Cluster-Based Routing Approach in Hierarchical Wireless Sensor Networks toward Energy Efficiency using Genetic Algorithm

Nguyen Thanh Huong[1], Dao Dang Toan[2]*

[1]*Hanoi University of Science and Technology – No.1, Dai Co Viet str., Hai Ba Trung dist., Hanoi, Vietnam*
[2]*Vingroup Big Data Institute - 458 Minh Khai str., Hai Ba Trung dist., Ha Noi*
Received: August 05, 2020; Accepted: November 12, 2020

Abstract

Energy efficiency is one of the important factors when exploiting Wireless Sensor Networks, especially for increasing lifespan and performance. In the network nowadays, the number of sensor nodes can reach hundreds or thousands and can be arranged in complex hierarchical architecture. Besides, the current sensor nodes have a small size, limited battery source but are operated in vast areas. The clustered-based method has been an effective and potentially extensible means of boosting the management and operation of such large-scale networks and minimizing the overall energy consumption. In this paper, the issue of arranging and routing the nodes in the sensor network in a hierarchical manner is investigated, in which each lowest level sensor nodes are grouped in a cluster with a common cluster head, then the cluster-head plays an intermediate role transmit the information back and forth between the sensor nodes and the base station. In this way, the route to exchange information can not only be optimized with respect to the distance but also for energy spent on the communication. In order to do so, this paper proposed a novel method based on a Genetic Algorithm to establish a routing protocol to achieve energy optimization. The results demonstrate that this approach can decrease the energy consumption according to the optimized routing through clustering and increase the performance superior to the other clustering schemes.

Keywords: Wireless sensor network, energy efficiency, clustering, routing, optimization, genetic algorithm

1. Introduction

Wireless sensor networks (WSNs) have been increasingly the target of research for over a few decades. It is composed of a set of sensing devices connecting wirelessly by means of radio, infrared or optical links. Nowadays a massive number of sensing devices are involved in WSNs but are desired to operate efficiently and able to survive for a long time.

Thanks to the current communication technology, wireless sensors are combined into a network to create many new possibilities for automatic interaction from humans. The sensors now are only equipped with very small microprocessors and radio devices that occupy a very small space and consume low energy. They can operate in dense environments with high-speed processing capability. Today, WSNs are used in many areas such as marine microbiology [1], monitoring of pollutants transportation [2], monitoring of complex ecosystems and environments [3], control surveillance in industry and military [4-5], security and defense applications [6] and in daily life [7-8].

A massive amount of data is generated from sensor nodes and must be transmitted to the sink and finally, the base station. This causes heavy load trafficking which results in collision or interference, even a high probability of packet errors. Moreover, DC batteries supply almost all sensor nodes in the network; therefore, the power source is very limited. For the communication within the network to function properly and last longer, the routing technique plays a key role. In the wireless sensor network, there are many power optimization techniques such as routing techniques, optimizing the energy use of each node, improving battery power, collecting wireless energy, etc. This technique is optimized by the wireless sensor network routing technology which is one of the most important factors to control the entire wireless sensor network, especially those with multiple nodes.

In the WSNs, there are many energy optimization techniques such as routing techniques, node-centric energy usage optimization, improving battery power, wireless energy scavenging, etc. Among these techniques, routing is one of the most important factors to control the entire WSNs, especially those with multiple nodes. Routing is the technique of finding the path from the source node to the destination node to ensure the optimal desired goal.

* Corresponding author: Tel.: (+84) 904.466.684
Email: huong.nguyenthanh3@hust.edu.vn

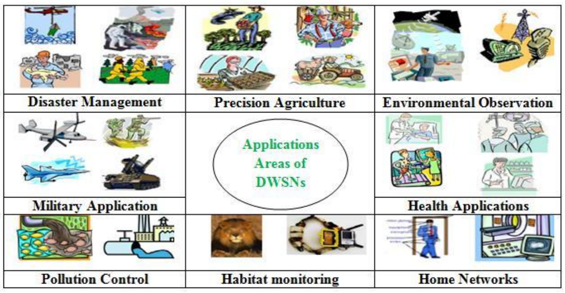


Fig. 1. WSN applications [9]

Conventional routing protocols involving centralized and distributed ones have been commonly used for WSNs [10]. However, when the network size grows, the number of routers increases, which causes a lot of difficulties in network trafficking. To overcome this problem, hierarchical routing was proposed first by [11]. Clustering or hierarchical routing protocol is a multilevel framework technique which divides the network into zones or clusters and assigns the coordinating task to a special node, Cluster Head (CH), to decrease the local computations and overhead of classic nodes. This approach is mainly aimed to reduce power consumption and increase network performance with consideration of the mobility of the nodes.

Numerous studies have been done to solve the sink location problem in WSNs with multiple sinks [12]. However, in many types of networks, sinks and sensors are randomly placed or scattered, and thus, careful sink placement is either difficult or even impossible. Algorithms are then needed to distribute the load of the sensors on random sinks to balance the burden in the WSN. Accordingly, we focus our attention on the remainder of this section on current load balancing techniques in multiple sink WSNs. In [13] a simple technique is used to allow the sensor to select the sink closest to them, as it minimizes their energy transfer. This solution results in a more energy-efficient link and the algorithm can split the sensor load roughly equally between sinks. In [14] the authors calculated the sink communication power by supplementing the communication costs of all the sensors in the cluster to balance the load in a network. The results of [15] show that the algorithm can balance a load of sensor nodes following distribution scenarios but cannot yield optimal results right away. As a result, we aim to improve the disadvantages of the proposed approaches by proposing a continuation using genetic algorithms [16] to discover a larger solution space and maintain efficiency. The result of our algorithm is to determine the best sink for each sensor that ensures sensor load balance between sinks. In addition, the results will determine the best routes that remote sensors should take to reach the sink.

In this paper, we propose a cluster-based routing approach toward energy efficiency using a genetic algorithm for hierarchical WSNs. The primary concern of this approach is to utilize formulation and operators of the genetic algorithm to determine the optimized routing based on fitness function. All sensor nodes can serve as transmitting or receiving nodes so as to make the shortest path and consume the least energy for the whole network. The paper is organized as follows. Section 2 describes the structure of hierarchical WSNs. In Section 3, the model of GA based routing protocol with the formulation is presented. Section 4 analyzes the results and comparison with the previously studied methods. Lastly, section 5 concludes the researched protocols.

2. Structure of Hierarchical WSN

The main purposes for all WSNs are to efficiently transmit and receive data within the network, extend the network lifetime, and prevent the reduction of connections by employing flexible energy management techniques. While designing routing protocols, we often encounter the following problems: network time allocation and arrangement, communication range, resource constraints, mobility issues and data transmission approach. In order to minimize errors and lengthen network lifespan, data exchange between sensor nodes and base stations can be done by using multi-hop packet transfer over short transmission range. This method significantly saves energy and reduces the transmission interference between nodes when sensor nodes compete for channel access, especially in high-density WSNs. In order to respond to queries from sinks or special events occurring in the environment, the collected data will be transmitted to base stations through different multi-hop paths.

Each sensor node consists of four basic components: sensor, microprocessor, wireless transceiver and power source. Depending on the specific application, the sensor node may also have additional components such as a location finder, power generator and mobile device. The components in a sensor node are shown in Fig. 2.

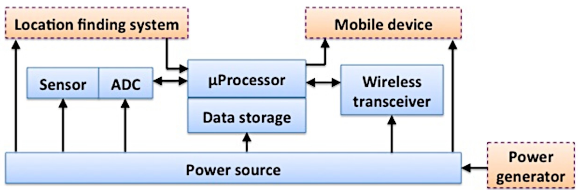


Fig. 2. Architecture of a sensor node

A WSN contains a vast number of nodes which are arranged densely inside or very close to the object to be explored and collected data. The position of the sensors does not need to be pre-determined; therefore, it allows random implementation in inaccessible or hazardous areas. The ability to organize the network itself and collaborate on the operation of wireless sensors is a basic feature of this network. With the coexistence of a variety of wireless sensors, multi-link communication is chosen; hence, the power consumption is minimal (compared to single-link communication) and provides better signal transmission efficiency (compared to long distance transmission). The basic structure of the wireless sensor network is shown in Fig. 3. The sensor nodes are placed in a sensor field. Each sensor node, which is distributed within the network, can collect data, routing the data to the receiver (Sink) for delivery to the user (User), and routing the requested messages from the Sink node to the sensor nodes. The data is routed towards the Sink in a hierarchical structure (Multihop Infrastructureless Architecture), in which there are no base stations or gateway. The receiver can communicate directly with the operating station (Task Manager Node) of the user or indirectly via Internet or Satellite….

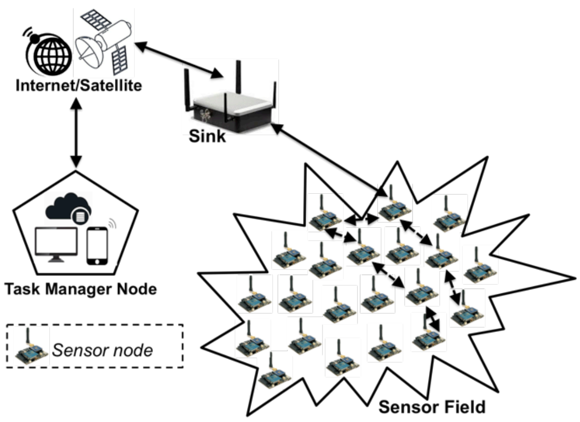


Fig. 3. Basic structure of a WSN

3. Formulation of Hierarchical WSN

For efficient energy routing, sensor nodes in the network are divided into three types: (1) Closed Nodes: nodes that can only select a single sink to connect to; (2) Open Nodes: nodes that have more than one sink choice to connect to; (3) Private Nodes: nodes that cannot connect directly to any sink in the network but through intermediary nodes to transfer data to the sink. Before describing how to implement the GA algorithm, we need to determine the energy model used in the network solution. The energy consumed by sensors can be divided into two parts, communication and computing energy. The main energy parameters involved in the communication are energy/bit consumed by the transmitter (α_t), the energy dissipated in the transmission op-amp (α_{amp}), and the energy/bit consumed by the receiving process (α_r). Assuming the loss of $1/d^2$ path, the energy consumed to send and receive bits is calculated by:

$$E_{tx} = (\alpha_t + \alpha_{amp} d^2) \times \mathrm{r} \tag{1}$$

$$E_{rx} = \alpha_r \times \mathrm{r} \tag{2}$$

where E_{tx} is the power consumed when transmitting r bits of data, E_{rx} is the energy consumed when receiving r data bits, r is the number of exchanged data bits, d is the transmission distance. According to […], the average transmitting and receiving energy density are $\alpha_t = \alpha_r = 50$ nJ/bit and $\alpha_{amp} = 100$ pJ/m^2. From equation (1), we can derive a method to calculate the amount of communication energy among nodes in the sensor network. The energy transmitted from node i to node j is calculated as follows:

- If node i has no other nodes dependent on it, the energy is calculated as:

$$C_{i,j} = E_{tx} + E_{rx} = (\alpha_t + \alpha_{amp} d^2_{i\to j}) \times \mathrm{r} + \alpha_r \times \mathrm{r} \tag{3}$$

- If node i has other n nodes that depend on it, then node i will be responsible for transmitting the data received from the other n nodes to the next hop of node i. The energy now will be:

$$\begin{aligned} C_{i,j} &= (n+1) \times (E_{tx} + E_{rx}) \\ &= (n+1) \times ((\alpha_t + \alpha_{amp} d^2_{i\to j}) \times \mathrm{r} + \alpha_r \times \mathrm{r}) \end{aligned} \tag{4}$$

Each routing solution in the network is an individual in the GA population. Every single individual is related to an optimal value resulted from the fitness function. Such optimal value corresponds to the quality of the mentioning individual, which states the position of this individual compared with others in the population. This value plays an important role in the GA process where the optimized value of the feasibly better individual can have a higher chance of being included in the next generation. The fitness function is constructed from a a sink node i based on the processing load PL_i and the communication load CE_i in the cluster. The load function of sink node i is defined as:

$$L_i = CE_i \times PL_i \tag{5}$$

where the communication load of sink node i is calculated by the total communication costs of all the sensor nodes in the cluster of sink node i.

$$CE_i = \sum_{j=1}^{n} C_{j,i} \tag{6}$$

where $C_{j,i}$ the cost of sensor node j in the cluster of sink node i. The processing load PL_i is the result of data processing procedure from sensors in the cluster. If the processing load is unbalanced, which means some clusters in the network is overloaded with sensor node data. In general, the overloaded clusters will decrease quickly, thus, part of the network will malfunction and require network rearrangement. Assume all sensor nodes are similar and process the data at the same energy level, the processing load PL_i of sink i is proportional to the n number of clusters. For every WSN, to achieve optimized handling and load balancing, each cluster should have the

$$n_{avg} = \left\lceil \frac{n_{\max}}{G} \right\rceil \tag{7}$$

where $n_{\max}$ is the number of sensor nodes and G is the number of sink nodes in the WSN. If the G number is small, the PL is small correspondingly. If n_i is supposed to be the number of sensor nodes in the cluster of sink i. If n_i is smaller than n_{avg}, the PL_i will reduce as a reward, otherwise, the PL_i will increase as a penalty. Hence, the PL_i of the sink G_i can be defined as:

$$PL_i = \frac{n_{\max}}{n_{\max} - (\max(n_i, n_{avg}))} \times \frac{\min(n_i, n_{avg})}{n_{avg}} \tag{8}$$

Although total load of each sink can be computed, the load of different clusters needs to be compared in order to determine whether the load distribution is balanced. A deviation function is used to evaluate this load difference among sinks:

$$\sigma^2 = \frac{1}{G}\sum_{i=1}^{G}(PL_i - PL_{avg})^2 \tag{9}$$

where PL_{avg} is the average processing load over all clusters. The equation (9) becomes the fitness function in which smaller deviation displays the more suitable chromosome representing better solution.

4. GA optimization algorithm for WSN

After realizing all the parameters and processes for selection, crossover and mutation operations necessary for deploying GA for WSN, we propose the optimization process applying the GA as follows:

In Fig. 4, after generating a population with P individuals, we need to calculate the fitness of each individual as in equation (9). At the first generation, after fitness assessment, the population is rearranged according to the best fitted to the worst fitted value. If this is not the final generation (the current G < maximum G), the selection, crossover and mutation are conducted. In the selection, $P/2$ best suitable individuals are paired together in order to produce $P/4$ parents and the gene difference $d_{i \to j}$ between consecutive parents. The crossover and mutation steps are implemented at this stage because they are the main manipulation in the evolution to create better next generations with better chromosomes than their parents. A parameter called Setting Difference-Degree (D_s) is used to determine whether crossover or selection should be chosen. If $d_{i \to j} \geq D_s$, the mutation should be performed; otherwise the crossover can be executed. The other $P/2$ individuals must follow the same previous steps so that no traits will be missed. After all these procedures, the whole population is rearranged, the least suitable half of which are removed to come to the next generation: the current G is $G + 1$. The final population with the lowest number is the optimal individual.

The fitness function in equation (9) is the decisive function that serves the GA execution and result. The calFitness() is drawn as the subprogram to compute the function as well as generate the final optimized result, displayed in Fig. 5.

For the fitness function, the parameters must be initialized first to decide the influenced and calculated constants and variables. It is necessary to fit each sensor node to one cluster G_i in the hierarchical WSN and compute the communication load to the next hop according to equations (3) and (4). This communication value is added to the CE_i. For each cluster, the processing load of the sink node is calculated according to equation (8) and then equation (5) is applied to determine the total load L_i of the cluster G_i. Finally, the L_{avg} is the average of all clusters and all the calculated parameters are replaced in equation (9) to find the fitness function value.

5. Results and discussion

After establishing the genetic algorithm for hierarchical WSN, the system model is computed, simulated and validated by the MATLAB program. The routing approach is realized for the following cases with different number of sensor and sink nodes.

Table 1. The experiment cases for routing protocol

Case	Number of sensor nodes	Number of sink nodes
A	30	4
B	50	5
C	100	10

These three cases represent the small, medium, and big size of hierarchical WSN structure and can be routed in Fig. 6. For around 30 to 50 sensor nodes, the number of sink nodes can be 4 or 5 to save the communication cost between the sink and the base station. However, when the number of sensors

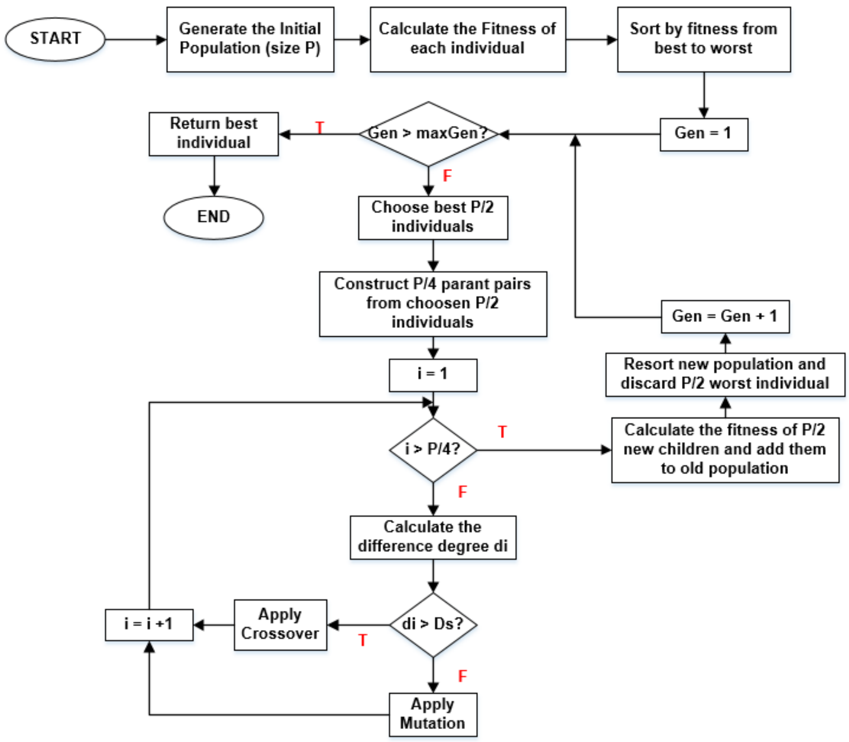


Fig. 4. Genetic Algorithm flow chart for Hierarchical Wireless Sensor Network

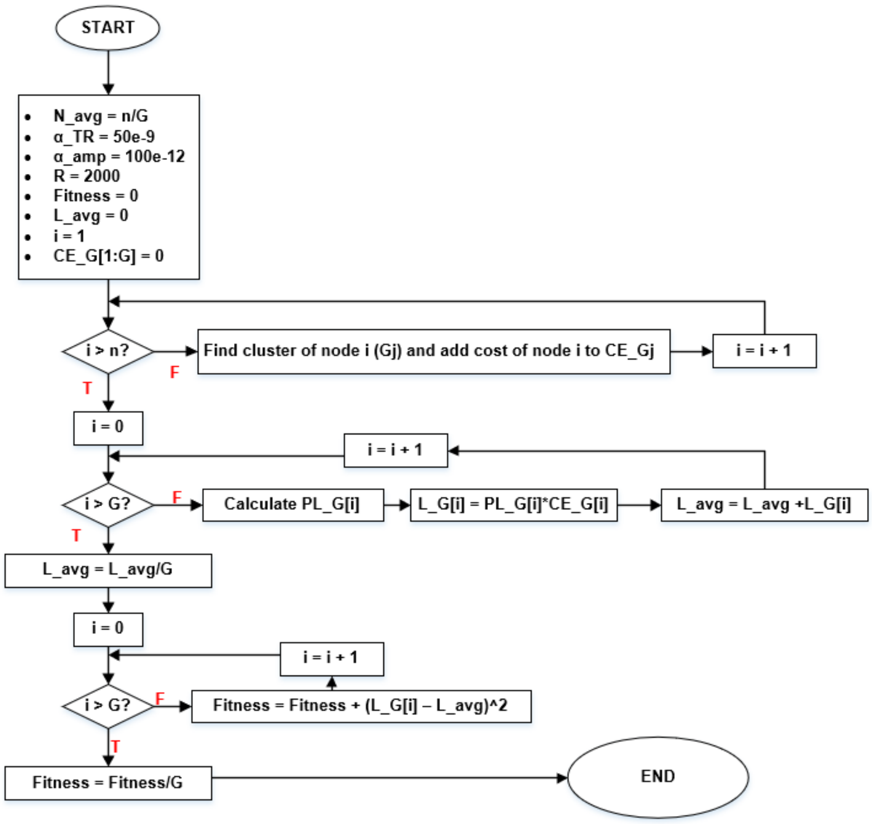
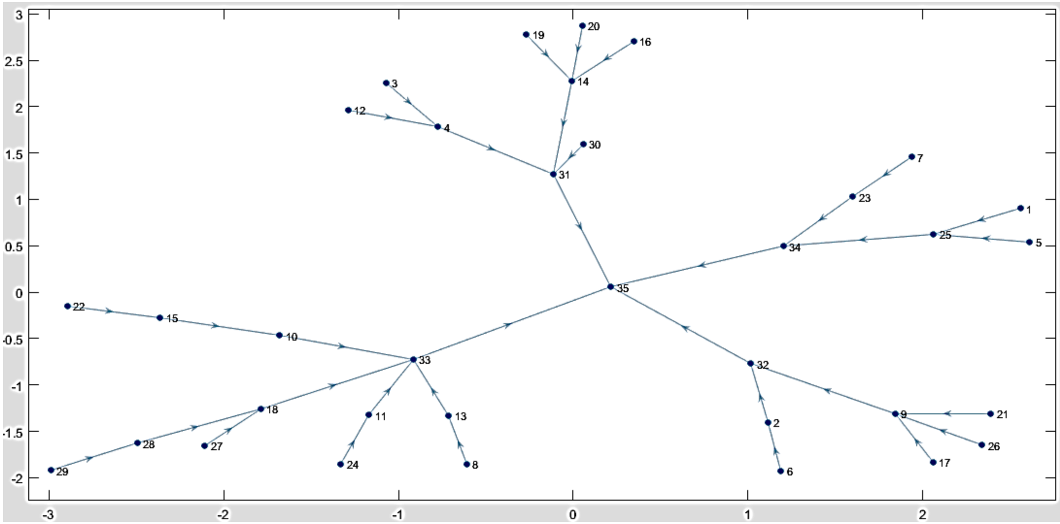
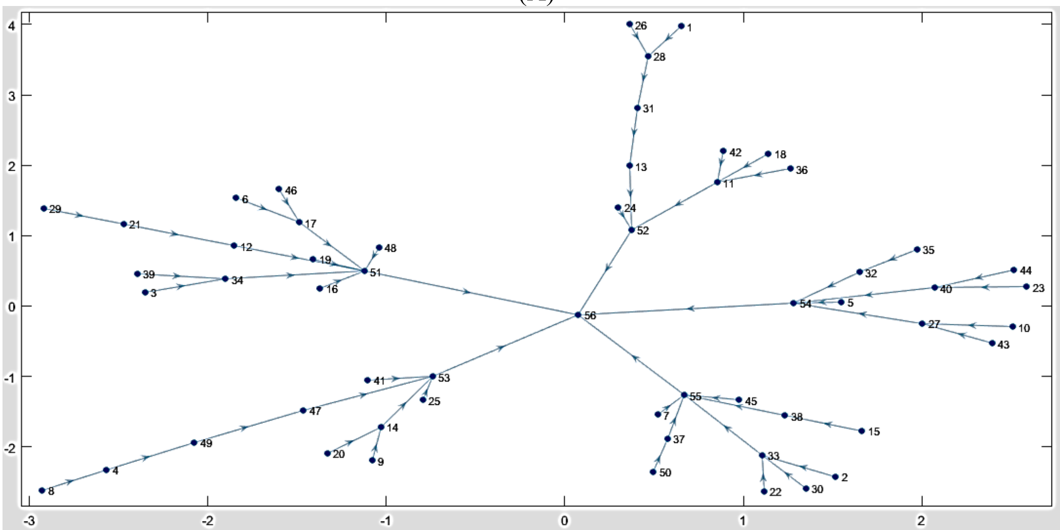


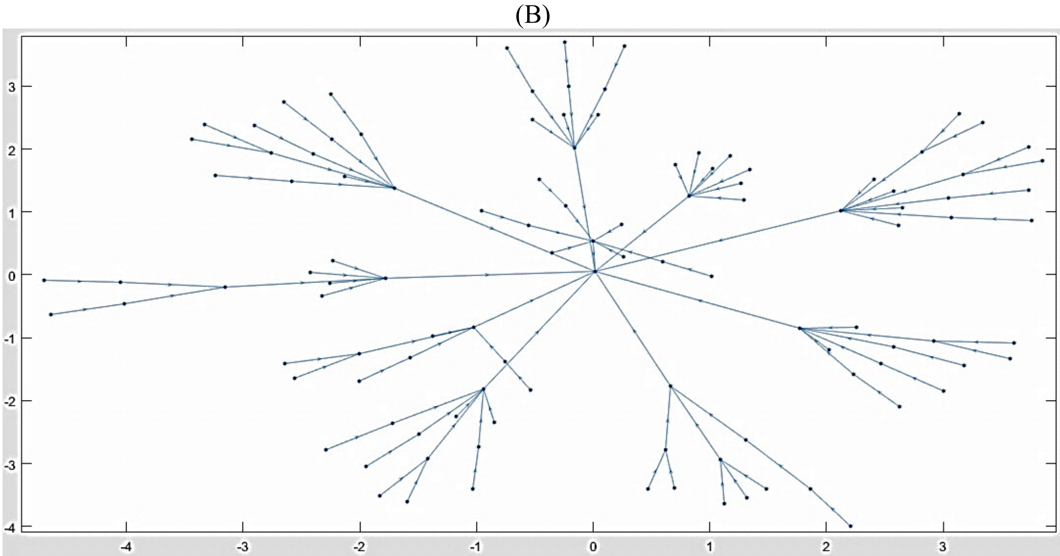
Fig. 5. Flow chart of Calculate Fitness function



(A)



(B)



(C)

Fig. 6. Routing protocol for different numbers of sensor and sink nodes

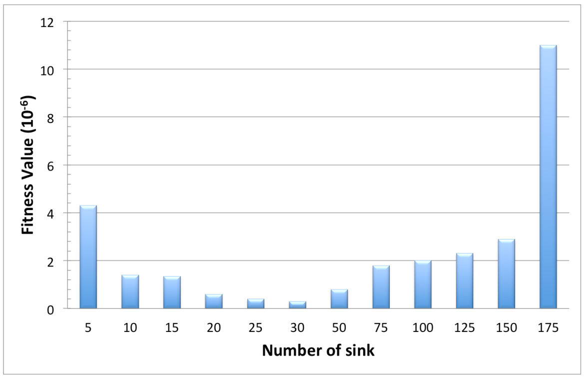


Fig. 7. Fitness function evaluation for different number of sinks

nodes increases to 100, the number of sinks must increase to save the total cost of sensor to sink nodes. The resulting routing protocol when applying GA. After routing data and acquiring the structure of the wireless sensor network, we go into assessing the dependence of fitness value on the number of sinks and the number of sensor nodes in the network.

Finally, we tested the influence of the number of sink nodes on the energy balance in WSN. The simulation results obtained with a network of 100 sensor nodes and the number of sink nodes vary from 5 to 30. The fitness value corresponding to each case is shown in the graph in Fig. 7.

From the graph we can see the more sink nodes in the wireless sensor network, the more beneficial it is in terms of energy balance in the network. Next, we will assess the influence of the number of sensor nodes on the energy balance in the wireless sensor network. Simulation results were performed with a 10-sink network and the number of sensor nodes varied from 50 to 175 nodes. The influence is shown in the graph in Fig. 7. From the graph, we can see that the less the sensor nodes in the wireless sensor network, the more beneficial it is in terms of energy balance in the network. For the low number of sinks, the fitness exhibits the decreasing tendency while for the high number of sinks, an increasing trend is obtained.

5. Conclusion

In this paper, based on an evolutionary algorithm, the GA, the routing protocol approach is proposed and exhibits the capability to adapt with the hierarchical WSN. By applying crossover, selection and mutation, the core operators of GA, the hierarchical WSN can be routed to save energy. The evaluation of the fitness demonstrates that the large network can even show better results than the small network.

Acknowledgments

This research is funded by the Hanoi University of Science and Technology (HUST) under project number T2020-PC-019.

References

[1] Xu, G., Shen, W., & Wang, X. (2014). Applications of wireless sensor networks in marine environment monitoring: A survey. Sensors, 14(9), 16932-16954.

[2] Jamil, M. S., Jamil, M. A., Mazhar, A., Ikram, A., Ahmed, A., & Munawar, U. (2015). Smart environment monitoring system by employing wireless sensor networks on vehicles for pollution free smart cities. Procedia Engineering, 107, 480-484.

[3] Chen, C. P., Chuang, C. L., & Jiang, J. A. (2013). Ecological Monitoring Using Wireless Sensor Networks-—Overview, Challenges, and Opportunities. In Advancement in Sensing Technology (pp.1-21). Springer, Berlin, Heidelberg.

[4] Aponte-Luis, J., Gómez-Galán, J. A., Gómez-Bravo, F., Sánchez-Raya, M., Alcina-Espigado, J., & Teixido-Rovira, P. M. (2018). An efficient wireless sensor network for industrial monitoring and control. Sensors, 18(1), 182.

[5] Mahamuni, C. V. (2016, December). A military surveillance system based on wireless sensor networks with extended coverage life. In 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC) (pp. 375-381).

[6] XU, J. F. (2011). A defense system for wireless sensor networks. The Journal of China Universities of Posts and Telecommunications, 18, 119-122.

[7] Abbas, A. H., Mohammed, M. M., Ahmed, G. M., Ahmed, E. A., & Seoud, R. A. A. A. A. (2014, April). Smart watering system for gardens using wireless sensor networks. In 2014 International Conference on Engineering and Technology (ICET) (pp. 1-5). IEEE.

[8] Balhwan, S., Gupta, D., & Reddy, S. R. N. (2019). Smart parking—a wireless sensor networks application using IoT. In Proceedings of 2nd International Conference on Communication, Computing and Networking (pp. 217-230). Springer, Singapore.

[9] Alharbe, N. and Abdullah, M., 2017. Routing Protocols for Dense Wireless Sensor Networks: Characteristics and Challenges. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 9(1-3), pp.7-13.

[10] Ketshabetswe, L. K., Zungeru, A. M., Mangwala, M., Chuma, J. M., & Sigweni, B. (2019). Communication protocols for wireless sensor networks: A survey and comparison. Heliyon, 5(5), e01591.

[11] Kleinrock, L., & Kamoun, F. (1977). Hierarchical routing for large networks performance evaluation and optimization. Computer Networks (1976), 1(3), 155-174.

[12] Kim, H., Seok, Y., Choi, N., Choi, Y., & Kwon, T. (2005, January). Optimal multi-sink positioning and energy-efficient routing in wireless sensor networks. In International conference on information networking (pp. 264-274). Springer, Berlin, Heidelberg.

[13] Tas, N. C., Sastry, C., & Mesrob, V. (2008, August). Noise-aware energy-efficient sensor binding. In 2008 Proceedings of 17th International Conference on Computer Communications and Networks (pp. 1-6).

[14] Gupta, G., & Younis, M. (2003, May). Load-balanced clustering of wireless sensor networks. In IEEE International Conference on Communications, 2003. ICC'03. (Vol. 3, pp. 1848-1852).

[15] Safa, H., Moussa, M., Artail, H., & Abbani, N. (2011, September). A heuristic based approach for sensor-to-sink binding in WSNs. In 2011 IEEE Symposium on Wireless Technology and Applications (ISWTA) (pp. 102-107).

[16] Srinivas, M., & Patnaik, L. M. (1994). Genetic algorithms: A survey. computer, 27(6), 17-26.