Research Article

Synthesized Cluster Head Selection and Routing for Two Tier Wireless Sensor Network

Keyur Rana[1] and Mukesh Zaveri[2]

[1] Department of Computer Engineering, Sarvajanik College of Engineering & Technology, Athwalines, Surat 395001, Gujarat, India
[2] Department of Computer Engineering, Sardar Vallabhbhai National Institute of Technology, Ichhanath, Surat 395007, Gujarat, India

Correspondence should be addressed to Keyur Rana; keyur.rana@scet.ac.in

Large scale sensor networks can be efficiently managed by dividing them into several clusters. With the help of cluster heads, each cluster communicates using some routing schedule. It is essential to rotate the role of cluster heads in a cluster to distribute energy consumption if we do not have dedicated high energy cluster heads. Usually routing and cluster head selection for such networks have been separately solved. If cluster heads are selected with the consideration of routing and routing schedule is prepared with the consideration of selected cluster heads, it can help each other. We have proposed an integrated approach of cluster head selection and routing in two tier wireless sensor network (WSN) using Genetic Algorithm based cluster head selection with A-Star algorithm based routing method to extend life of WSN. This approach can lead to significant improvements in the network lifetime over other techniques.

1. Introduction

Wireless Sensor Networks are composed of a large number of sensor nodes with limited resources in terms of energy, memory, and computation. They are operated by a small battery attached to it. This battery has some initial energy, and in every communication it dissipates a fraction of the energy. Many such communications take place during the network lifetime, and every time sensor node consumes some energy which makes battery exhaust eventually. When nodes are deployed in hostile environment or in a kind of environments where it is hard to reach, in most of the cases there is no way to recharge these batteries.

Sensor nodes are used for monitoring physical phenomena like temperature, humidity, acoustic, seismic, video, and so on [1]. For large scale wireless sensor networks, applications exist in a variety of fields, including medical monitoring [2–4], environmental monitoring [5, 6], surveillance, home security, military operations, and industrial machine monitoring [7]. To fulfill the requirements of these applications, sensor network should have a lifetime long enough to cater for several months. How to prolong the network lifetime to

such a long time is the vital question to design and manage sensor network systems.

Randomly deployed sensor nodes in the field collect required data and send towards the base station after processing them. If the optimal path (in terms of energy consumption or quality of service) is chosen for each round of communications, nodes of that particular path may get drained of energy, and network can get partitioned soon. We consider the end of network life as soon as the network gets partitioned as in [8]. Many approaches have been proposed in the literature for routing and for cluster head selection in WSN to extend the lifetime [9–13].

Data transmission in WSN can be single hop or multihop. In either case, data collected by sensor nodes are sent to the base station. The large and dense network is divided into several clusters, and each cluster contains one cluster head that is responsible to collect and send data to the base station. The sensor nodes are grouped into clusters geographically and are capable of operating into two roles. It can work as a sensor node and as the cluster head node [14, 15]. As a sensor node, the node senses the task and sends the sensed data to its cluster head. As a cluster head node, a node gathers data from
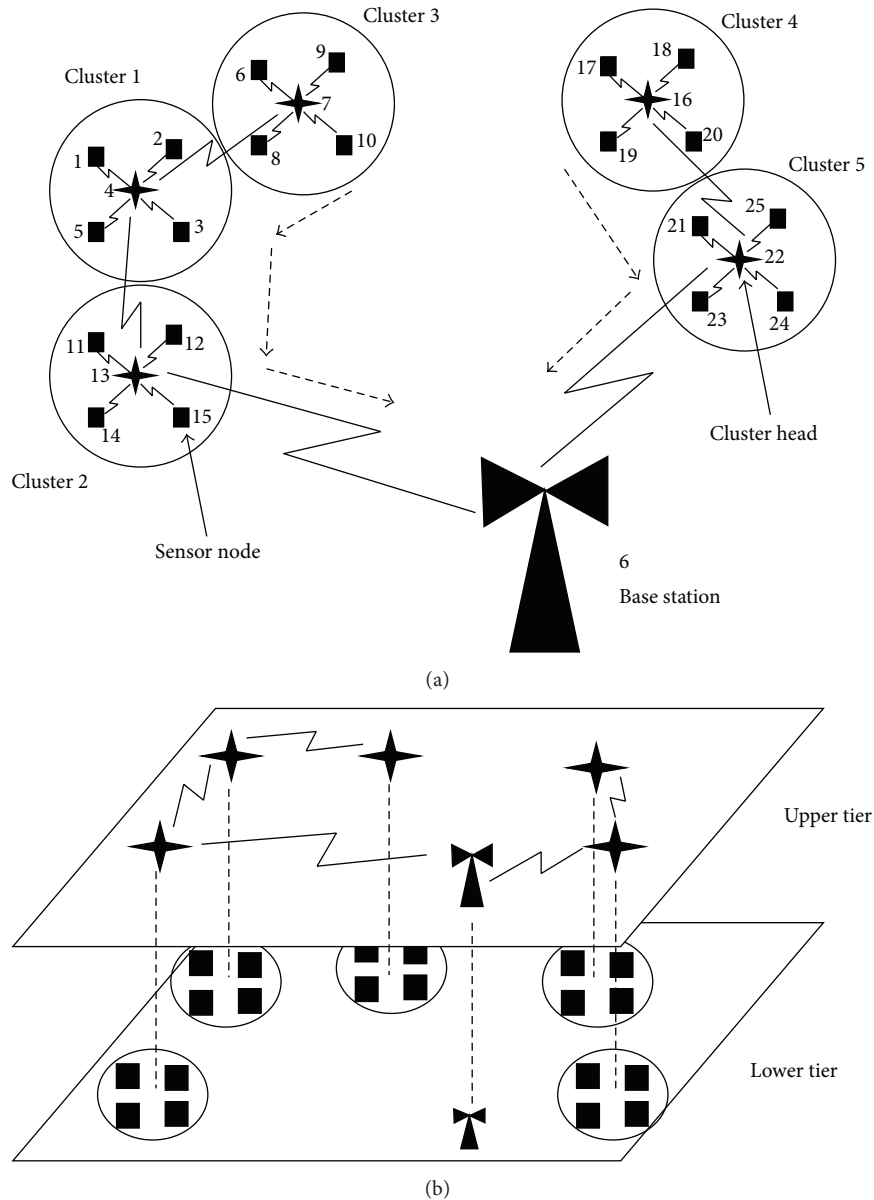
Figure 1: (a) Two Tier Sensor Network. (b) A Logical Topology of Two Tier.

its cluster members, performs data fusion, and transmits the data to the base station.

In the single-hop data transmission model [11, 12], the cluster heads send data directly to the base station. The transmission power dissipated by a sender node to transmit each bit of data to a receiver node is proportionate to the distance between the sender and the receiver [1, 12, 16–18]. Therefore, cluster heads that located farther away from the base station will get drained quickly due to the large distance for communication. For the large network, multihop data transmission model [19–22] can be used in which such farther located nodes use some intermediate nodes (cluster heads) to forward the data towards the base station. These cluster heads form a network among themselves to send data to the base station. In this case, the cluster head not only transmit data

gathered from the sensor nodes in their respective clusters but also forward data from other clusters towards the base station. This kind of network is depicted in Figure 1. Entire network with clusters and cluster head is shown in Figure 1(a). We have spotted cluster head selection and routing using these selected cluster heads as different problems. The network shown in Figure 1(a) is logically divided into two tiers, namely, lower tier and upper tier. The cluster head selection is to be solved at the lower tier. Selected cluster head forms a network amongst themselves which is shown in the upper tier of Figure 1(b). Routing using this selected cluster head is to be solved at the upper tier.

In [17, 23–25], authors have focused on the upper tier of the network to find the routing schedule. It is assumed in their approach that the relay nodes (cluster heads) contain

relatively high energy compared to normal sensor nodes, and these nodes are used throughout the network lifetime as cluster heads. But if we do not have such relatively high energized special nodes, we must think about the changeover of cluster heads with the other member nodes.

By following nonadaptive routing where energy is not taken into account, if a fixed node serves as a cluster head throughout then the battery of this unlucky node will get depleted and will die quickly. This will result in the end of useful lifetime of all nodes belonging to that cluster. Selection of cluster head periodically and strategically can avoid such circumstances. Low Energy Adaptive Clustering Hierarchy (LEACH) [11, 12], Hybrid Energy Efficient Distributed Clustering (HEED) [13], Energy Efficient Dynamic Clustering Algorithm (EEDCA) [26], Genetic Algorithm based Weighted Clustering Algorithm (GA-WCA) [27], and Location Aware 2 Dimensional Genetic Algorithm (LA2D-GA) [28] are algorithms for cluster head selection while GA based routing [23] and Minimum Hop Routing Model (MHRM) [17, 25] are centralized routing algorithms for WSN available in the literature.

In this paper, we have proposed Genetic Algorithm (GA) based approach to select a cluster head for each cluster optimally for each round of data transmission such that the overall lifetime of a sensor network is increased. This is done at the lower tier. For the upper tier, we have proposed A-Star algorithm based routing which uses these selected cluster heads to send data to the base station. Simulation result shows that our synthesized approach of GA based proposed technique to select a cluster head with A-Star algorithm based routing certainly improves the life of sensor network and outperforms LEACH, HEED, EEDCA, and GA-WCA methods of cluster head selection (at lower tier), GA [23], and MHRM routing algorithms (at upper tier).

Routing and cluster head selection are two important factors which affect life time of two tier WSN, but they have been focused and solved separately traditionally. We have proposed integration of these two approaches to extend lifetime of WSN in this paper. We assume that the cluster head selection and routing schedule generation are carried out at the base station which is broadcasted subsequently.

The rest of the paper is organized as follows. In the next section we briefly review correlated matter. In Section 3 GA based cluster head selection is discussed and Section 4 contains A-Star algorithm based routing. Section 5 discusses synthesized cluster head selection and routing algorithm followed by Section 6 which shows simulation results and analysis.

2. Review

Many protocols have been proposed for the sensor networks in the last few years. Since sensors are typically battery operated with limited energy supply, many researchers have focused on issues like energy aware routing [29]. Reducing energy consumption due to wasteful sources has been primarily addressed in the context of adaptive MAC protocols, such as EAR [30] and S-MAC [31] which periodically puts nodes to sleep to avoid idle listening and overhearing.

Distributed clustering approaches like Distributed Clustering Algorithm (DCA) [32] assumes quasi-stationary nodes with real valued weights, and the Weighted Clustering Algorithm (WCA) [33] combines several properties in one parameter called weight, that is used for clustering. Boukerche et al. [34] proposed a clustering routing protocol which uses the following: the nearest neighbor approach, alternation of nodes responsible for the intercluster communication, and alternation of possible routes to the base station. In [35], Hao et al. proposed a geographical based multihop clustering algorithm which divides the network area into small regions by adopting multihop links for the intercluster communication.

In the literature, many cluster head selection methods are proposed; LNCA [36], ACE [37], LEACH [11, 12], HEED [13], EEDCA [26], GA-WCA [27], and LA2D-GA [28] are few of them. Local Negotiated Clustering Algorithm (LNCA) [36] presents a novel clustering algorithm, which employs the similarity of nodes readings as an important criterion in cluster formation. ACE clusters the sensor network in a constant number of iterations using the node degree as the main parameter. LEACH utilizes a randomized fair rotation of cluster heads to evenly distribute the energy load among the sensor nodes in the network. HEED periodically selects cluster heads according to a hybrid of sensor node's residual energy and node proximity, a secondary parameter. EEDCA maintains a neighbor table for the cluster head which contains information about its member nodes such as ID number, location, residual energy, and so forth. If cluster head's residual energy falls below threshold energy then the process of selecting new cluster head starts. In GA-WCA, load balanced factor is considered as one of the weights along with a sum of distance from all neighbor nodes to cluster heads. LA2D-GA takes only distance as a parameter to calculate fitness function; however, representation of a chromosome is a two-dimensional grid which represents valid statistics of a WSN. Having only chromosome representation different with a traditional fitness function, it does not help much in improving the lifetime of WSN.

In the upper tier, Bari et al. [23] uses a GA based routing in which a chromosome represents the routing schedule. Every ith gene's value is a node number to whom ith node has to forward the packet. The fitness value is the number of rounds (network lifetime) which is dependent on the maximum dissipated energy of any node of a routing schedule. Each period of data gathering is recognized as a round [38]. In MHRM routing method, each relay node finds a path to the base station that minimizes the number of hops.

The problem identified is to find the best cluster heads so that using them, routing at upper tier, helps in extending the lifetime of WSN. This is appeared as an optimization and a search problem having enormous search space. In order to find out a better solution, we need some heuristic search methods; hence, we have proposed GA based and A-Star algorithm based solutions which would work at lower tier for cluster head selection and at upper tier for routing, respectively.

We have proposed a GA based solution to select cluster head to overcome these limitations. Out of many possible cluster head solutions, we search for a cluster head by which

overall communication cost and the number of weak nodes are minimized. We have discussed these parameters in detail in Section 3.1. In Figure 2, as per our GA approach, node 2 will be selected as a cluster head irrespective of the current solution and by doing so, overall nodes with enough residual energy level are also maintained.

Out of many possible routing schedules, one typical routing schedule is shown in Figure 1(a) which is represented by dotted arrows. It can be seen that cluster head of cluster 3 sends data to the base station through cluster 1 and cluster 2. In the next instance, it may be through clusters 4 and 5 or directly forwarded to the base station for the new routing schedule. Hence, cluster heads for each routing would be different to conserve energy. Considering the current routing schedule, we select cluster heads in our synthesized approach. It helps in improving the network life.

2.1. Overview of GA. GA is modeled after the natural process of survival of the fittest. For most genetic algorithms, the main concept is that the strongest individuals survive a generation and recombine with other survivors, producing an even stronger child. We have used standard GA terminology as in [39, 40]. The individuals are represented by character strings or an array of genes, often referred to as chromosomes. From the initial solution, using the principle of survival of the fittest by applying selection, cross-over, and mutation operator of GA, new generation is prepared. This process is repeated until GA converges or fixed numbers of iterations are performed [39, 41].

2.2. Overview of A-Star Algorithm. A-Star algorithm is used to find a path and to traverse the graph efficiently by using heuristics for the decision making. The A-Star algorithm [42] is a best-first search algorithm that finds the optimal path from source to the destination.

It uses a distance and a cost heuristic function (usually denoted $f(n)$ for node n) to determine the order in which the search visits nodes in the tree. The distance-plus-cost heuristic is a sum of the following two functions [43]:

(1) the path-cost function, which is the cost from the starting node to the current node (usually denoted $g()$) and

(2) an admissible "heuristic estimate" of the distance to the goal (usually denoted $h()$).

Thus for node n, $f(n) = g(n) + h(n)$ is, intuitively, the estimate of the best solution that goes through n.

2.3. Network Model. We have considered two tier sensor network model as shown in the Figure 1 with the assumption of the following properties. (1) Nodes always have data to send, and there is only one base station located far away from the network. (2) All nodes have similar capabilities of processing and communication and are energy constrained. (3) Nodes are randomly deployed having the same initial energy and left unattended after the deployment; therefore, battery recharge is not possible. (4) All nodes communicate through an ideal shared medium where communication

Chromosome	4	13	7	16	22
Index value	1	2	3	4	5

Figure 2: Chromosome representation (for cluster head selection) of WSN shown in Figure 1(a).

between nodes is handled by the proper MAC protocol as in [17, 44]. (5) We are focusing on nonflow splitting routing model which avoids many limitations of the flow splitting model [19]. (6) All the sensor nodes are homogeneous and most of the time stationary.

As shown in Figure 1, assume that there are total n nodes and m number of clusters. The value of m can be computed as given in [11, 12]. Our goal is to select such m cluster heads optimally for each round of data transmission and using these cluster heads, to search for routing schedule so that the overall life of WSN is improved.

3. GA Based Cluster Head Selection

Dividing network in several clusters helps in the distribution of energy dissipation. Ideal number of clusters in a large scale network supports proper distribution. It has been studied in [12] that 5% of total nodes as cluster heads (i.e., number of clusters) give optimal results. We have kept the number of clusters fixed and worked on assignment of cluster heads per cluster.

3.1. Initial Population and Fitness Function. The chromosome is represented as a string of node numbers. Each number at index i represents the cluster head for the cluster i. For the network diagram shown in Figure 1(a), chromosome representation is shown in Figure 2.

This indicates that for cluster 1 node 4, for cluster 2 node 13, for cluster 3 node 7, for the cluster 4 node 16, and for the cluster 5 node 22 are chosen as cluster head. From each cluster, randomly one node is selected as a cluster head. This process of creating chromosome for initial population is repeated for each chromosome of the first generation.

The fitness of a chromosome represents its credentials of the solution and helps in finding stronger or weaker nodes. Following parameters are considered to calculate fitness value in our approach.

(1) Total Distance as per current Cluster Head selection ($\text{TDCH}_{\text{dist}}$).

(2) Measure of Weak Nodes (WN_{count}).

It is represented as a pair of ($\text{TDCH}_{\text{dist}}$, WN_{count}).

$TDCH_{dist}$. Total distance as per current cluster head is considered as one of the parameters in the fitness calculation because the energy will be lost in proportion to the distance of communication. It can be calculated as follows:

$$\text{TDCH}_{\text{dist}} = \left(\sum_{i=1}^{m} \left(\sum_{j=1}^{k} d_{j,i} \right) + d_{i,\text{bs}} \right), \tag{1}$$

where m is the number of clusters and k is the total number of member nodes in a particular cluster. $d_{j,i}$ is the distance from node j to cluster head of cluster i, whereas $d_{i,\text{bs}}$ is the distance between cluster head of cluster i and the *base station*. For a better solution, this value must be as low as possible.

WN_{count}. Measure of weak nodes is another and prime parameter of fitness calculation. It signifies the health of nodes for a current set of cluster head selection which is measured with respect to the predefined threshold level of residual energy, *TLevel*. It is used to explore whether a node is weak or not and if yes then to what extent?

This can be achieved by introducing different levels of energy of node. A node having initial energy E_{init}, will also have another mark of energy, *TLevel* of energy (say 40% of E_{init}). It has been experimentally found in [45] that the value of *TLevel* can be set from 30% to 40% of E_{init} to achieve better results.

Calculation of Weak Node. WN_{count} is calculated by keeping count of the total number of nodes which is on the current selected schedule and is below predefined threshold level of energy *TLevel*. The residual energy of the node will be checked, and if it is found less than *TLevel* energy, then this count will be incremented by *Incremental_factor*. Value of *Incremental_factor* is determined as per the strategy.

Network having sensor nodes with a wide transmission range and located farther away from each other will cause more amount of energy consumption at each transmission. Energy consumption factor will be more in such case and once residual energy is below *TLevel*, only for a few more rounds that node can serve as a cluster head and soon it will get exhausted. Thus, strategy for assigning value to *Incremental_factor* can be promptly 1. For the dense network having sensor nodes located nearer to each other, energy consumption factor will be less as compared to the previous discussed case. If we still continue with adding 1 to the *Incremental_factor*, it will treat those nodes whose residual energy has just entered below *TLevel* and nodes that were already below *TLevel* and now about to reach to zero energy as the same level of weak node. To have precise consideration of energy utilized and available energy (which is required in this type of case), we have divided residual energy from 0 to *TLevel* further into several subparts. It will differentiate a node having 5% of residual energy left and a node having 25% of residual energy left, quite significantly. The intensity of a node's weakness is calculated by observing where the current residual energy of a node falls in this subpart. *Incremental_factor* value will be incremented more and more as the residual energy (of course below *TLevel*) goes nearer to the zero value.

3.2. Selection, Crossover, and Mutation. The selection of individuals is carried out using the roulette wheel, rank based selection, or tournament selection method [39]. We have applied tournament selection method with tournament size = 5 which is 10% of the population size. To produce new offspring from the selected parents, the uniform cross-over or k-point cross-over can be used for each cross-over operation.

We know that much of the power of GA comes from the fact that it contains a rich set of strings of great diversity. Mutation helps to maintain that diversity throughout the GA iterations. According to the mutation probability, one random gene of a chromosome is replaced by a better node if available, otherwise replaced by a random node.

3.3. Cluster Head Selection. While searching for the cluster head the criteria will be as follows,

(1) total energy consumption should be minimum and

(2) in a cluster head selection, by selecting any node as a cluster head, if WN_{count} is found with comparatively larger value (it means that some nodes are affected by this selection and become weak node), that node is avoided to be considered as a part of cluster head selection schedule. Instead, a new node having lesser value of WN_{count} is sought although total energy consumption would be higher than the previous selection.

The member nodes of each cluster send data to respective cluster head. Cluster head aggregates data and sends towards the base station by following routing schedule which is searched at the upper tier.

4. A-Star Algorithm Based Routing in Two Tier WSN

Given a collection of m cluster heads, numbered from 1 to m, and a base station, numbered as $m + 1$, along with their locations, the objective of the A-Star algorithm is to find a schedule for data gathering in a sensor network, such that the lifetime of the network is maximized. Each sensor node transmits fixed number of packets of data containing a fixed number of bits, in each round. Each period of data gathering is referred to as a round [38], and the lifetime is measured by the number of rounds until the first relay node runs out of power (N-of-N metric) [46].

Routing schedule is computed at the central entity (base station), and we have assumed that the average amount of data transmitted by each relay node is fixed and is known to the base station. Base station calculates optimal routing schedule and broadcasts it. Every cluster head follows this schedule. This process of finding optimal path, broadcasting it in the network, and sending data from all clusters to the base station by following this schedule is repeated in every round. Computation of routing schedule is done dynamically with the consideration of current level of energy of each cluster head. For this, normally it may require the cluster heads to report their residual energy periodically to the base station to coordinate their status. The base station can then determine the routing schedule based on this updated information.

4.1. Routing Schedule. A-Star Algorithm can be used to find an efficient path between any sources to destination. In a network if the source node is S and the destination node is D, then for every intermediate node n on the path, $g(n)$ will be the actual cost to reach node n from source node S,

To (next node)	2	6	1	5	6
From (cluster head)	1	2	3	4	5

FIGURE 3: Routing schedule represented in form of an array for Figure 1(a).

and $h(n)$ will be estimated heuristic cost from the current node n to the destination node D. We have considered actual energy so far expensed to send packet from the source node to node n, to calculate $g(n)$, while $h(n)$ is calculated as the estimated energy consumption to send packet from node n to the destination node D. This can be calculated by estimating distance between nodes. Energy consumption is calculated as per (2).

A-Star algorithm will be applied for each cluster head. The cluster head where this algorithm is applied will be the source node, and the base station will be destination node. Such m different routes will be created, and these all information is consolidated. This consolidated route information is put in an array. This is shown in Figure 3 for a network shown in Figure 1(a). Note that the base station is numbered as 6 (i.e., $m + 1$), and dotted arrow shows the flow of data from one cluster head to the other.

The array has m number of indices. Value at ith index will represent node number as to where node i will be sending data, which in turn, will go to the base station in a same way. For example, in Figure 3, node 3 will send data to node 1, then node 1 will send data to node 2, and finally node 2 will send it to the node 6 (base station).

4.2. Routing Parameters

4.2.1. Node Strength (Path Cost Count). In routing, only considering the total amount of energy consumed will not be efficient because it will drain some of the nodes which are on the efficient path. To avoid network partitioning due to this, energy usage should be balanced and distributed. As discussed earlier, concept of *TLevel* is used here as well.

While making decision for the routing, we observe the route strength by considering whether a node is a strong node or not and if yes, then to what intensity. In a route, if a node is found having below *TLevel* of residual energy, then alternate route is selected with a node having more energy than previous case. Taking this into account, overall energy consumption would be increased more than the previous case. However, this alternate route will give life extension to those nodes which were selected in the first attempt. This helps in making healthy nodes participate in routing and weak nodes getting rest, thus overall network lifetime can be extended. Calculating the value of Node Strength (path cost count) is the same as "Calculation of Weak Node" of Section 3.1.

4.2.2. Total Energy Consumed. This is another parameter for searching better routing schedule. As shown in (2), E_{cons} is the total energy consumed in transmitting sensed data by all sensor nodes to their cluster head ($E_{T(j)}$), energy dissipated by cluster head to receive such data from all its member nodes (E_R), energy used in the aggregation of data at cluster head

side (E_{Aggr}), and finally energy dissipation by cluster head to send aggregated data towards the base station ($E_{T(\text{CH})}$). Here, m is the total number of clusters and k is the number of member nodes in the particular cluster

$$E_{\text{cons}} = \sum_{i=1}^{m} \left(\left(\sum_{j=1}^{k} E_{T(j)} \right) + k \times E_R + E_{\text{Aggr}} + E_{T(\text{CH})} \right). \quad (2)$$

A-Star algorithm creates a tree structure in order to search for the optimal route from a given source to the destination. In our approach, in addition to $g()$ and $h()$, we have also taken another parameter $l()$ to measure the strength of a route which is the path cost count of weak nodes having less energy.

Thus, for a node n, estimated cost function $f()$ carries two parameters. First parameter is a summation of $g(n)$ and $h(n)$, and the other parameter is $l(n)$. $f()$ can be defined as a two argument function $f(g(n) + h(n), l(n))$.

$l(n)$ is a primary parameter, and a routing decision will be made based on its value. If $l(n)$ is same for multiple routes then only, the value of $f(n)$ is checked for further comparisons to search for the better route. Value of $l(n)$ is a path cost count which is calculated as discussed earlier.

5. Synthesized Cluster Head Selection and Routing Algorithm

We are proposing an integrated solution of cluster head selection and routing for two tier sensor network. Pseudocode 1 represents proposed solution. Line number 5 to 11 in the given pseudo code is for cluster head selection using a GA approach which is integrated with A-Star algorithm based routing [24].

Explanation of functions and variables of Pseudocode 1 is as follows.

INITIALIZE_NETWORK (). This function will initialize the network, in terms of node id, node energy, node coordinates, and so forth. This will also find the estimated distance amongst various nodes and the base station. Calculation of estimated distance from current node to the base station can be carried out from the methods mentioned in [22, 47] or location can be found using GPS system which needs to be operated for a very small period of time [25, 44]. It can also be kept fixed by placing sensor nodes at a predetermined location. This distance will be useful to calculate energy consumption between two nodes and also to get value for heuristic function, (h).

END_ASTAR (). It will check whether to terminate the process of A-Star based routing or not. This will check the residual energy of each node per cluster of the network. If any cluster is found totally drained and the cluster is not able to send data, this function will return true, otherwise it will return false.

INITIALIZE_SOL_ARRAY (). It will initialize solution array where routing schedule is to be stored. Routing schedule is an array having m number of index values. This routing schedule is discussed and shown in Figure 3.

INITIALIZE_POPULATION (). It will initialize population of chromosomes (solutions). Each chromosome will carry m

```
BEGIN
  (1) INITIALIZE_NETWORK ()
  (2) COUNT_ROUND = 0
  (3) WHILE NOT END_ASTAR()
  (4)        INITIALIZE_SOL_ARRAY ()
  (5)        INITIAL_POPULATION()              // GA STARTS
  (6)        COUNT=0
  (7)        WHILE COUNT <= ITERATIONS
  (8)             NEXT_GENERATION ()
  (9)             CALCULATE_FITNESS()
  (10)            COUNT=COUNT+1
  (11)       END WHILE                         // GA ENDS
  (12)       CH_DECLARATION()
  (13)       FOR EACH NODE i IN THE NETWORK
  (14)            CREATE_TREE(i)
  (15)            PREPARE_SOL_ARRAY(i)
  (16)       END FOR
  (17)       BROADCAST_SOLUTION ()
  (18)       INVITE_STATUS( )    // After a fixed interval
  (19)       COUNT_ROUND = COUNT_ROUND + 1
  (20) END WHILE
  (21) PRINT COUNT_ROUND
END
```

Pseudocode 1: Proposed synthesized cluster head selection and routing algorithm.

number of index values. The value at index i will contain id no of the node which is cluster head of cluster i. This chromosome representation is shown in Figure 2. This is needed for cluster head selection using GA.

NEXT_GENERATION (). It will perform all operations of GA and generate a new pool of population for the next generation. Parent selection, cross-over, and mutation operators are performed in this function. The outcome of this function is a new set of chromosomes which would be better than the previous pool.

CALCULATE_FITNESS (). It will calculate the fitness value of each chromosome in the population. Fitness is calculated as discussed in Section 3.1.

CH_DECLARATION (). It will declare selected cluster heads which have been found by GA based approach. Every node will obtain information about their cluster head for the current round. This information will be taken into account while searching for routing schedule by A-Star algorithm at the upper tier.

CREATE_TREE (i). It will expand the whole tree for a node i using A-Star algorithm, to search for the optimal path from source node i, to the destination node, the base station. The found solution is stored in the solution array (routing schedule) as shown in Figure 3. This shall fill solution array

partially. For example, for node 3, it will fill indices 3, 1, and 2 only as shown in Figure 3.

PREPARE_SOL_ARRAY(i). It will be called after a tree for a node is created to prepare the solution array which shall get partially filled. Solution array will be filled up in the same manner by every node, and eventually it will be completely filled. When all nodes are covered, this solution array is ready which we call as a routing schedule.

BROADCAST_SOLUTION (). It will be called after a routing schedule is prepared to broadcast it from the base station. All cluster heads in the network will follow this solution and accordingly send data towards the base station.

INVITE_STATUS (). This function is to ask each sensor node about their status of residual energy. This is not required to be called after every round, rather called after some fixed interval. Base station can keep track of the status of all sensor nodes because of two known parameters, namely, location and amount of data to transfer. Using this, energy consumed by each node can be computed and can be upgraded. But to improve accuracy, this function is called and statuses sent by sensor nodes are matched with the status available at the base station and updated accordingly.

COUNT_ROUND. This is a variable which keeps track of the total number of rounds that the network works for. After

every round, the energy of all nodes will be updated, new cluster heads for each cluster (by GA approach) will be searched, using these new cluster heads, and routing schedule will be searched. This will be broadcasted, and all cluster heads will follow this new schedule for transferring their data. This variable is incremented by one at every iteration. When any cluster is exhausted by energy depletion, our algorithm will stop. We are measuring lifetime of WSN in terms of rounds.

6. Experimental Results

We have performed simulation for different size of network to analyze performance of our algorithm. For our experiments, we have used first order radio model for communication energy dissipation of [12] which has been widely used by many researchers as in [23, 48–52]

$$E_{T_i}\left(b_{t_i}, d_{i,j}\right) = \alpha_2 b_{t_i} + \beta b_{t_i} d_{i,j}^m, \tag{3}$$

where $d_{i,j}$ is the Euclidian distance between node i and j, α_2 is the transmit energy coefficient, β is the amplifier coefficient, b_{t_i} is the amount of data to transmit from node i to another node, and M is the path loss exponent, $2 \leq M \leq 4$. E_{T_i} is total transmit energy dissipated.

Similarly, the receive energy, E_{R_i} is calculated as follows:

$$E_{R_i}\left(b_{r_i}\right) = \alpha_1 b_{r_i}, \tag{4}$$

where b_{r_i} is the number of bits received by node i and α_1 is the receive energy coefficient.

Hence total energy dissipated by a node i for data to receive and then to transmit it further is E_i. Consider the following:

$$E_i = E_{T_i} + E_{R_i}. \tag{5}$$

We consider both types of energy in computation of energy consumption. For simulation, the values for the constants are taken same as in [12] as follows:

(i) $\alpha_1 = \alpha_2 = 50$ nJ/bit,

(ii) $\beta = 100$ pJ/bit/m^2, and

(iii) the path loss exponent, $M = 4$.

The initial energy of each node $E_{\text{init}} = 5$ J as in [22].

We have assumed that the average data generation rate of the sensor nodes and the allocation of sensor nodes to the clusters are known. For GA, the cross-over rate is relatively kept higher than the mutation rate as it helps to evolve new offspring and takes an entire generation forward. The following is the list of parameters for GA and their values: (1) population size = 50, (2) number of iterations = 200, (3) cross-over probability = 0.8 and mutation probability = 0.2, and (4) selection method is tournament selection with tournament size = 5 and used single point cross-over.

In our simulation, we have assumed that sensor network is spread in 150×150 meters. We have experimented and observed several statistics for network sizes of 30, 50, and 80.
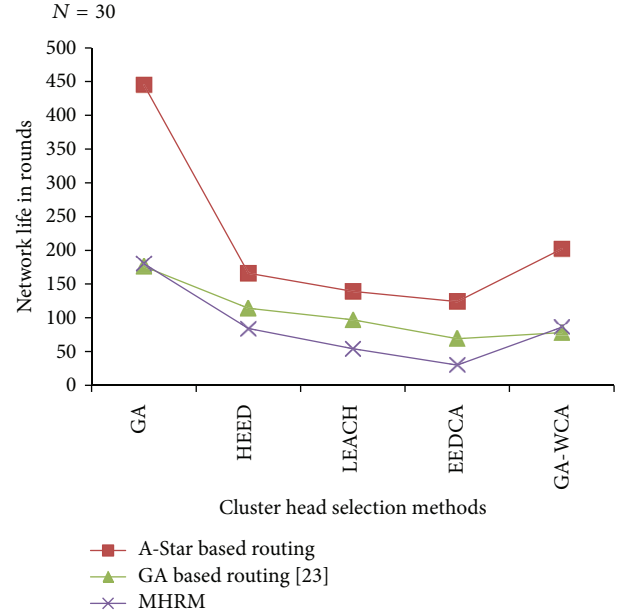

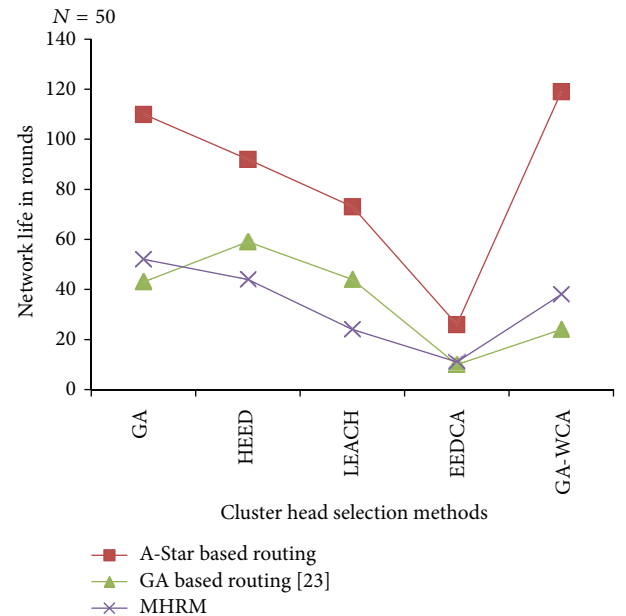FIGURE 4: Network life in rounds for different routing and cluster head selection methods for $N = 30$.

FIGURE 5: Network life in rounds for different routing and cluster head selection methods for $N = 50$.

Figures 4, 5, and 6 show network lifetime in terms of rounds for several cluster head selections at the lower tier and some routing methods at the upper tier.

For different network size, we have compared proposed GA based cluster head selection with HEED, LEACH, EEDCA, and GA-WCA cluster head selection methods for different routing methods at the upper tier. For routing, we have used proposed A-Star based routing and compared it with GA based routing of [23] and MHRM.
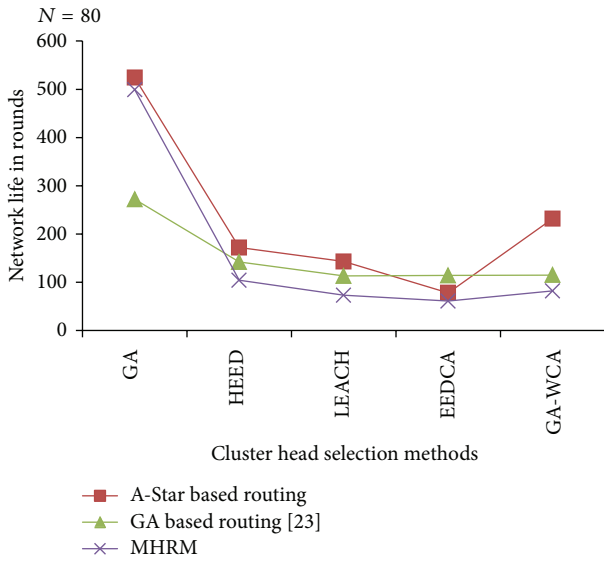
FIGURE 6: Network life in rounds for different routing and cluster head selection methods for $N = 80$.
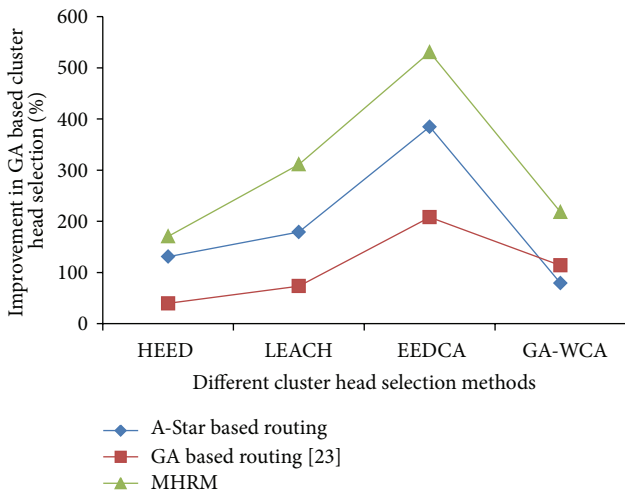

FIGURE 7: Percentage improvement in GA based cluster head selection.

We have simulated all 5 cluster head selection methods for A-Star algorithm, GA based approach of Bari et al. [23], and MHRM. It is observed that proposed A-Star based routing algorithm outperforms GA based of [23] and MHRM routing for all cluster head selection methods for different networks. It is also perceived that for different routing methods running at upper tier, proposed GA based cluster head selection at the lower tier outperforms HEED, LEACH, EEDCA, and GA-WCA. Percentage improvement in the network lifetime for GA based cluster head selection against other cluster head selection methods is shown in Figure 7. These statistics are observed for different routing methods at the upper tier of the network.

Improvements in the network lifetime using A-Star algorithm based routing as compared to other routing methods
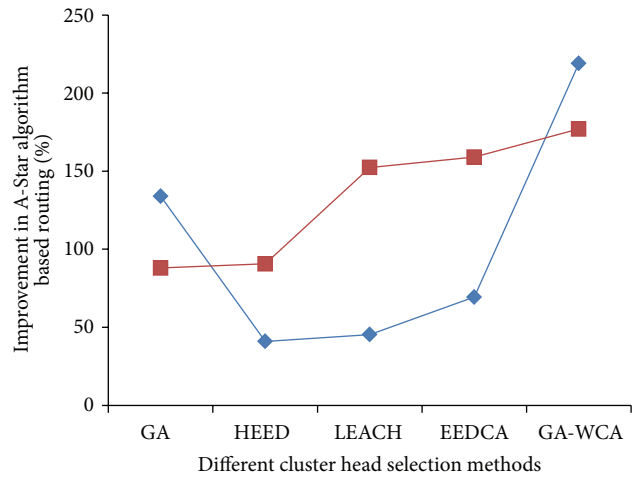

FIGURE 8: Percentage improvement in A-Star algorithm based Routing.

are shown in the Figure 8. These routing algorithms take cluster heads selected by different cluster head selection methods from the lower tier. We have taken average improvement for different network sizes and depicted in these figures. Figures 7 and 8 show significant improvement in the network lifetime using our proposed solution.

Our proposed algorithm works at the base station, and routing schedule is broadcasted by the base station. Sensor nodes follow this schedule, hence no computation is required at the sensor node regarding routing. There is an overhead of sending status of sensor nodes to the base station after a fixed interval, but it is quite low as compared to other message handshake and communication required for clustering and routing by other methods.

7. Conclusion

In this paper, we have proposed synthesized cluster head selection and routing algorithm for two tier WSN to extend life of the network. Cluster head selection and routing are two important aspects when sensed data is required to send at the base station in a large network. We have proposed GA based cluster head selection with A-Star algorithm based routing mechanism and compared it with routing algorithms like GA based [23], MHRM, and cluster head selection methods like HEED, LEACH, EEDCA, and GA-WCA. Current cluster head selection according to the current routing and routing schedule selection with the consideration of cluster head chosen helps in searching for the best combinations. Experimental results clearly exhibit that our proposed approach significantly extends the network lifetime for different sizes of the networks.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

[2] C. Kidd, O. Robert, and D. Gregdory, "The Aware Home: a living laboratory for ubiquitous computing research," in *Proceedings of the 2nd International Workshop on Cooperative Buildings (CoBuild '99)*, 1999.

[3] S. S. Intille, "Designing a home of the future," *IEEE Pervasive Computing*, vol. 1, no. 2, pp. 76–82, 2002.

[4] L. Schwiebert, S. K. S. Gupta, and J. Weinmann, "Research challenges in wireless networks of biomedical sensors," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, pp. 151–165, July 2001.

[5] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, pp. 88–97, September 2002.

[6] D. C. Steere, A. Baptista, D. McNamee, C. Pu, and J. Walpole, "Research challenges in environmental observation and forecasting systems," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MOBICOM '00)*, pp. 292–299, August 2000.

[7] M. Perillo and W. Heinzelman, "Wireless sensor network protocols," in *Fundamental Algorithms and Protocols for Wireless and Mobile Networks*, CRC Press, Boca Raton, Fla, USA, 2005.

[8] D. M. Blough and P. Santi, "Investigating upper bounds on network lifetime extension for cell-based energy conservation techniques in stationary ad hoc networks," in *Proceedings of The 8th Annual International Conference on Mobile Computing and Networking*, pp. 183–192, September 2002.

[9] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Networks*, vol. 3, no. 3, pp. 325–349, 2005.

[10] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: a survey," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 6–27, 2004.

[11] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[12] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Siences (HICSS-33, '00)*, pp. 3005–3014, Maui, Hawaii, USA, January 2000.

[13] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.

[14] Y.-Y. Liu, H. Ji, and G.-X. Yue, "Routing protocol with optimal location of aggregation point in wireless sensor networks," *Journal of China Universities of Posts and Telecommunications*, vol. 13, no. 1, pp. 1–5, 2006.

[15] L. Li, D. Shu-Song, and W. Xiang-Ming, "An energy efficient clustering routing algorithm for wireless sensor networks," *Journal of China Universities of Posts and Telecommunications*, vol. 13, no. 8, pp. 71–75, 2006.

[16] C.-Y. Chong and S. P. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.

[17] G. Gupta and M. Younis, "Load-balanced clustering of wireless sensor networks," in *Proceedings of the International Conference on Communications (ICC '03)*, pp. 1848–1852, May 2003.

[18] E. Falck, P. Floren, P. Kaski, J. Kohonen, and P. Orponen, "Balanced data gathering in energy-constrained sensor networks," in *Proceedings of the 1st International Workshop on Algorithmic Aspects of Wireless Sensor Networks (ALGOSENSORS '04)*, S. Nikoletseas and J. D. P. Rolim, Eds., vol. 3121 of *Lecture Notes in Computer Science*, pp. 59–70, Springer, Berlin, Germany, 2004.

[19] Y. T. Hou, Y. Shi, J. Pan, and S. F. Midkiff, "Lifetime-optimal data routing in wireless sensor networks without flow splitting," in *Proceedings of the Workshop on Broadband Advanced Sensor Networks*, San Jose, Calif, USA, 2004.

[20] Y. T. Hou, Y. Shi, H. D. Sherali, and S. F. Midkiff, "Onenergy provisioning andrelay node placement for wireless sensor networks," in *Proceedings of the 2nd Annual IEEE Communications Society Conference on Sensor and AdHoc Communications and Networks (SECON '05)*, vol. 32, September 2005.

[21] K. Kalpakis, K. Dasgupta, and P. Namjoshi, "Maximum lifetime data gathering and aggregation in wireless sensor networks," in *Proceedings of the IEEE International Conference on Networking*, pp. 685–696, 2002.

[22] J. Tang, B. Hao, and A. Sen, "Relay node placement in large scale wireless sensor networks," *Computer Communications*, vol. 29, no. 4, pp. 490–501, 2006.

[23] A. Bari, S. Wazed, A. Jaekel, and S. Bandyopadhyay, "A genetic algorithm based approach for energy efficient routing in two-tiered sensor networks," *Ad Hoc Networks*, vol. 7, no. 4, pp. 665–676, 2009.

[24] K. Rana and M. Zaveri, "A-Star algorithm for energy efficient routing in wireless sensor network," in *Trends in Network and Communications*, D. C. Wyld, M. Wozniak, N. Chaki, N. Meghanathan, and D. Nagamalai, Eds., vol. 197 of *Communications in Computer and Information Science*, pp. 232–241, Springer, Berlin, Germany, 2011.

[25] G. Gupta and M. Younis, "Performance evaluation of load-balanced clustering of wireless sensor networks," in *Proceedings of the 10th International Conference on Telecommunications*, vol. 2, pp. 1577–1583, 2003.

[26] M. Zhang and C. Gong, "Energy-efficient dynamic clustering algorithm in wireless sensor networks," in *Proceedings of the International Symposium on Computer Science and Computational Technology (ISCSCT '08)*, pp. 303–306, December 2008.

[27] J. Nie, D. Li, Y. Han, and S. Xie, "The optimized method of cluster-head deployment based on GA-WCA in Wireless Sensor Networks," in *Proceedings of the International Conference on Computer Application and System Modeling (ICCASM '10)*, pp. 449–452, October 2010.

[28] H.-S. Seo, S.-J. Oh, and C.-W. Lee, "Evolutionary genetic algorithm for efficient clustering of wireless sensor networks," in *Proceedings of the 6th IEEE Consumer Communications and Networking Conference (CCNC '09)*, pp. 1–5, January 2009.

[29] M. Younis, M. Youssef, and K. Arisha, "Energy-aware routing in cluster-based sensor networks," in *Proceedings of the 10th IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS '02)*, Fort Worth, Tex, USA, October 2002.

[30] K. Sohrabi, J. Gao, V. Ailawadhi, and G. J. Pottie, "Protocols for self-organization of a wireless sensor network," *IEEE Personal Communications*, vol. 7, no. 5, pp. 16–27, 2000.

[31] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the IEEE Infocom*, pp. 1567–1576, June 2002.

[32] S. Basagni, "Distributed clustering algorithm for ad-hoc networks," in *Proceedings of the International Symposium on Parallel Architectures, Algorithms, and Networks (I-SPAN '99)*, 1999.

[33] M. Chatterjee, S. K. Das, and D. Turgut, "WCA: a weighted clustering algorithm for mobile ad hoc networks," in *Cluster Computing*, pp. 193–204, 2002.

[34] A. Boukerche, A. Martirosyan, and R. Pazzi, "An inter-cluster communication based energy aware and fault tolerant protocol for wireless sensor networks," *ACM Transaction on Mobile Network Application*, vol. 13, no. 6, pp. 614–626, 2008.

[35] X. Hao, Y. Kang, and Y. Wang, "Geographical-based multihop clustering algorithm for distributed wireless sensor network," in *Proceedings of the 7th World Congress on Intelligent Control and Automation (WCICA '08)*, pp. 3230–3235, June 2008.

[36] D. Xia and N. Vlajic, "Near-optimal node clustering in wireless sensor networks for environment monitoring," in *Proceedings of the 21st International Conference on Advanced Information Networking and Applications (AINA '07)*, pp. 632–641, May 2007.

[37] H. Chan and A. Perrig, "ACE: an emergent algorithm for highly uniform cluster formation," in *Proceedings of the 1st European Workshop Sensor Networks (EWSN '04)*, January 2004.

[38] A. Bari, A. Jaekel, and S. Bandyopadhyay, "Maximizing the lifetime of two-tiered sensor networks," in *Proceedings of the IEEE International Conference on Electro Information Technology*, pp. 222–226, May 2006.

[39] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison Wesley, Reading, Mass, USA, 1989.

[40] K. Sastry, D. Goldberg, and G. Kendall, "Genetic Algorithms," in *Introductory Tutorials in Optimization and Decision Support Techniques*, E. Burke and G. Kendall, Eds., pp. 97–125, Kluwer, 2005.

[41] M.-T. Chen and S.-S. Tseng, "A genetic algorithm for multicast routing under delay constraint in WDM network with different light splitting," *Journal of Information Science and Engineering*, vol. 21, no. 1, pp. 85–108, 2005.

[42] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.

[43] S. Russell and P. Norving, *Artificial Intelligence: A Modern Approach, 2/E*, Pearson Prentice Hall.

[44] G. Gupta and M. Younis, "Fault-tolerant clustering of wireless sensor networks," in *Proceedings of IEEE WCNC*, pp. 1579–1584, IEEE Computer Society Press, Los Alamitos, Calif, USA, 2003.

[45] K. M. Rana and M. A. Zaveri, "Genetic algorithm based routing technique to extend lifetime of wireless sensor network," *International Journal of Advanced Research in Computer Science*, vol. 1, no. 2, pp. 126–132, 2010.

[46] J. Pan, Y. T. Hou, L. Cai, Y. Shi, and S. X. Shen, "Topology Control for Wireless Sensor Networks," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom '03)*, pp. 286–299, September 2003.

[47] W. C. Y. Lee, *Mobile Cellular Telecommunications*, McGraw-Hill, 1995.

[48] B. Chang and X. Zhang, "An energy-efficient routing algorithm for data gathering in wireless sensor networks," in *Proceedings of the Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC '12)*, pp. 137–141, 2012.

[49] N. Ababneh, A. Viglas, H. Labiod, and N. Boukhatem, "ECTC: energy efficient topology control algorithm for wireless sensor networks," in *Proceedings of the IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks and Workshops (WOWMOM '09)*, pp. 1–9, June 2009.

[50] H. Yang, F. Ye, and B. Sikdar, "A swarm-intelligence-based protocol for data acquisition in networks with mobile sinks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 8, pp. 931–945, 2008.

[51] W.-Y. Zhang, Z.-Z. Liang, Z.-G. Hou, and M. Tan, "A power efficient routing protocol for wireless sensor network," in *Proceedings of the IEEE International Conference on Networking, Sensing and Control (ICNSC '07)*, pp. 20–25, April 2007.

[52] S. Wazed, A. Bari, A. Jaekel, and S. Bandyopadhyay, "Genetic algorithm based approach for extending the lifetime of two-tiered sensor networks," in *Proceedings of the 2nd International Symposium on Wireless Pervasive Computing*, pp. 83–87, February 2007.