Sensor Review

Genetic optimization of hybrid clustering algorithm in mobile wireless sensor networks

Sangeetha M., Sabari A.,

Genetic optimization of hybrid clustering algorithm in mobile wireless sensor networks

Sangeetha M. and Sabari A.
IT Department, K.S.Rangasamy College of Technology, Namakkal, India

Abstract

Purpose – This paper aims to provide a prolonging network lifetime and optimizing energy consumption in mobile wireless sensor networks (MWSNs). MWSNs have characteristics of dynamic topology due to the factors such as energy consumption and node movement that lead to create a problem in lifetime of the sensor network. Node clustering in wireless sensor networks (WSNs) helps in extending the network life time by reducing the nodes' communication energy and balancing their remaining energy. It is necessary to have an effective clustering algorithm for adapting the topology changes and improve the network lifetime.

Design/methodology/approach – This work consists of two centralized dynamic genetic algorithm-constructed algorithms for achieving the objective in MWSNs. The first algorithm is based on improved Unequal Clustering-Genetic Algorithm, and the second algorithm is Hybrid K-means Clustering-Genetic Algorithm.

Findings – Simulation results show that improved genetic centralized clustering algorithm helps to find the good cluster configuration and number of cluster heads to limit the node energy consumption and enhance network lifetime.

Research limitations/implications – In this work, each node transmits and receives packets at the same energy level throughout the solution. The proposed approach was implemented in centralized clustering only.

Practical implications – The main reason for the research efforts and rapid development of MWSNs occupies a broad range of circumstances in military operations.

Social implications – The research highly gains impacts toward mobile-based applications.

Originality/value – A new fitness function is proposed to improve the network lifetime, energy consumption and packet transmissions of MWSNs.

Keywords Energy consumption, Cluster head, Cluster member, Dynamic topology, Network lifetime

Paper type Research paper

1. Introduction

Genetic algorithm (GA) was initially suggested as a search algorithm. GA is also known as a global heuristic algorithm. A GA estimates an optimal elucidation through generating different individuals. Focused fitness function is one of the main procedures of GA. It is taken from natural evolution of new species in natural environment. The natural assessment has the accompanying features:

- the individual characteristics are implied on a chromosome;
- every chromosome has a specific fitness function value as indicated by the location in which it is present;
- individual chromosomes with best fitness value can survive and produce next generations of better individual chromosomes.

In this work, Base Station (BS) is a centralized, high-capacity node capable of co-coordinating the entire communication among nodes in an environment (Gherbi *et al.*, 2017). GA-based cluster heads (CHs) selection is proposed, for better CH selections. On every round, new populations of CHs are selected by BS on evaluating fitness function. In dynamic clustering, i.e. re-clustering, it is significant to avoid early death of CHs. Hence, such efficient CHs can be obtained by applying selection, crossover and mutation processes. This proposed centralized approach is projected to overcome the main limitations of Low Energy Adaptive Clustering Hierarchy (LEACH) protocol (Heinzelman *et al.*, 2002).

In our proposed GA, binary representations are used to represent each sensor node bit by bit. The representation of a node in mobile wireless sensor network (MWSN) is called a Chromosome or Genome, a collection of bits. In this network, cluster members (CMs) are represented as "0" and CHs are represented as "1". Initially, GA starts with any initial random population, a predefined number of chromosomes; each chromosome consists of a potential solution. Then, GA calculates fitness value of each chromosome by using fitness function. After evaluation, GA selects the best fit chromosomes from the population, by using a specific selection method based on their fitness values and then applies crossover and mutation operator, respectively. These procedures are repeated iteratively to obtain a new population better than the previous one.

The current issue and full text archive of this journal is available on Emerald Insight at: **www.emeraldinsight.com/0260-2288.htm**

Sensor Review

[DOI 10.1108/SR-08-2017-0149]

Received 2 August 2017
Revised 11 September 2017
14 December 2017
Accepted 4 January 2018

To improve the lifetime and energy consumption of sensor node, the BS will execute the proposed GA algorithm in each round. The new fitness function is calculated and executed based on some networking parameters, i.e. node residual energy, degree of connectivity, distance from sensor node to BS and mobility factor. The proposed GA encodes the election of CHs and forms the dynamic clusters in an entire network according to the CH distance. Following, the expected energy rate is derived, together with other energy metrics, to achieve balanced energy consumption across all nodes and improve the lifetime of WSN.

Challenges of MWSNs differ from stationary sensor networks (Singh and Soni, 2017). In a stationary sensor network, distance clusters once formed are not disturbed till re-clustering happens. But, using mobile nodes, distance between nodes keeps on changing and node mobility can quickly disturb a good cluster configuration. So, clustering algorithm for mobile sensor nodes needs to create viable clusters so that created clusters do not get disturbed for as much time as possible.

1.1 Problem statement

Consider a multi-hop WSN represented by an undirected graph G = (N, A), where N = {1, . . ., N} is the set of nodes, N is the number of sensor nodes in the network and A is the edge representing the wireless link between a pair of sensors in the sensor network. Each sensor node is assumed to use a fixed transmission range r (meters). The problem is to find in a cluster of N nodes, where H nodes are available for the CH role, the M nodes that minimize the average distance between a node and its assigned CH while at the same time, equalizing the number of nodes and selecting the optimum number of clusters and CH in each cluster. This is a nondeterministic polynomial time hard problem of a discrete and combinatorial nature, which would be solved by hybrid techniques effectively to improve the network lifetime and reduce the energy consumption.

1.2 GA fundamental operations

The GA requires the determination of six fundamental aspects: the initial population creation, chromosome representation, selection of new population, crossover function, mutation and termination are the basic GA operations (Krishna and Murty, 1999). The step-by-step process of GA is given in the following sections.

1.2.1 Initialization

In GA, each chromosome corresponds to a likely solution. The random immigrant scheme generates a new offspring into the population. The initial population in GA is given in equation (1):

$$P_{GA} = \{chr1,\ chr2,\ \ldots.,\ chr\,n-1\} \tag{1}$$

Where P represents population and chr represents chromosome.

1.2.2 Genetic representation of chromosomes

The initial population consists of sensor nodes having chromosomes. Initially, the number of sensor nodes in WSNs is considered as a population. It is denoted by equation (2):

$$Population = \{sn\#1, \mathrm{sn\#2},\ldots, \mathrm{sn\#m}\} \tag{2}$$

Where sn represents sensor node.

Each node in network {sn#1}, {sn#2}, . . ., {sn#m} represents a genetic factor. The chromosome representation of WSNs is as shown in Figure 1. Here, the CH nodes are represented as "1" and member nodes are represented as "0".

1.2.3 Parent selection mechanism

The parent selection mechanism is separation among individual parents based on quality to allow better offspring from cutting edge parents. Parent selection is probabilistic and subsequently the selection process is used to elect the best chromosomes from a generation. There are many selection algorithms like Tournament, Truncation and Roulette Wheel.

In this work, the Roulette Wheel selection algorithm is used to perform the parent selection where fitness of each chromosome is normalized to 1 using the total fitness of all chromosomes (Mekkaoui *et al.*, 2015).

1.2.4 Crossover operator

Recombination or crossover is a binary variation operator which blends data from two parent genotypes into half offspring genotypes. Crossover operator chooses which part of the parents is consolidated and the method used for joining. The policy behind crossover is accomplished by mating two individuals with various, attractive characteristics creating an offspring consolidating both characteristics.

In this work, a two-point crossover operator over individuals is actualized and is implemented. The two-point crossover calls for two points to be selected on the parent organism strings. Everything between the two points is swapped between the parent organisms, versioning two child organisms.

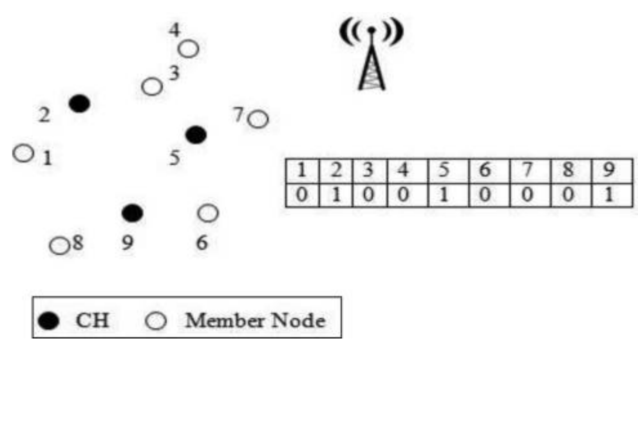
Example: Suppose Parent 1 is **11**010**111** and Parent 2 is 11**101**110 and after performing the crossover, the resulted output contains some part of Parent 1 and other from Parent 2:

- **11**010**111** + 11**101**110 = **11101111 (Offspring 1)**
- 11**010**111 + **11**101**110** = **11010110 (Offspring 2)**

1.2.5 Mutation operator

A single variant operator is known as mutation and connected to one genotype assigning an altered mutant, the child or its offspring. The mutation is used to avoid the super chromosome problem. It means if one chromosome is recursively selected as many times in the same generation, the crossover will not produce new chromosomes, as the parents are still the same.

Figure 1 Representation of chromosomes in WSNs



Hence, the mutation is used to change an arbitrary bit. Several uncertain methods have been performed so as to produce the best-run GA parameters in terms of runtime and convergence.

Example:

- Original chromosome: 11100101
- Mutated chromosome: 1**01**001**1**1

1.2.6 GA termination

GA terminates by selecting the best chromosomes with well-attained qualities of parental ones over evolution. Every GA terminates like this and these chromosomes are considered as CHs in our proposed work, as it is obvious each CH possess better lifetime and energy.

Finally, a centralized GA clustering in WSNs keeps the best chromosomes on every round, thereby holding only "the Best" and neglects others which are determined from their fitness function. Each chromosome, in fact, represents a potential solution.

2. Related work

LEACH (Heinzelman *et al.*, 2002) is a first dynamic clustering protocol. Generally, LEACH has a motivation for many proposed dynamic clustering protocols. In LEACH, the selection of CH is done randomly to uniformly distribute the energy consignment between the sensors nodes in the network. The CHs have responsibility of gathering data from their CMs and aggregating their data for transferring information to BS. These results came up with less energy dissipation, and efficient network lifetime. LEACH consists of a set of rounds. Each round consists of a Setup phase and a Steady-state phase. The nodes are organized into clusters in the Setup phase. The sensed data are transferred to the BS in the Steady-state phase. During cluster formation, each node decides to become a CH for the current round based on the given threshold value. The LEACH algorithm selects the CH based on random probability. This random selection of CH may produce faster death in few nodes. LEACH permits single hop routing only in WSNs. If BS is located far away from the nodes, the energy consumption is more. So, it is unsuitable for a huge network.

Mobility-Based Clustering protocol in WSN elect a CH based on its residual energy and mobility factor (Deng *et al.*, 2011). The role of CH is to allocate time division multiple access (TDMA) schedule for each of its CMs to transmit data in ascending order based on the estimated association time. The CM decides to join a CH by taking into account the assessed connection time, residual energy, node degree of the CH and the distance between the sensor node and the CH. This protocol aims to produce a better adaptability in mobile environment. Also, TDMA scheduling creates the protocol simplex, as new scheduling must be completed and broadcast to the members whenever a node enters or leaves a corresponding cluster.

Energy-Efficient Heterogeneous Clustered (EEHC) scheme is a randomized and distributed clustering algorithm to maximize the lifetime of WSN (Bandyopadhyay and Coyle, 2003). EEHC algorithm is executed in two phases. In the first phase, initial volunteer nodes, may decide to be CHs based on the probability value and they broadcast their decisions to their one hop neighbors to form the initial clusters. In the second level, the clustering algorithm is recursively executed to form hierarchical clustering, where new CHs are selected from the already formed CHs, until a final BS is reached. This approach is effective in prolonging the lifetime of WSN.

Genetic Algorithm-Based Energy Efficient Clusters have been developed to determine the number of clusters, the CHs, the CMs and the TDMA schedules to optimize the lifetime of sensor network (Bayrakh and Erdogan, 2012). This method consists of two phases, which are setup phase and steady-state phase. In setup phase, clustering is executed at first and it is not changed in network lifetime. So, in each round, there are static clusters with dynamically changing CHs. In steady-state phase, selection of CH is done. The selection of the new CH is based on the residual energy of the current CH and its member nodes. Forming optimal clusters and CHs such that has a longer lifetime for static network.

GAs are a kind of optimization and heuristic search algorithms that come from the idea of natural evolution (Alex, 1957). GA covers the concept of chromosomes, inheritance, mutation and crossover in natural evolution. In a GA, the input parameters are treated as chromosome vectors. The GA evaluates the fitness function based on the parameters and selects the best chromosomes as CH for current round using fitness value. The GA performs mutation and crossover of the chromosomes and produces the best offspring for future generation. GAs is very valuable in attaining the global maximum without being trapped in a local optimum because the initial population covers the entire solution space so that local optimized values would be avoided. This algorithm is suitable for finding the global optimization value to improve the network lifetime.

A Genetic K-means Algorithm (GKA) was proposed to find a globally optimal partition of a population into a number of clusters (Krishna and Murty, 1999). To circumvent the local optimum, the hybrid GA was proposed with K-means algorithm. In GKA, K-means algorithm is executed first, and then later GKA applied as a search operator instead of crossover. It defines a biased mutation operator-specific clustering called distance-based mutation. Using finite Markov chain theory, it proves that the GKA converges to the global optimum. The simulation result shows that GKA converges to the best-known optimum corresponding to the given data in concurrence with the convergence result. It is also observed that GKA searches faster than some of the other evolutionary algorithms used for clustering.

Multi-Objective Genetic Algorithm (MOGA) was proposed to optimize the coverage and lifetime of an entire WSN (Jourdan and De Weck, 2014). In this network, all sensor nodes need to transmit data to the BS directly or through multi-hop communication. Data transmission is the main concern of energy consumption in MOGA. So, with different communication range of sensors, the data transmission will have different number of hops by changing the energy consumed. The concept of MOGA deals with the goal of minimizing energy consumption and maximizing coverage using a MOGA to find the optimal tradeoff between coverage and lifetime. Different optimum layouts are obtained with respect to different sensing-to-communication range ratios.

3. Proposed work

Generally almost all GA-based clustering algorithms have been produced with an objective of reducing the energy consumption and improvement of network lifetime (Hussain *et al.*, 2007). In this paper, a centralized dynamic clustering algorithm is proposed (i.e. GA is executed at the BS). At each round, re-clustering is performed to avoid early death of CHs, i.e. each sensor node can then be either a CH or a CM. This centralized approach is anticipated to overcome the main limitations of LEACH protocol, where the number of CHs is fixed and selection is random, routing is single-hop and their spatial distribution is arbitrary, i.e. there is no coordination (Heinzelman *et al.*, 2002).

This work consists of two centralized dynamic GA-constructed algorithms for achieving the objective in WSNs. The first algorithm is based on improved Unequal Clustering-Genetic Algorithm (UC-GA), and the second algorithm is Hybrid K-means Clustering-Genetic algorithm (KC-GA).

3.1 Unequal clustering-genetic algorithm

The UC-GA is proposed to optimize the lifetime and energy consumption of WSNs. In UC-GA, there are two phases used:

1 setup phase (CH selection and cluster formation) and
2 steady-state phase (data transmission).

In the setup phase, numbers of predefined sensor nodes are elected as CHs based on fitness value represented in equation (3). The number of CHs indicates the total number of independent clusters present in the WSNs. Then CMs are associated with the nearest clusters based on their distance to CH. Here, dynamic clustering has been considered, i.e. re-clustering is considered to avoid early death of CHs. At each re-clustering round, each sensor node becomes either a CH or a CM. The setup state for generating new population is represented in Algorithm 1.

In the steady-state phase, all CMs start to communicate with their CHs according to TDMA schedule. After CH receives the data from all CMs, it fuses the data packets into one and forwards it to the BS. When all CHs send their data to BS, a single round is finished. At the completion of each round, the BS checks the energies of CHs and the CMs. The steady-state phase procedure is represented in Algorithm 2.

3.1.1 Proposed UC-GA algorithm

This solution is appropriate for randomly distributed widespread network. At each round, various chromosomes are evaluated using a fitness function. The best fitness value of chromosome is used for generating the next population. After the evaluation step, the population for the next round is made up using Selection, Crossover and Mutation operations.

```
Algorithm 1: Phase 1 - Setup state for generating
  new population
  Input: N chromosomes
  Output: Best chromosomes
  Begin
    Randomly initialize population P(0)
    Repeat
    Evaluate population // Evaluate fitness f,
    for each chromosome
    Select chromosomes based on fitness f(x)
    while the new population is not complete
    do
       Select parent chromosomes // based on
       fitness f(x)
       Crossover (P0(f))
       Mutate (P0(f))
       Place new offspring in a new population, P(1)
       till the termination condition is met
       Use new generated population for future
       rounds
    end
  Return the best population
  End
```

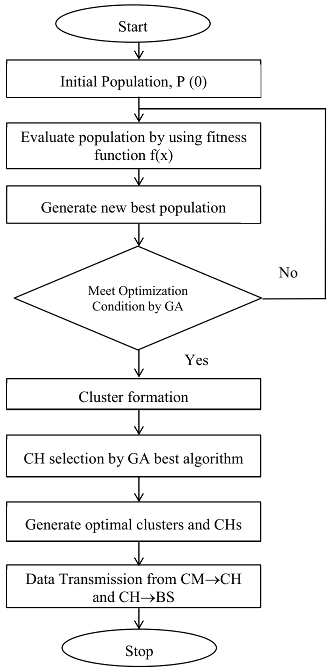
```
  Algorithm 2: Phase 2 - Steady-state UC-GA
  Input: Best chromosomes obtained during phase1
  Output: Optimized Clusters and CHs
  Deploy best sensor nodes generated during Phase 1
  Round = 1
  GA2:
    Nodes (id, mobility, energy & distance) → BS
      BS→Evaluate fitness f(x)
      Selection (CHs)
      Broadcast (CHs→CMs)
       Join (CMs→CHs)
    Unequal cluster formation according to CHs
    strength
    CHs collect data←CMs
    Data Transfer (CHs→BS)
    Round++
    If the required number of rounds is achieved
    or the
    Network dies, then Stop. Otherwise, go to
    GA2.
```

Initially, each node sends its identifier, mobility metric (velocity and position), energy and distance information to BS. BS applies UC-GA algorithm to optimize the clusters and CHs. Then, BS selects the optimal number of CHs out of phase 1. This selected CHs form the initial clusters based on signal strength. TDMA schedule is generated and broadcasted to the CM by each CH. Nodes wake up and sense the data in their timeslot and then forward to respective CH. Finally, the CMs transmit data to their CHs. The CHs aggregate the data received from all the member nodes and transmit the aggregated data to the BS. At the end of each round, energy of each node is checked for data transmission and determines the energy proficient clusters and sensor nodes for the next round. Re-clustering is done after each round is completed. The Figure 2 explains the detailed view of the entire UC-GA algorithm.

3.1.2 Fitness function

Each chromosome is evaluated based on its fitness function which qualities have a higher chance to provide the best

Figure 2 Flowchart for UC-GA algorithm



solution for CH selection. Each node uses mobility metric, energy and distance from the sensors in a cluster and from the CH to the BS. The fitness of a candidate chromosome *f(x)* can be expressed in equation (3):

$$f(x) = \sum_{i=1}^{4} f_i\,,\ \forall f_i \in \{f_1, f_2, f_3 f_4\} \tag{3}$$

The sub-functions f_1, f_2, f_3 and f_4 proposed in equations (4), (5), (6) and (7) together contributes to the fitness function $f(x)$.

For each cluster, energy of each node is represented in equations (4) and (5):

$$f_1 = \sum_{i=1}^{n \in C_k} E_{(i)} \tag{4}$$

$$E_{(i)} = \sum_{i=1}^{C_K} \Big(\Big(E_{\min\{T(i \to CH)\}} + K \times E_{\min R\{CH \to i\}} + E_{\min\{T(CH \to BS)\}} + E_{\min(Sens+idle)} \Big) \tag{5}$$

In equation (4), $E_{(i)}$ denotes the energy of node i. In equation (5), the first term depicts the energy spent to transmit (T) messages from member nodes to CH. The second term shows the energy consumed by the CH to receive (R) messages from the member nodes. The third term represents the energy needed to transmit from the CH to the BS. Finally, the fourth term represents the energy spent whenever the node is sensing in the environment or idle in the environment.

For each cluster, node mobility is represented as in equation (6):

$$f_2 = \sum_{i=1}^{n \in C_k} M_{(i)} \tag{6}$$

where $M_{(i)}$ represents mobility of node(i). The node mobility factor is represented as in equation (7):

$$Mobility\ factor\ M(i) = \sum_{i=1}^{n} \frac{sensor\ current\ field - sensor\ origin\ field}{Mobility\ speed} \tag{7}$$

For a cluster with k sensor nodes, the cluster distance from CH to CM is denoted as in equation (8):

$$f_3 = \min(i = 1, 2 \ldots C_k)\{\sum_{i=1}^{C_K} \frac{d(CM_{ik}, CH_k)}{C_k} \tag{8}$$

where C_k is the cluster count, and $d(CM_{ik}, CH_k)$ defines the distance between CM and CH from each cluster.

For each sensor network containing k CH nodes, the distance from CH to BS is denoted as in equation (9):

$$f_4 = \min(i = 1, 2 \ldots C_k)\Big\{\sum_{i=1}^{C_k} d(CH_k, BS)\Big\} \tag{9}$$

where C_k, is the cluster count, CH_k is the CH count, i.e. $\{CH_1, CH_2, \ldots, CH_k\}$ and $d(CH_k, BS)$ represents distance between CH and BS.

In simulating UC-GA, CH selection is done prior followed by cluster formation. The results obtained are quite convincing too. On the other hand, an inverse process over this methodology is also attempted in our follow up process. Here cluster formation is done followed by CH selection when using K-means Clustering and improved GA (KC-GA). To our surprise, the result obtained through this latter method is more efficient than the former one. To balance energy distribution, a K-means-based algorithm is better suited to dynamically cluster the network. The simulations in NS-2 show the proposed KC-GA algorithm has longer network lifetime than other comparative algorithms like LEACH and UC-GA.

The subsequent section explains the use of KC-GA algorithm in a detailed fashion. This algorithm also executes in two phases, setup state and steady-state phases, with a newer convention as Hybrid KC-GA.

3.2 Hybrid K-means clustering-genetic algorithm

The goal of this proposed methodology is clubbing K-means along with GA. It aims to develop a clustering algorithm for mobile nodes in order to extend network lifetime and reduce energy consumption in WSNs. Figure 3 shows the working principle of UC-GA and KC-GA.

The KC-GA consists of two phases. They are:

1 setup phase and
2 steady-state phase.

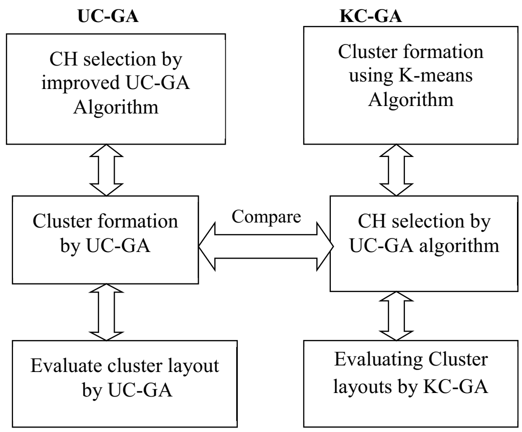
In setup phase, the Hybrid KC-GA cluster formation algorithm is twofold:

- K-means clustering algorithm is used to partition the network into K per cent of clusters; and
- the improved GAs are used to select the best CHs in each cluster obtained by K-means algorithm.

3.2.1 Cluster formation by K-means algorithm

K-means is an unsupervised clustering algorithm which solves several dynamic clustering problems on account of large-scale networks. K-means performs the clustering at very high speeds (Sasikumar and Khara, 2012). It is mainly applied to partition a network into K number of clusters based on the distance

Figure 3 Working principle of UC-GA and KC-GA



between an elected CH and the rest of the nodes in the network. The elected CH always focuses on the centroid (K) of the cluster. The centroid of the cluster is calculated by using the equation (10):

$$Centroid(K) = \left(\frac{1}{n}\right) \times \sum_{j}^{mind^2(x_i,m_j)} \quad ,for\ i = 1 \ldots n \tag{10}$$

where n represents number of nodes, j represents cluster centroids and $d(x_i,m_j)$ signifies the Euclidean distance between the node x_i and mean value m_j denoting the mean or center of the cluster and x_i denotes the location of the nodes in that cluster.

The distance between any two member nodes (m1, m2) is formulated using the Euclidean distance which is shown in equation (11):

$$D(m1, m2) = \sqrt{(x_{m1} - x_{m2})^2 + (y_{m1} - y_{m2})^2} \tag{11}$$

where x and y represent the position of the node's x coordinate and y coordinate, respectively, in the network environment.

K-means Clustering Algorithm: The following steps portray the general K-means clustering algorithm.

- *Step 1*: Arbitrarily choose K points identified as center of clusters. K is defined as total number of clusters expected.
- *Step 2*: The distance between each sensor and the cluster center is calculated and each sensor assigned to the neighboring cluster center.
- *Step 3*: Calculate the new cluster center by calculating the mean distance between each cluster center and all sensors in its range by equation (10).
- *Step 4*: With the new cluster centers, repeat Step 2. If any changes occur in the assigned cluster center, repeat Step 3; else clusters are finalized and the clustering process ends.

3.2.2 Proposed hybrid KC-GA algorithm

Hybrid KC-GA is proposed to find the optimal clusters and CHs to solve the clustering problem of dynamic nodes in WSNs to increase the network lifetime and reduce the energy consumption. The Algorithm 3 shows the process of Hybrid KC-GA.

Initially all nodes send their mobility metric (speed, position), energy level and distance information to the BS. BS executes K-means clustering to partition the network into K clusters. Then BS runs the improved GA algorithm which is applied to elect the optimal number of CH from each cluster obtained by the K-means algorithm. Generally, Chromosome with high fitness value closer to the optimal solution is selected as the CH for next round based on equation (3). After completing setup state, steady-state phase is established in the usual manner, i.e. CHs to BS and vice versa. Finally, this entire process is repeated until termination condition is reached, i.e. when the centroid value is repeated again for the next round.

```
Algorithm 3 Hybrid KC-GA
  Input: Initial population, N
  Output: Optimized Clusters and efficient CHs
  for each round do
  Begin
    Round = 1;
    BS → N sensor nodes//
    BS Requesting node's mobility, energy level,
    Distance information
    BS runs K-means algorithm
    BS dispenses role (CH/CMs) to each sensor node //
    using centroid and distance parameters
    respectively
    Each CH sends one hop communication to CMs
    for initial
    Cluster formation
    BS execute fitness function to select best
    CH from each cluster
    Cluster formation
    Round++
  End
  end for
```

4. Experimental setup

To validate the performance of UC-GA and KC-GA with existing protocol LEACH, there are different metrics to be used. In this work, the metrics number of clusters formed, number of alive nodes, network lifetime, energy consumption and packet delivery ratio have been considered as the QoS parameter for evaluation. Table I shows the simulation parameter of UC-GA and KC-GA.

In each scenario, a given number of sensors are randomly deployed and performed in a monitored area. In this work, simulations were carried out in a 100 $\times$ 100 m^2 network consisting of 200 nodes by using network simulator under various parameters. The coordinate of BS is assumed to be at 50 $\times$ 175. The energy consumption due to communication is calculated using the first-order radio model. It is considered that all homogeneous nodes have batteries with initial energy of 2 Joules and the sensing range is fixed as 25 m. The transmission range of each node is up to 20 m and node mobility is set to 10 m/s.

In these experiments, IEEE 802.11 Distributed Coordination Function with Carrier Sense Multiple Access/ Collision Avoidance is used as the medium access control

Table I Simulation parameters of UC-GA and KC-GA

Simulation parameters	Values
Simulation area	100×100 m^2
Number of nodes (N)	200
Population size	N
Length of chromosome	N
Base station	50×175
Sensing range	25 m
Transmission range	20 m
Propagation model	Two-ray ground
Number of rounds	2000
Crossover type	Two-point
Selection type	Roulette Wheel
Crossover rate	0.5
Mutation rate	0.2
Initial energy	2 Joules
Mobility	Random waypoint model
Node speed	10 m/s
Packet size	512 bytes

protocol, and two-ray ground as the propagation model. The packet size is set to 512 bytes and Constant Bit Rate traffic sources are used to generate traffic with a rate of 1 p/s (packet per second). Each simulation experiment is run for 500 s, and the obtained results are averaged over 2,000 different rounds.

4.1 Results and discussion

The observations from Figure 4 portray that the number of efficient clusters formed for KC-GA is optimal for 200 nodes than UC-GA and LEACH.

The observations from Figure 5 suggest that the lifetime computation for number of nodes alive of KC-GA is better with an increased lifetime than number of nodes alive of UC-GA and number of nodes alive of LEACH. When the number of rounds increases within the network, the lifetime computation decreases based on the alive nodes. In simulation, at the end of 2,000 rounds, 69 of 200 nodes are alive in KC-GA, 58 of 200 nodes are alive in UC-GA and 50 of 200 nodes are alive in LEACH, i.e. KC-GA performed 5.5 per cent better than UC-GA and 9.5 per cent better than LEACH.

Figure 6 shows the comparison of KC-GA, UC-GA and LEACH protocol. For UC-GA and LEACH, 100 per cent of nodes death occurs at 1,637 rounds and 1,689 rounds, respectively, whereas in KC-GA, this occurs only after 1,750 rounds, which means that the cumulative cluster energy level using KC-GA lasts longer than UC-GA and LEACH. The simulation results show that KC-GA protocol proved 3.5 and 6.1 per cent better network lifetime than UC-GA and LEACH. Energy efficiency is the total energy consumed by all the nodes in the network based on transmitting packets and node speed in a dynamic environment. It is concluded from the simulation results that for 200 nodes with 2 Joules of initial energy, KC-GA required 1.58 J, which is less than the 1.65 J of UC-GA and the 1.71 J of LEACH as shown in Figure 7.

Figure 4 Number of cluster formed for UC-GA and KC-GA

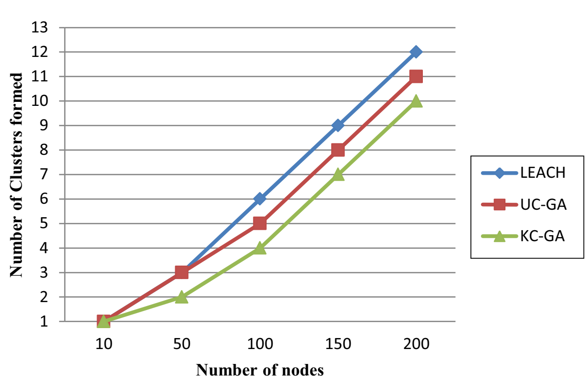


Figure 5 Number of alive nodes for UC-GA and KC-GA

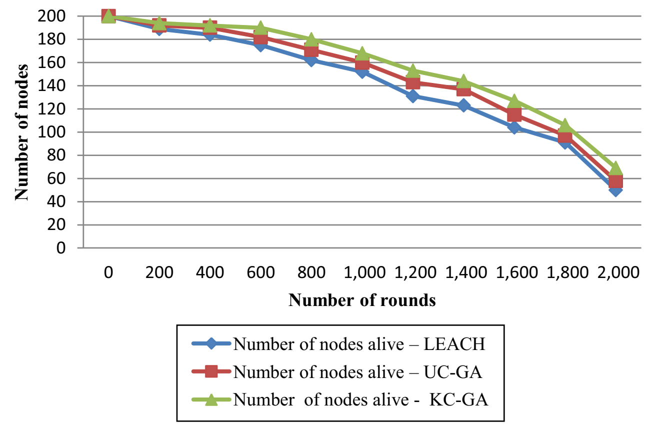


Figure 6 Network life time for UC-GA and KC-GA

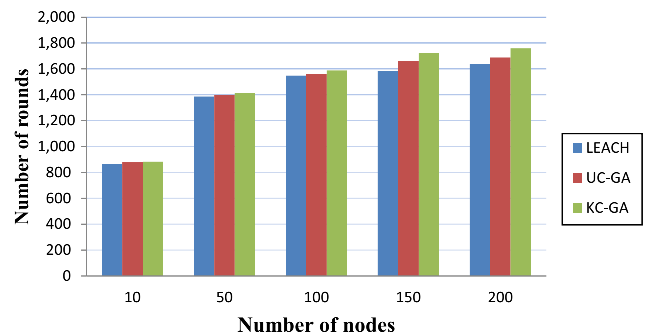


Figure 7 Energy consumption of UC-GA and KC-GA

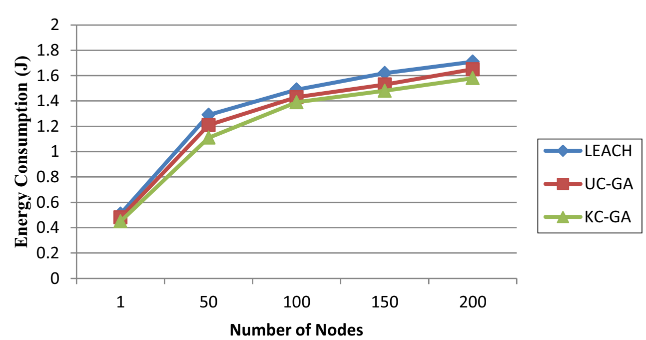
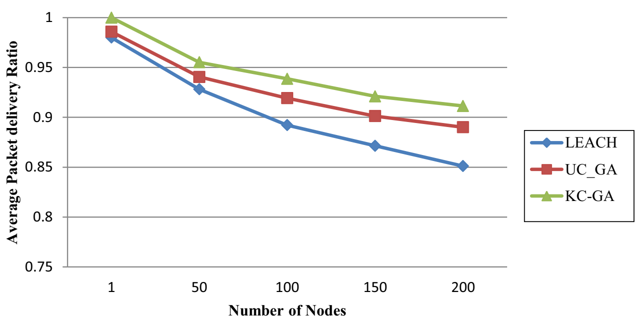


Figure 8 Packet delivery ratio for UC-GA and KC-GA



Packet delivery ratio represents the number of packets that are sent successfully to the BS. The percentage of average packet delivery ratio of KC-GA is better than UC-GA and LEACH by 2.12 and 6 per cent, respectively, for 200 nodes as shown in Figure 8.

5. Conclusion

Certainly, it is a challenging task to handle the dynamic clustering problem in unstable network topology of MWSN. The proposed algorithms KC-GA and UC-GA are implemented in NS-2 for selecting optimal number of clusters and CHs in dynamic environment. The KC-GA and UC-GA algorithms are considered mobility factor, energy and distance metric for calculating fitness function and then applied genetic operation. So, the proposed GAs have more amount of stability in solving dynamic clustering problems and network-based optimization issues. Based on the simulation results, KC-GA performs better by reducing energy consumption and improving network lifetime than UC-GA and LEACH. Finally, KC-GA suits well for dynamic network environment by avoiding faster convergence and obtaining the optimum solution.

References

Alex, F. (1957), "Simulation of genetic systems by automatic digital computers. I. introduction", *Australian Journal of Biological Sciences*, Vol. 10 No. 4, pp. 484-491.

Bandyopadhyay, S. and Coyle, E.J. (2003), "An energy efficient hierarchical clustering algorithm for wireless sensor networks", *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications (INFOCOM), 2003, IEEE*, Vol. 3, pp. 1713-1723.

Bayrakh, S. and Erdogan, S.Z. (2012), "Genetic algorithm based energy efficient clusters (GABEEC) in wireless sensor networks", *Procedia Computer Science, ELSEVIER*, Vol. 10, pp. 247-254.

Deng, S., Li, J. and Shen, L. (2011), "Mobility-based clustering protocol for wireless sensor networks with mobile nodes", *IET Wireless Sensor System*, Vol. 1 No. 1, pp. 39-47.

Gherbi, C., Aliouat, Z. and Benmohammed, M. (2017), "A survey on clustering routing protocols in wireless sensor networks", *Sensor Review*, Vol. 37 No. 1, pp. 12-25, available at: https://doi.org/10.1108/SR-06-2016-0104.

Heinzelman, W.B., Chandrakasan, A.P. and Balakrishnan, H. (2002), "An application-specific protocol architecture for wireless micro sensor networks", *IEEE Transactions on Wireless Communications*, Vol. 1 No. 4, pp. 660-670.

Hussain, S., Matin, A.W. and Islam, O. (2007), "Genetic algorithm for hierarchical wireless sensor networks", *Journal of Networks*, Vol. 2 No. 5.

Jourdan, D.B. and De Weck, O.L. (2004), "Layout optimization for a wireless sensor network using a multi-objective genetic algorithm", *IEEE Semiannual Vehicular Technology, Los Angeles, Vol.* 5, pp. 2466-2470.

Krishna, K. and Murty, M.N. (1999), "Genetic K-means algorithm", *IEEE Transactions on Systems, Man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society*, Vol. 29 No. 3, pp. 433-439.

Mekkaoui, K., Rahmoun, A. and Ferrari, G. (2015), "Genetic centralized dynamic clustering in WSN", *5th International Conference on Computer Science and its Application, 2015: CIIA 2015, Springer*, pp. 503-511.

Sasikumar, P. and Khara, S. (2012), "K-means clustering in wireless sensor networks", *Proceedings of the Fourth International Conference on Computational Intelligence and Communication Networks, 2012: CICN 2012*, pp. 140-144.

Singh, M. and Soni, S.K. (2017), "A comprehensive review of fuzzy-based clustering techniques in wireless sensor networks", *Sensor Review*, Vol. 37 No. 3, pp. 289-304, available at: https://doi.org/10.1108/SR-11-2016-0254

Corresponding author

Sangeetha M. can be contacted at: sangeecse@gmail.com