GAR: An Energy Efficient GA-Based Routing for Wireless Sensor Networks

Suneet K. Gupta, Pratyay Kuila, and Prasanta K. Jana

Department of Computer Science and Engineering
Indian School of Mines, Dhanbad-826 004, India
suneet.banda@gmail.com, {pratyay_kuila,prasantajana}@yahoo.com

Abstract. Routing with energy consideration has paid enormous attention in the field of Wireless Sensor Networks (WSNs). In Some WSNs, some high energy sensors called relay nodes are responsible to route the data towards a base station. Reducing energy consumption of these relay nodes allow us to prolong the lifetime and coverage of the WSN. In this paper, we present a Genetic algorithm based routing scheme called GAR (Genetic Algorithm-based Routing) that considers the energy consumption issues by minimizing the total distance travelled by the data in every round. Our GA based approach can quickly compute a new routing schedule based on the current network state. The scheme uses the advantage of computational efficiency of GA to quickly find out a solution to the problem. The experimental results demonstrate that the proposed algorithm is better than the existing techniques in terms of network life time, energy consumption and the total distance covered in each round.

Keywords: Wireless sensor networks, routing, genetic algorithm, network lifetime.

1 Introduction

Wireless Sensor Networks (WSNs) have been proven to be an effective technology for their wide range of applications, such as disaster warning systems, environment monitoring, health care, safety and strategic areas such as defense reconnaissance, surveillance, intruder detection etc. [1], [2]. However, the main task of WSNs in such application is to route the sensed data from randomly deployed sensor nodes to a remote Base Station (BS) called sink [3], [4]. Clustering of sensor nodes is an effective means to collect and route the sensed data using multi-hop communication usually in two-layer architecture. In the first layer, sensed data are locally aggregated by a special node called cluster head (CH) and then these aggregated data are routed to the base station through other CHs in the second layer [5], [6], [7]. By this way, the energy of the sensor nodes is conserved, the network lifetime of the WSN is prolonged and the scalability of the WSN can be increased.

However, the main constraints of the WSNs for such operation is the limited and irreplaceable power sources of the sensor nodes and in many scenarios it is difficult to replace the sensor nodes after complete depletion of their energy. Therefore, energy

C. Hota and P.K. Srimani (Eds.): ICDCIT 2013, LNCS 7753, pp. 267–277, 2013.

consumption for the sensor nodes is the most challenging issue for the long time operation of WSNs [8], [9], [10]. Moreover, in many WSNs the cluster heads are usually selected amongst the normal sensor nodes which can die quickly due to extra work load for data aggregation and data forwarding with their limited energy. In this context, many researchers [5], [11], [12], [13] have proposed the use of relay nodes, which are provisioned with extra energy. These relay nodes are treated as the cluster heads and responsible for local data collection, their aggregation and communication of the aggregated data to the sink via other relay nodes. The functionality of a two tier WSNs with the relay nodes is shown in Fig. 1.

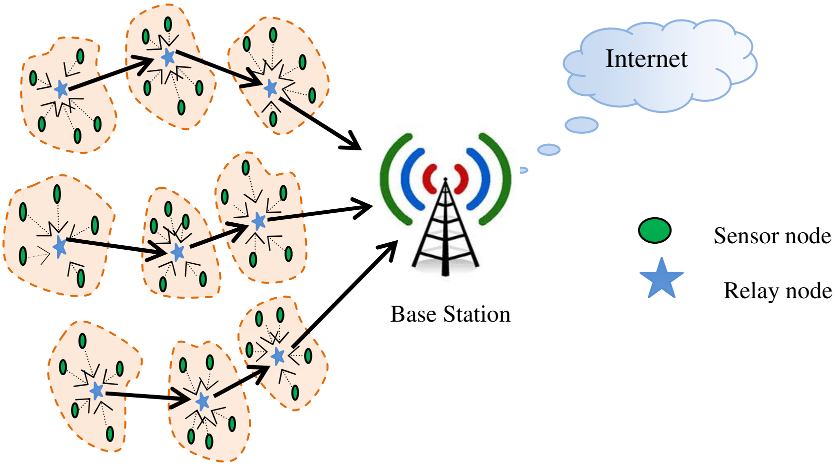


Fig. 1. An example of two-tiered wireless sensor network

Although the relay nodes are given extra energy, they are also battery operated and hence power constrained. Life time of the relay nodes is very crucial as the network life time of the WSN and the coverage fully depends on it. Therefore, energy efficient routing of the data through relay nodes is extremely important for reducing their energy consumption. It is noteworthy that for a WSN, with n relay nodes, each having an average of d valid one-hop neighbor relay nodes, the number of valid routes is d^n. This implies that the computational complexity of finding the best route for a large scale WSN is very high by a brute force approach. In order to select an energy efficient routing scheme from such a large solution space, a heuristic approach like Genetic Algorithm (GA) is extremely useful.

In this paper, we address the problem of routing the aggregated data from all the relay nodes to the sink via other relay nodes and propose a GA based routing scheme called GAR (Genetic Algorithm-based Routing) which considers the energy efficiency of the relay nodes. The GAR finds out the route from all the relay nodes to the BS by minimizing their overall distances with the view that the consumed energy to transmit the diffused data is proportional to the square of the distance between sender and receiver. Our approach can quickly compute a new routing schedule based on the current network state. We perform simulation run of the proposed method and evaluate with several metrics, namely network lifetime, energy consumption and distance covered in each round. The results are also compared with MHRM [14] as is done in [5]. The experimental results demonstrate the effectiveness of the proposed algorithm.

The paper is organized as follows. The related work is presented in Section 2. An overview of GA is given in section 3. The Network model and the terminologies are described in Section 4. The proposed algorithm and the experimental results are presented in Section 5 and 6 respectively, followed by the conclusion in Section 7.

2 Related Work

A number of routing protocols have been developed for WSN which can be found in [15], [16], [17] and their references. However, quite a few routing algorithms have been reported which are GA based. We review only those research works as they are related to our proposed algorithm. In [18], Huruiala et al. presented a GA based clustering and routing algorithm using a multi-objective function by choosing the optimal cluster-head and minimizing the transmission distance. Chakraborty [19] et al. developed a GA based protocol called GROUP in which a chain is formed to communicate with base station. The network lifetime is increased by allowing individual sensor nodes to transmit the message to the base station in non-periodical manner depending on their residual energy and location. Thus, this approach avoids forming greedy chains. In [20], L. Badia et al. presented a model by using a signal-to-interference ratio for scheduling and routing. Their results show that the model performs well for both small and large topologies. In [5], Ataul Bari et al. proposed a GA based algorithm for data routing using relay nodes in a two-tire wireless sensor networks. The main three operations namely selection, crossover and mutation used in their approach are as follows. Selection of individuals is carried out using the Roulette-wheel selection method and the fitness function is defined by network lifetime in terms of rounds. For mutation operation, they select a critical node from the relay nodes, which dissipates the maximum energy due to receiving and/or transmitting data. Mutation is done by either replacing the next-hop node of this critical node by a new next-hop relay node or by diverting some incoming flow towards that critical node to other relay node. The proposed work is also based on GA. However, it is different from [5] in respect of the following issues. 1) For selection of individuals, we use Tournament selection in contrast to Roulette-wheel selection. 2) Fitness function is defined in terms of total distance covered in a round rather than network life time in rounds. 3) In mutation, we select relay node, which uses maximum distance to transmit the data to its neighbour in contrast to a critical node defined in [5].

3 Overview of Genetic Algorithm

Genetic Algorithms (GAs) are adaptive methods that can be used in solving optimization problems. Their basic principles in the context of mathematical optimization have been rigorously described by John Holland [21]. GA begins with a set of randomly generated possible solutions, known as initial population. An individual solution is represented by a simple string or an array of genes and termed as a chromosome. The length of each chromosome in a population is same. Once an individual is generated, a fitness function is employed to evaluate its performance as a solution to the problem. This fitness function is based on how close an individual is to the optimal solution. Once the initial population of chromosomes is generated, two randomly selected chromosomes (parents) can produce two child chromosomes by a process called crossover in which the parent chromosomes exchange their genetic information. To produce a better solution, the child chromosomes undergo another process called mutation, in which the lost genetic values of the chromosomes are restored. Whenever the crossover and mutation are completed, the fitness function of the child chromosomes is evaluated and their fitness values are compared with that of all the chromosomes of the previous generation. In order to confirm that the current generation produces better result, two chromosomes of previous generation with poorest fitness values are replaced with the newly generated child chromosomes. The various steps of simple GA used in our proposed work are depicted in Fig. 2.

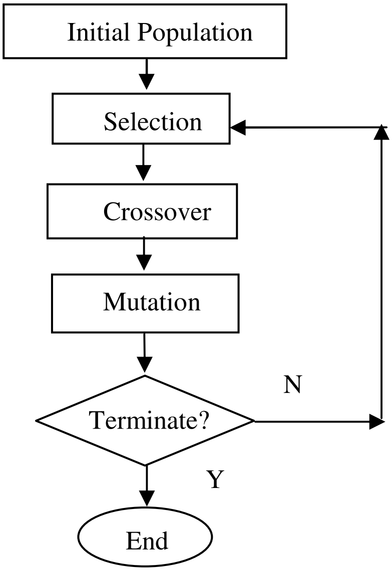


Fig. 2. Flow chart of Genetic Algorithm

4 Network Model and Terminologies

Here, we assume a WSN model, where all the sensor nodes along with some relay nodes are deployed randomly and they are stationary after deployment. The sensor nodes collect the local data and send it to their corresponding relay nodes. On receiving the data, the relay nodes aggregate them to reduce the redundant data within their cluster. Being a centralized approach, our algorithm assumes the exact location of the relay nodes which are known a priori through some location finding system such as GPS [22]. The relay nodes then route their aggregated data to the BS directly or via other relay nodes. Therefore, as per as data routing is concerned, all the relay nodes act as source or intermediate nodes treating the BS as the destination. It is obvious to note that for multi-hop communication, each relay node has to select a neighbor relay node in single-hop distance. Our algorithm deals with the selection of the best neighbour relay node to find the optimal route with respect to energy consumption. Each period of full data gathering and transmission of aggregated data from all the relay nodes to the BS is referred as a round [4]. The life time of the network is measured in terms of number of rounds until the first relay node dies. All communication is over wireless link. A wireless link is established between two nodes only if they are within communication range of each other. We use the following terminologies in the proposed algorithm.

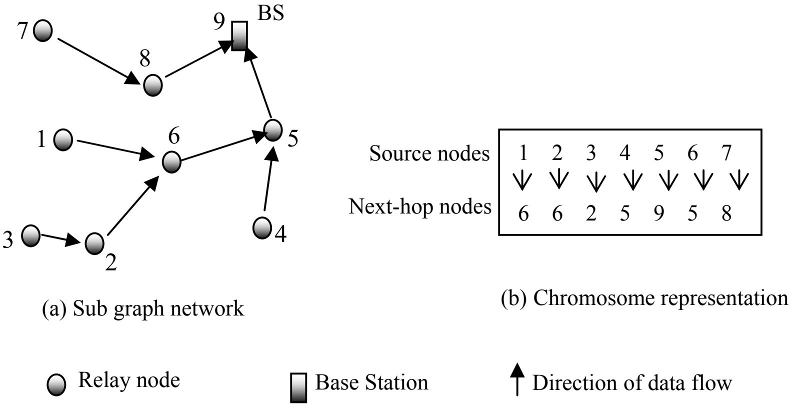
- $V = \{C_1, C_2, \ldots, C_N\}$: The set of relay nodes including BS.
- $Dist\,(C_i, C_j)$: The distance between two relay nodes C_i and C_j.
- $Com\,(C_i)$: The set of all those relay nodes, which are within communication range of C_i. The BS may be a member of $Com(C_i)$.
- $Next\_Hops(C_i)$: The set of all relay nodes which can be selected as a next hop of C_i such that the next hop relay node must be towards the BS. In other words,
  $$Next\_Hop(C_i) = \{C_j \mid C_j \in Com(C_i)\, and\; Dist(C_j, BS) < Dist(C_i, BS)\}$$

5 Proposed Algorithm

Now, we present our proposed algorithm. However, as the basic steps of the GA are repeated until the termination criterion is met, we present all basic steps, namely crossover and mutation along with the chromosome representation, creation of initial population, determining fitness function and the selection process as follows.

5.1 Chromosome Representation

We represent the chromosome as a string of relay nodes. The length of each chromosome is always equal to the number of relay nodes. A chromosome represents the route from each relay node to the BS. A routing schedule for a sub graph network along with the corresponding chromosome of eight relay nodes and one BS is shown in Fig. 3.



(a) Sub graph network

(b) Chromosome representation

Fig. 3. Sub graph network of WSN and the corresponding Chromosome Representation

This is a directed acyclic graph *G*(*V*,*E*), where *E* represents the edges $(C_i \rightarrow C_j)$, C_i is a relay node and C_j is either a relay node or BS such that C_j is closer to BS than C_i and C_j is within communication range of C_i. Fig. 3(b) is the corresponding chromosome. In this example, the value of the gene in position 1 is 6, indicating that node 1 selects the node 6 for data transmission. The value at 6 is 5 indicates that 5 is the next selected relay node and the value at 5 is 9 indicating the next relay node which is the BS (destination). Therefore the route is expressed as the path 1→6→5→9.

5.2 Initial Population, Fitness Function and Selection

The initial population is a collection of randomly generated chromosomes which correspond to a valid routing schedule. The valid chromosomes are generated in such a way that the value say *j* of the gene at position *i* is randomly selected such that $C_j \in Next\_Hop(C_i)$. As an example, the value at gene position 4, may select 6 or 5 in Fig.3 (a). However, it is shown to select 5 in Fig.3 (b). It is noteworthy that our GA based approach does not depend on any particular algorithm for generating the initial population. Moreover, it neither attempts to find a best route nor it considers the energy consumption of any relay node at the stage of initial population.

Now, we construct a fitness function to evaluate the individuals of the initial population. This helps us for the next step computation of GA, i.e., selection. We note that the total transmission distance covered by all source relay nodes to transmit the aggregated data to the BS is the main factor which we need to minimize. Therefore, our proposed work constructs the fitness function in such a way that the overall distance covered by the relay nodes is minimized in each round. The shorter the total transmission distance, the higher is the fitness value. Thus, each individual is evaluated by the following fitness function

$$Fitness(k) = \frac{1}{\sum_{i=0}^{N-1} Dist(C_i, Crom(k,i))}$$

where, *Fitness* (*k*) denotes the fitness value of the k^{th} chromosome and *Crom*(*k*,*i*) represents the value of the gene at i^{th} position in the k^{th} chromosome. It is important to note that *Crom(k,i)* actually represents the next hop neighbour relay node of C_i in the k^{th} chromosome.

For the selection process, we select some valid chromosomes with higher fitness value. In our proposed method, we use tournament selection for selecting the chromosomes with best fitness values from the initial population. All selected chromosomes are used as parents to produce new child chromosomes (offspring) with the help of crossover operation discussed in the following section.

5.3 Crossover

To produce the new offspring from the selected parents, we use 1-point crossover. The crossover operation takes place between two randomly selected chromosomes. Parts of the selected chromosomes are exchanged at some crossover point which is randomly selected. The whole process is shown in Fig.4.

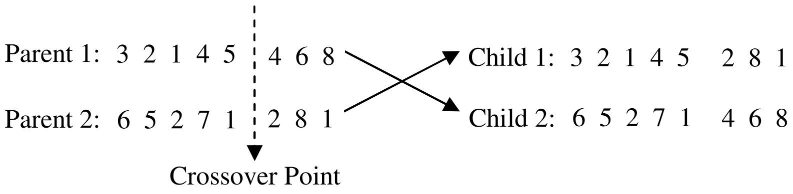


Fig. 4. Crossover Operation

***Lemma* 5.1:** The two child chromosomes produced by the above crossover operation is valid.

***Proof*:** A chromosome is treated valid if it corresponds to a proper routing schedule that indicates a path from every relay node to the BS without any cycle. As mentioned in section 5.1, the chromosomes are generated in such a way that the value *j* of the i^{th} gene is randomly selected such that $C_j \in Next\_Hop(C_i)$. Therefore, each next-hop relay node is valid. Moreover, there must be at least one relay node which can directly communicate with the BS so that each chromosome must have a route up to the BS. At the time of crossover, GAR just alters two parent chromosomes such that the value of each chromosome at a single gene position only is interchanged. As all parents are valid, therefore, new offspring must also be valid.

5.4 Mutation

We apply mutation at a selected gene rather than randomly selected gene. We select that gene which contributes maximum distance in a round and we call it critical node. We replace the critical node with other node in such a way that the new node must lead to BS covering minimum distance. The replacement of node is occurred as follows. Let i^{th} gene contributes maximum distance, i.e., the distance from C_i to the next-hop relay node is the largest. In that case, our approach searches for any other relay node C_r as next-hop node of C_i such that C_r is closer to C_i. This implies that $Dist\ (C_i, C_r)$ should be minimum. We also require that the total energy consumption via new relay node should be lesser than any other relay nodes, which implies that $\{(Dist(C_i, C_r))^2 + (Dist(C_r, BS))^2\}$ should be smallest. Combining these two, C_r is selected as the new next-hop relay node, if

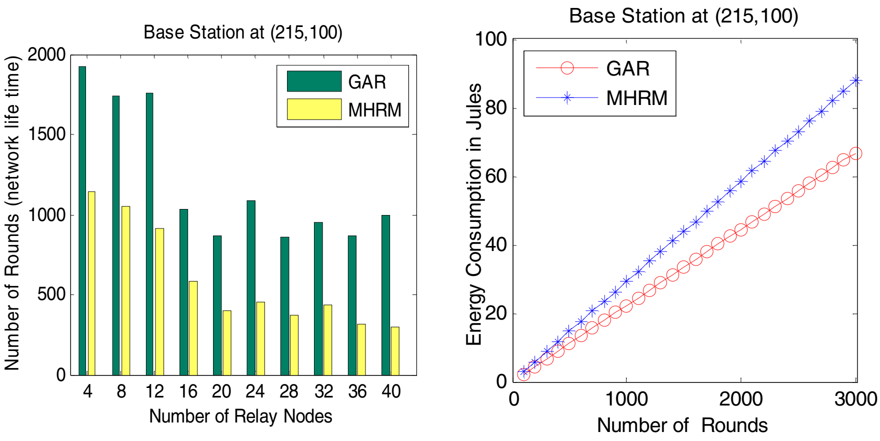
$Dist(C_i, C_r) * \{(Dist(C_i, C_r))^2 + (Dist(C_r, BS))^2\}$ is minimum.

***Lemma* 5.2:** The new chromosome after the above mutation process is valid.

***Proof*:** At the time of mutation, GAR just selects a critical node and replaces a new next hop relay node. Since all valid offspring are generated in crossover operation, we can say that mutation does not hamper the validity of these offspring by replacing a proper next-hop relay node for a critical node.

6 Experimental Results

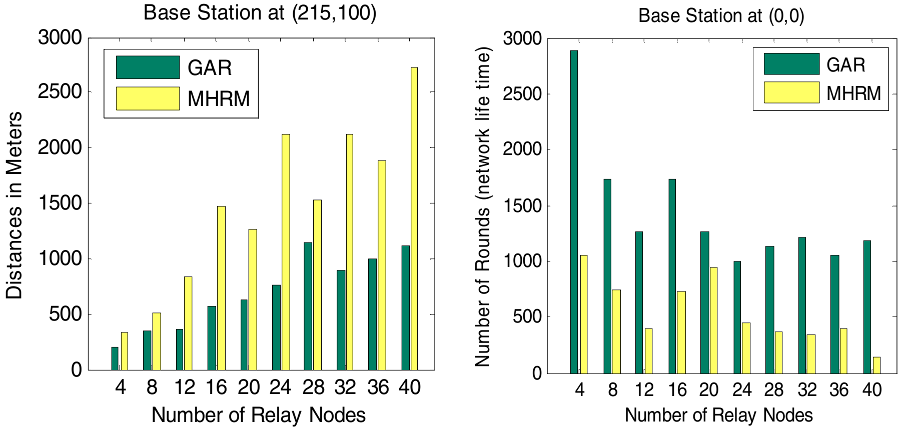
We performed extensive experiments on the proposed algorithm. The experiments were performed using MATLAB (version 7.5) on an Intel Core 2 Duo processor with T9400 chipset, 2.53 GHz CPU and 2 GB RAM running on the platform Microsoft Windows Vista. For the experiments, we considered a WSN scenario assuming that the relay nodes are deployed in 200×200 square meter area and the position of the sink is (215,100). We use the same energy model and the corresponding typical parameters as have been used in LEACH [4]. We considered an initial population of 300 chromosomes. For crossover operation, we selected the best 5% chromosomes using tournament selection procedure. The proposed algorithm was run for 150 iterations. However, it started showing high-quality results after 30 epochs. We ran it by varying the relay nodes from 4 to 40. For the comparison purpose, we also ran the MHRM [14] algorithm as also done in [5], keeping the experimental parameters same as that of the proposed algorithm. The simulated results of both the algorithms are compared in terms of in terms of number rounds (i.e., network life time) and energy consumption, as shown in Fig. 5.



(a) (b)

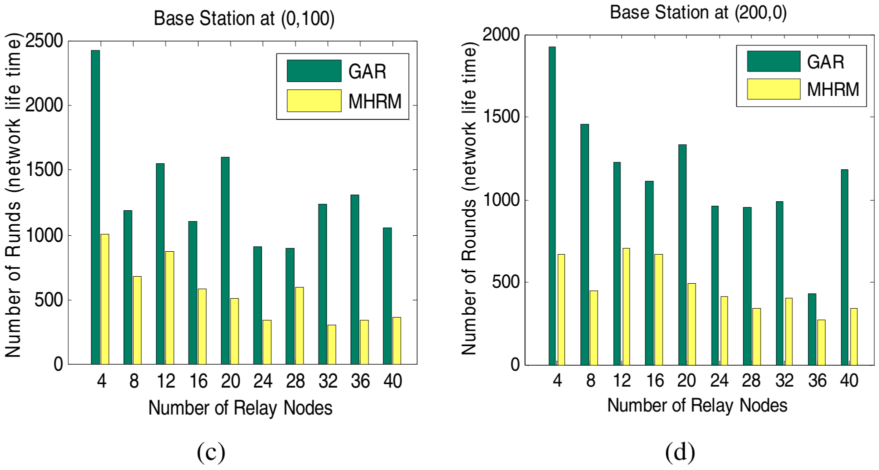
Fig. 5. Performance Comparison between GAR and MHRM in (a) Number of rounds (network life time) (b) Energy consumption in Jules

We also compare both the algorithms with respect to the total distance covered in each round. This can be observed from Fig. 6(a) that GAR covers significantly less distance than MHRM in a single round. We also ran GAR and MHRM after changing the position of the BS at several points i.e., (0, 0), (0, 100) and (200, 0) and compared results as shown in Fig. 6(b-d). It demonstrates that GAR outperforms the MHRM in network life time for all these cases.



(a) (b)

Fig. 6. (a) Total distance covered in a round and Number of rounds (i.e., network life time) (b) BS at (0, 0), (c) BS at (0, 100) and (d) BS at (200, 0)



(c) (d)

Fig. 6. (*Continued*)

It is clear from all these figures that the proposed algorithm can significantly extend the network life time in terms of rounds by approximately 230% compared to MHRM and hence is an improvement over the algorithm by the GA based algorithm by Ataul Bari et al. [5].

7 Conclusions

In this paper, we have presented a GA based routing scheme for wireless sensor networks. The algorithm has been provided with suitable chromosome representation, the fitness function, selection, crossover and mutation with their proofs of validity. We have presented the experimental results of the simulation runs of the proposed algorithm and shown that it outperforms the MHRM algorithm by extending the network life time of approximately 230 % in contrast to 200% as reported in the GA based algorithm [5]. However, the algorithm lacks the consideration of residual energy of the relay nodes for energy efficiency. Currently, we are working on this issue and trying to improve the GA based approach for a combined scheme of clustering and routing.

References

1. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless Sensor Networks: A Survey. Computer Networks 38(4), 393–422 (2002)
2. Yick, J., Mukherjee, B., Ghosal, D.: Wireless sensor network survey. Computer Networks 52, 2292–2330 (2008)
3. Chong, C.-Y., Kumar, S.P.: Sensor networks: evolution, opportunities, and challenges. Proceedings of the IEEE 91(8), 1247–1256 (2003)

4. Heinzelman, W., Chandrakasan, A., Balakrishnan, H.: Energy efficient communication protocol for wireless micro-sensor networks. In: Proceedings of the 33rd HICSS, pp. 3005–3014 (2000)
5. Bari, A., Wazed, S., Jaekel, A., Bandyopadhyay, S.: A genetic algorithm based approach for energy efficient routing in two-tiered sensor networks. Ad Hoc Networks 7(4), 665–676 (2009) ISSN: 1570-8705
6. Hamed Abbasi, A., Younis, M.: A Survey on clustering algorithms for wireless sensor networks. Computer Communications 30, 2826–2841 (2007)
7. Boyinbode, O., et al.: A Survey on Clustering Algorithms for Wireless Sensor Networks. In: 13th International Conference on Network-Based Information Systems, pp. 358–364. IEEE (2010)
8. Anastasi, G., et al.: Energy conservation in wireless sensor networks: A survey. Ad Hoc Networks 7, 537–568 (2009)
9. Kim, K.T., et al.: An Energy Efficient Routing Protocol in Wireless Sensor Networks. In: Proc. of Int. Conf. on Comp. Sc. and Engg., pp. 132–139. IEEE (2009)
10. Lattanzi, E., et al.: Energetic sustainability of routing algorithms for energy-harvesting wireless sensor networks. Computer Communication 30, 2976–2986 (2007)
11. Tang, J., Hao, B., Sen, A.: Relay node placement in large scale wireless sensor networks. Computer Communications 4(29), 490–501 (2006)
12. Kuila, P., Jana, P.K.: Improved Load Balanced Clustering Algorithm for Wireless Sensor Networks. In: Thilagam, P.S., Pais, A.R., Chandrasekaran, K., Balakrishnan, N. (eds.) ADCONS 2011. LNCS, vol. 7135, pp. 399–404. Springer, Heidelberg (2012)
13. Gupta, G., Younis, M.: Load-balanced clustering of wireless sensor networks. In: ICC 2003. IEEE International Conference, vol. 3, pp. 1848–1852 (2003)
14. Chiang, S.-S., Huang, C.-H., Chang, K.-C.: A Minimum Hop routing Protocol for Home Security Systems Using Wireless Sensor Networks. IEEE Transactions on Consumer Electronics 53(4), 1483–1489 (2007)
15. Akkaya, K., Younis, M.: A survey on routing protocols for wireless sensor networks. Ad Hoc Networks 3, 325–349 (2005)
16. Saleem, M., Di Caro, G.A., Farooq, M.: Swarm intelligence based routing protocol for wireless sensor networks: Survey and future directions. Information Sciences 181, 4597–4624 (2011)
17. Boselin Prabhu, S.R., et al.: A Survey of Adaptive Distributed Clustering Algorithms for Wireless Sensor Networks. International Journal of Computer Science & Engineering Survey (IJCSES) 2(4), 165–176 (2011)
18. Huruiala, P.C., Urzica, A., Gheorghe, L.: Hierarchical Routing Protocol based on Evolutionary Algorithms for Wireless Sensor Networks. In: Roedunet International Conference, RoEduNet (2010)
19. Chakraborty, A., Mitra, S.K., Naskar, M.K.: A Genetic Algorithm inspired Routing Protocol for Wireless Sensor Networks. Int. J. of Computational Intelligence Theory and Practice 6(1) (2011)
20. Badia, L., Botta, A., Lenzini, L.: A genetic approach to joint routing and link scheduling for wireless mesh networks. Journal Ad Hoc Networks (2009)
21. Goldberg, D.E.: Genetic Algorithms: Search Optimization and Machine Learning. Addison Wesley, Massachusetts
22. Han, A., Liu, Y.: Research on Routing Algorithm Based on the ZigBee-GPRS Technology and the Hypergraph Model. Journal of Computational Information Systems 8(9), 3895–3902 (2012)