 100 m by 100 m. The solid line with circles depicts the performance of using fitness with local node density; whereas the dash line with solid dots depicts the performance of using fitness without local node density. One hundred nodes are randomly placed in the field, and the average of ten experiments is reported. It is clear that by including the local node density in the GA fitness function, the network life is significantly extended. The network transmission rounds at which the first and the last nodes become unavailable are reported in Table 6. The network life improvement is about fourfold with respect to the round of first node discontinuation.

However, when the field size increases to 400 m by 400 m, the advantage of including local node density disappears, as shown in Fig. 6b. The lifespans of WSNs using these two fitness functions are almost the same. The network
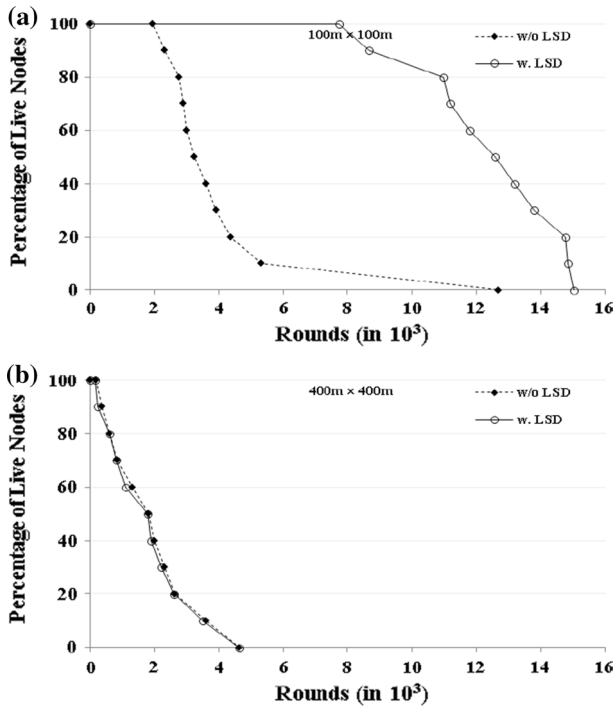
Fig. 6 The percentage of live nodes using fitness function with/without the number of neighbors. The base station is located in the center of the field. **a** An area of 100 m by 100 m. **b** An area of 400 m by 400 m

Table 6 The network life span with and without using the number of neighbors in the fitness function: FND—the round at which first node dies, LND—the round at which the last node dies

Sensor field	Without local node density		With local node density	
	FND	LND	FND	LND
100 m by 100 m	1943	12,688	7752	15,034
400 m by 400 m	175	4633	150	4610

transmission rounds at which the first and last nodes became unavailable in a field of 400 m by 400 m are reported in Table 6, which shows the closeness of the performances. Since the number of nodes in these two cases is the same, it is evident that using local node density yields a better performance in a dense sensor network.

4.4 Impact of the Base-Station Location to Network Structure and Life

Placement of the base station determines the energy expenditure of the nodes; it also affects the network structure. Figure 7 depicts the network clusters with different
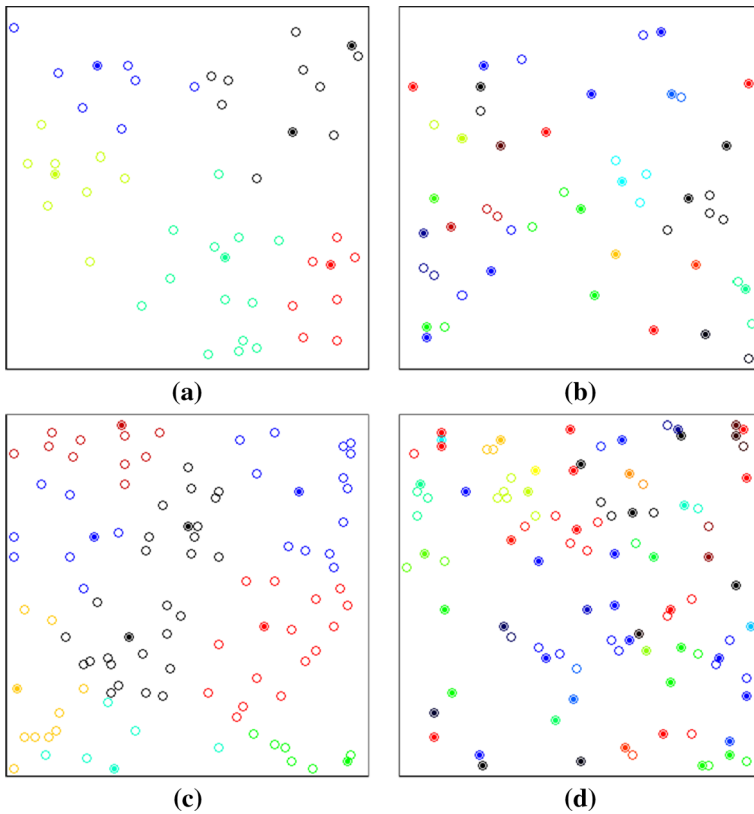
Fig. 7 Node placements and clusters (depicted in different *colors*). The sensor field is a 100 m by 100 m area. **a** Six clusters are formed out of 50 nodes with the base station at the field center. **b** Twenty-six clusters are formed out of 50 nodes with the base station at a distance of 300 m from the field center. **c** Nine clusters are formed out of 100 nodes with the base station at the center. **d** Fifty-six clusters are formed out of 100 nodes with the base station at a distance of 300 m from the field center

base station placements. Figure 7a, b shows two cases with the base station placed in the field center and at 300 m away from the center. The cluster heads are highlighted with solid dots and the other nodes are marked with colored circles. Each color represents a cluster. These demonstrate the early network clusters when the energy level of nodes is close to the maximum. The number of clusters increased when the base station was placed far from the sensor field. A similar phenomena was observed with a denser sensor field with 100 nodes in a 100 m by 100 m area, as shown in Fig. 7c, d.

Intuitively, when the base station is placed far from the sensor field, it is preferred that clusters are formed to conserve energy. However, this contradicts maximizing the network life because a significant amount of energy is usually required by the cluster head to relay messages from its member nodes. Table 7 lists the average number of clusters in a sensor field of 100 m by 100 m. It is clear that including the local node density in the GA fitness function makes little difference in the number

of clusters. In all cases, the cluster counts are very close. Despite that the number of clusters increases when the number of nodes increases, the proportion of the cluster count and node count remains similar. As the distance to the base station enlarges, the number of clusters increases. With the goal of minimizing the energy discrepancy among the nodes, small clusters are favored by GASONeC when the base station is far from the field.

Table 8 presents the average network life with respect to the base station distance to the sensor field. Both fitness functions with and without local node density are considered in this comparison. Clearly the network life reduces when the base station is placed farther away from the sensor field. It is, however, interesting to observe that as the number of nodes increases, the network life is slightly improved. Despite that the volume of data acquired and transmitted within the network increases, the larger number of nodes helps to subside the demand of long-distance data transmission, which facilitates extended network life. Also in this 100 m by 100 m field, local node density also serves as a factor to extend network life when the base station is relatively closer. As shown in the table, when the distance of the base station to the field center is >200 m, the network life (with respect to the first node die) based on the fitness function with the local node density that greatly decreases. As the distance to the base station increases, the energy consumption of the cluster head increases dramatically. LSD favors clusters with many nearby sensors, which usually results in large clusters. For a large cluster, however, the amount of data to be transmitted greatly magnifies the energy cost (recall that the energy cost of a long distance transmission is proportional to d^4), which drains the energy of the cluster head quickly and hence, exhausts the energy of those cluster heads. On the other hand, the effect of such a distance is trivial when the base station is close. In these cases, using local node density improves network life.

Figure 8 illustrates the average network life (in terms of network transmission rounds) with respect to the node density. Figure 8a, b depict the results with 50 nodes in the 100 m by 100 m field, and Fig. 8c, d depict the results with 100 nodes in the same field. It is evident that the location of the base station dramatically affects the network life. When the base station is placed far away from the sensor

Table 7 The average number of clusters in a sensor field of 100 m by 100 m

Base-station distance (m)	100 nodes		50 nodes	
	Without LSD	With LSD	Without LSD	With LSD
300	58	60	29	30
250	51	49	18	21
200	28	32	14	16
150	18	20	10	10
100	11	12	6	7
0	6	8	4	4

The distance of the base station is in respect to the center of the field

Table 8 Average network life with respect to the base station locations

| Base-station distance (m)	Without LSD		With LSD	
	FND	LND	FND	LND
50 nodes				
300	269	1927	110	2429
250	646	2086	315	3860
200	875	2700	1169	2229
150	1199	5851	3304	8326
100	1682	10,434	4986	11,566
0	2005	12,206	7838	16,448
100 nodes				
300	317	2139	231	2663
250	628	2486	588	4307
200	870	2875	1429	2778
150	1412	5661	3507	8681
100	1796	9748	5084	12,818
0	2044	13,624	8549	18,411

The sensor field size is 100 m by 100 m

field, more energy is consumed to transmit data, which shortens the network life. Comparing the network life change between the base station placed at 0 (i.e., center of the field) and 100 (or 100 and 150, etc.) shown in Fig. 8b and that shown in Fig. 8a, such an impact is more significant when the local node density is employed in the optimization. Similar observations can be obtained from Fig. 8c, d.

It is interesting to note that when local node density is used in the optimization, the network life is not inversely proportional to the base station distance. The exceptional cases are shown in Fig. 8b, d when the base station is placed at 200 m and 250 m. The network life is surprisingly longer when the base station is placed at 250 m compared to that of when the base station is placed at 200 m. Note that the location of base station is randomized on the radius of certain distances to the sensor field center and the network life reported is the average of ten independent runs. This is possibly due to the instability of including local node density in the fitness function. In contrast, such an anomaly is absent in Fig. 8a, c.

4.5 Efficiency

In the GASONeC method, the most time consuming process is fitness evaluation. Given a chromosome, clusters are formed according to the nearest neighbor rule and fitness is then evaluated according to the cluster. In our experiments, GA employs 30 chromosomes in any generation and 60 offspring are created. Thirty generations are performed to conclude the optimization. The methods are implemented in C# language and experiments are conducted in a computer with an Intel core i5 2.6 GHz CPU, 4 GB memory, and Windows 7 operating system.

Table 9 lists the average time (in s) and the standard deviation that is used to form clusters in each transmission round by the GASONeC method. The fitness

(a)

(b)

(c)

(d)
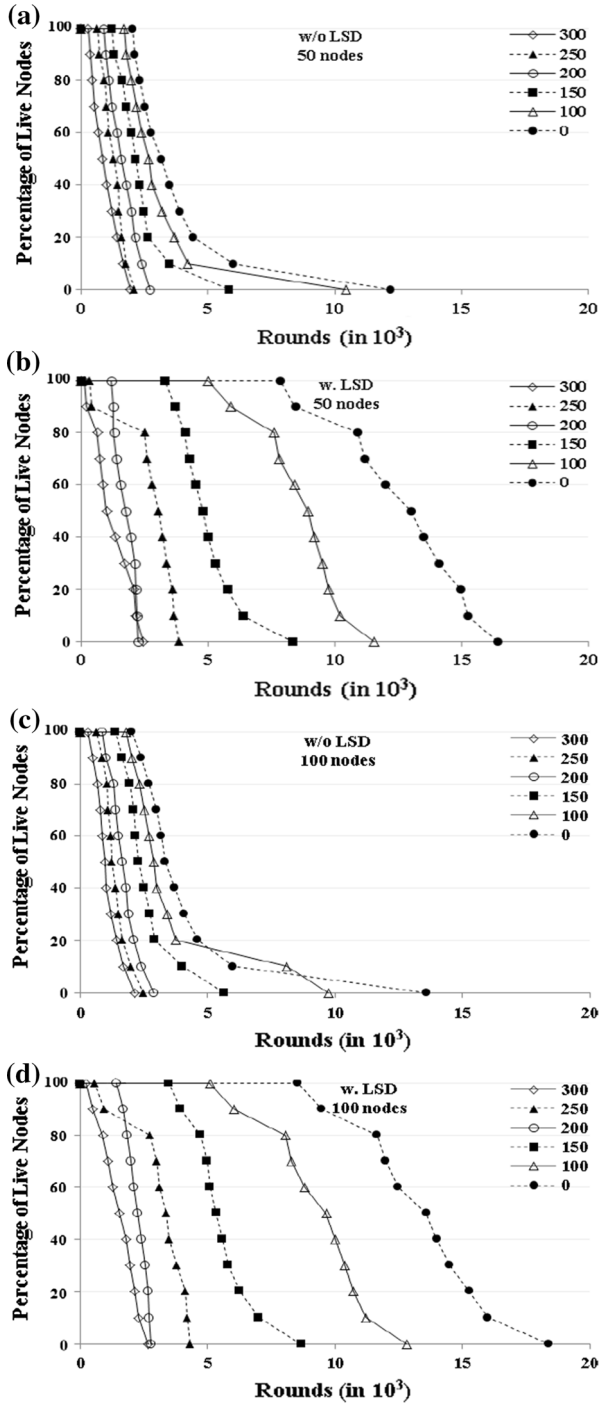
◀**Fig. 8** Network life in a sensor field of 100 m by 100 m. **a** Fifty nodes in the field and the fitness function without local node density. **b** Fifty nodes in the field and the fitness function with local node density. **c** One hundred nodes in the field and the fitness function without local node density. **d** One hundred nodes in the field and the fitness function with local node density

Table 9 Average time (in s) is used to identify optimal network structure in each transmission round using GASONeC

Node count BS to field	50		100		200		400	
	0 m	100 m	0 m	100 m	0 m	100 m	0 m	100 m
Field size: 100 m by 100 m								
Ave. time	0.51	0.57	0.66	0.74	0.89	0.88	0.98	0.89
SD	0.05	0.05	0.39	0.16	0.39	0.12	0.19	0.14
Node count BS to field	50		100		200		400	
	0 m	400 m	0 m	400 m	0 m	400 m	0 m	400 m
Field size: 400 m by 400 m								
Ave. time	0.49	0.47	0.59	0.57	0.65	0.61	0.72	0.63
SD	0.09	0.15	0.36	0.32	0.27	0.11	0.29	0.33

function includes energy terms, distance and local node density. The "BS to field" denotes the distance from the base station to the sensor field center. The time reported is before the first node became unavailable due to energy exhaustion. That is, the number of nodes remains the same. Despite that the standard deviation increases when the number of nodes is doubled, the average time is very close for all cases. It is evident that the efficiency of GASONeC is mostly independent from the field size as well as the number of nodes. The overall average time across all experiments is 0.58 s with a standard deviation of 0.05.

5 Conclusion

Forming network clusters is an effective way of improving the scalability and longevity of WSNs. A pre-determined communication structure or a randomized clustering scheme is far from fulfilling the critical need of maximizing the network life. Despite the great efforts in automatic organizing nodes, the dynamic nature of sensor network and numerous possible cluster configurations make searching for an optimal network structure on-the-fly an open challenge. To address this problem, we propose a GA-based, self-organizing network clustering method that provides a framework to integrate multiple factors and optimize dynamic node clustering. In the GASONeC method, we devise a concise way of encoding nodes and propose fitness functions that include residual energy, expected energy expenditure, distance to the base station, and local node density in search for an optimal, dynamic network

structure. Balancing these factors is the key of organizing nodes into appropriate clusters and designating a surrogate node as the cluster head.

Compared with state-of-the-art methods, the GASONeC method greatly extended the network life and the improvement is up to 43.44 %. The results showed that as the distance between nodes and the base station increases, the average network life is shortened. This is due to the extra energy required to forward data to the base station. Moreover, when the base station is placed far from the sensor field, it is preferred that more clusters are formed to conserve energy. The node density greatly affects network longevity. Due to the increased distance between nodes, the network life is shortened. The average running time of GASONeC is very close for all cases. It is evident that the efficiency of GASONeC is mostly independent from field size and number of nodes. The overall average time across all experiments is 0.58 s with a standard deviation of 0.05. The efficiency of GASONeC is satisfactory.

In future work, we plan to explore the effectiveness of GASONeC in heterogeneous network structures. In addition, the parallel programming will be investigated to reduce the optimization time for constructing the network clusters.

References

1. Li, B., Li, H., Wang, W., Yin, Q., Liu, H.: Performance analysis and optimization for energy-efficient cooperative transmission in random wireless sensor network. IEEE Trans. Wirel. Commun. 12(9), 4647–4657 (2013)
2. Xie, D., Zhou, Q., You, X., Li, B., Yuan, X.: A novel energy-efficient cluster formation strategy: from the perspective of cluster members. IEEE Commun. Lett. 17(11), 2044–2047 (2013)
3. Liao, Y., Qi, H., Li, W.: Load-balanced clustering algorithm with distributed self-organization for wireless sensor networks. IEEE Sens. J. 13(5), 1498–1506 (2013)
4. Elhoseny, M., Yuan, X., Yu, Z., Mao, C., El-Minir, H.K., Riad, A.M.: Balancing energy consumption in heterogeneous wireless sensor networks using genetic algorithm. IEEE Commun. Lett. 19(12), 3194–3197 (2015)
5. Tripathi, K., Singh, N., Verma, K.: Two-tiered wireless sensor networks—base station optimal positioning case study. IET Wirel. Sens. Syst. 2(4), 351–360 (2012)
6. Wang, L., Wang, C., Liu, C.: Optimal number of clusters in dense wireless sensor networks: a cross-layer approach. IEEE Trans. Veh. Technol. 58(2), 966–976 (2009)
7. Heinzelman, W., Chandrakasan, A., Balakrishnan. H.: Energy-efficient communication protocol for wireless microsensor networks. In: The Hawaii International Conference on System Sciences, Maui, Hawaii (2000)
8. Heinzelman, W., Chandrakasan, A., Balakrishnan, H.: An application-specific protocol architecture for wireless microsensor networks. IEEE Trans. Wirel. Commun. 1(4), 660–670 (2002)
9. Chengfa, L., Mao, Y., Guihai, C., Lie, W.: An energy-efficient unequal clustering mechanism for wireless sensor networks. In: IEEE International Conference on Mobile Adhoc and Sensor Systems, Washington, DC (2005)
10. Shirmohammadi, M., Faez, K., Chhardoli, M.: LELE: leader election with load balancing energy. In: International Conference on Communications and Mobile Computing, pp. 106–110 (2009)
11. Raj, E.: An efficient cluster head selection algorithm for wireless sensor networks EDRLEACH. J. Comput. Eng. 2(2), 39–44 (2012)
12. Lindsey, S., Raghavendra, C.: Pegasis power-efficient gathering in sensor information systems. IEEE Aerosp. Conf. Proc. 3, 1125–1130 (2002)
13. Nadeem, Q., Rasheed, M., Javaid1, N., Khan, Z., Maqsood, Y., Din, A.: 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 (2013)

14. Nayak, P., Devulapalli, A.: A fuzzy logic-based clustering algorithm for wsn to extend the network lifetime. IEEE Sens. J. **16**(1), 137–144 (2016)
15. Diallo, C., Marot, M., Becker, M.: Single-node cluster reduction in WSN and energy-efficiency during cluster formation. In: 9th IFIP Annual Mediterranean Ad Hoc Networking Conference, France (2010)
16. Smaragdakis, G., Matta, I., Bestavros. A.: SEP: a stable election protocol for clustered heterogeneous wireless sensor network. In: Second International Workshop on Sensor and Actor Network Protocols and Applications (2004)
17. Elbhiri, B., Rachid, S., Elfkihi, S.: Developed distributed energy-effecient clustering (DDEEC) for heterogeneous wireless sensor. In: Communications and Mobile Network, pp. 1–4, Rabat (2010)
18. Kashaf, A., Javaid, N., Khan, Z., Khan, I.: TSEP: threshold-sensitive stable election protocol for WSNs. In: Conference on Frontiers of Information Technology, pp. 164–168 (2012)
19. Mahmood, D., Javaid, N., Mahmood, S., Qureshi, S., Memon, A., Zaman, T.: MODLEACH: a variant of LEACH for WSNs. In: Eighth International Conference on Broadband and Wireless Computing and Communication and Applications, pp. 158–163 (2013)
20. Arunraja, M., Malathi, V., Sakthivel, E.: Distributed energy efficient clustering algorithm for wireless sensor networks. J. Microelectron. Electron. Compon. Mater. **45**(3), 180–187 (2015)
21. Chatterjee, M., Das, S., Turgut, D.: WCA: a weighted clustering algorithm for mobile ad hoc networks. Clust. Comput. **5**, 193–204 (2002)
22. Younis, O., Fahmy, S.: HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans. Mob. Comput. **3**(4), 366–379 (2004)
23. Torghabeh, N., Akbarzadeh, M., Yaghmaee, M.: Head selection using a two-level fuzzy logic in wireless sensor networks. In: 2nd International Conference on Computer Engineering and Technology, pp. 357–361 (2010)
24. Kannammal, K., Purusothaman, T., Manjusha, M.: An efficient cluster based routing in wireless sensor networks. J. Theor. Appl. Inf. Technol. **59**(3), 683–689 (2014)
25. Bhaskar, N., Subhabrata, B., Soumen, P.: Genetic algorithm based optimization of clustering in ad-hoc networks. Int. J. Comput. Sci. Inf. Secur. **7**(1), 165–169 (2010)
26. Bayrakl, S., Erdogan, S.: Genetic algorithm based energy efficient clusters in wireless sensor networks. Procedia Comput. Sci. **10**, 247–254 (2012)
27. Attea, B.A., Khalil, E.A.: A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks. Appl. Soft Comput. **12**(7), 1950–1957 (2012)
28. Wu, Y., Liu, W.: Routing protocol based on genetic algorithm for energy harvesting-wireless sensor networks. IET Wirel. Sens. Syst. **3**(2), 112–118 (2013)
29. Nandi, B., Barman, S., Paul, S.: Genetic algorithm based optimization of clustering in ad-hoc networks. Int. J. Comput. Sci. Inf. Secur. **7**(1), 165–169 (2010)
30. Seo, H., Oh, S., Lee, C.: Evolutionary genetic algorithm for efficient clustering of wireless sensor networks. In: Sixth IEEE Consumer Communications and Networking Conference, p. 2009 (2009)
31. Ming, Y., Leung, K., Malvankar, A.: A dynamic clustering and energy efficient routing technique for sensor networks. IEEE Trans. Wirel. Commun. **6**(8), 3069–3079 (2007)
32. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Professional, Reading (1989)

Xiaohui Yuan is an Associate Professor in the Department of Computer Science and Engineering at University of North Texas and a visiting Professor at the College of Information and Engineering, China University of Geosciences. His research interests include computer vision, pattern recognition, and artificial intelligence. He is an IEEE senior member and a recipient of Ralph E. Powe Junior Faculty Enhancement award and the Air Force Summer Faculty Fellow.

Mohamed Elhoseny is a lecture in the Faculty of Computers and Information Sciences at Mansoura University, Mansoura, Egypt and a visiting scholar in the Department of Computer Science and Engineering at University of North Texas. His research interests include artificial intelligence and wireless sensor network.

Hamdy K. El-Minir is an Assistant Professor in the Department of Computer Science, Faculty of Computers and Information at Kafrelsheikh University, Kafr El-Shaikh, Egypt. His research interests include image processing, artificial intelligence, and solar energy.

Alaa M. Riad is a Professor and Dean in the Faculty of Computers and Information Sciences at Mansoura University, Mansoura, Egypt. His research interests include intelligent information systems, mobile agents and networks, E-learning, and big data.