Relay Node Placement in Hierarchical Wireless Sensor Networks

Yan Wang, Liansheng Tan, and Yanjiao Zhang

Abstract—To obtain satisfactory performance in hierarchical wireless sensor networks (WSNs), it is a challenge to properly place the relay nodes in an efficient and effective way. The key focus of the current paper is the design of practical algorithm to achieve the solution of placement of relay nodes in a hierarchical WSN. We propose a relay node cover algorithm (RNCA) to generate all the possible positions for relay nodes and a genetic simulated annealing hybrid algorithm (GA-H-SA) to optimize the positions. Our goal is to find a acceptable trade-off among the three important objectives in order to maximize the lifetime and improve performance of the given WSN. The considered objectives are (1) the number of relay nodes is minimized, (2) the energy consumption is minimized, and (3) the connectivity degree between relay nodes and sensor nodes is maximized. Both theoretical analyses and numerical results demonstrate that, the proposed algorithms within limited number of iterations are promising. It is observed that they outperform other relevant state-of-the-art approaches.

Index Terms—Genetic simulated annealing hybrid algorithm, fault tolerant, relay node cover algorithm, relay node placement, wireless sensor networks.

I. INTRODUCTION

Event detection is one of the most critical applications in WSNs. For implementing the event detection task, the sensor nodes deployed in monitoring region are capable of gathering data, processing data, transmitting data to sink nodes. They are deployed in various fields such as industrial, agricultural, health care and military applications. Sensor nodes are very small, low-cost and low-power and they are usually deployed outdoors in harsh environments where we can not reachable. So extending lifetime of wireless sensor networks is a major issue. In WSNs, the main source of energy consumption is data transmission which is a super-linear function of the transmission distance [1], [2]. Thus, we design sensor nodes communicate to more powerful special nodes, relay nodes, to deliver data through multiple hops to the base station (BS) in order to short communication range.

Relay node placement is important for optimizing important network parameters and designing goals such as lifetime, cost, effectiveness, and connectivity [3]. In this paper, we establish a two tier communication network architecture, which consists of sensor nodes, relay nodes and a BS. Relay nodes aggregate data from sensor nodes forming the lower tire, meanwhile, relay nodes communicate with BS via wireless single-hop or multi-hop paths forming the upper tire. In this hierarchical architecture, we consider only the lower tire as is depicted in Fig. 1.



Fig. 1. The lower tire of a hierarchical model of WSNs.

Even though relay nodes are provisioned with higher energy and enhanced capabilities as compared to normal sensor nodes, they are also small battery operated, they have the possibility of failure (maybe due to the inhospitable surrounding or hardware malfunction). When a relay node fails, the relative sensor nodes can not sent their data to the BS. The effects of this can be fatal, such as the system is used in the forest fire prevention. Therefore, it is extremely necessary to have adequate redundancy of relay nodes, which in this paper we use connectivity to measure in order to guarantee that most of the sensor nodes will have multiple relay nodes which can communicated with [4], [5]. It is also very important that WSNs system consume less energy to prolong lifetime, and reduce the number of relay nodes in terms of management and save energy of the system. They are three goals of this paper, what we should do is to trade-off among them, to achieve the overall optimal of the system [6]. We use a multi-objective fitness function using weight sum approach and a genetic simulated annealing hybrid algorithm (GA-H-SA), which can be integrated the advantages of the two algorithms, avoiding the deficiencies of genetic algorithm in global searching and convergence speed. The experiments prove that the hybrid algorithm greatly improve the performances compared with the single algorithm.

The paper is organized as follows. We review the related works in Section 2 and propose our algorithms based on the new approaches to generate and optimize placements of relay nodes in Section 3 and Section 4 separately. Then we report the experimental results in Section 5, and conclude the paper in Section 6.

Manuscript received January 12, 2017; revised April 20, 2017. The work described in this paper was supported by National Natural Science Foundation of China (No. 61370107 and No. 61672258), and 2016 graduate student education innovation projects in Central China Normal University (No. 2016CXZZ082).

Yan Wang, Liansheng Tan and Yanjiao Zhang are with the Computer Science Department, Central China Normal University, Wuhan, China (e-mails: wangyan@mails.ccnu.edu.cn, l.tan@mail.ccnu.edu.cn, zhangyanjiao@mails.ccnu.edu.cn).

II. RELATED WORKS

In recent times, several protocols and algorithms have been developed for two-tiered sensor networks. The authors of [7] consider a two-tiered wireless sensor network. Their goals are each sensor node can communicate with at least one relay node and the network of relay nodes is connected. Ebrahim Farahmand *et al.* applied a load balanced energy-aware genetic algorithm clustering technique in [8]. First, the technique picks up the optimal cluster heads. Next, this technique assigns a

ppropriate cluster members to these cluster heads. The authors in [9] present a new routing protocol based on a lightweight genetic algorithm. In this method, the sensor nodes are aware of the data traffic rate to monitor the network congestion. In addition, the fitness function is designed from both the average and the standard deviation of the traffic rates of sensor nodes.

Misra *et al.* [10] propose an approximation algorithm for a constrained relay model for the single-tiered relay node placement problem, which is classified as NP-Hard problem. The relay nodes can only be placed in intersection lines in a grid. Lee *et al.* in [11] propose a multi-hop clustering that consider proximity to neighbor nodes and residual energy in cluster head selection process. Mafuta *et al.* [12] develop an optimal greedy relay node placement (OGRNP) featuring full coverage, communication cost in terms of received power and minimizing the number of relay nodes required for connectivity.

In this paper, we use a directed graph to model a fault tolerant hierarchical WSN. Our aim is to find the appropriate number of relay nodes which is a trade-off among the three objectives to achieve a better performance of the WSN. We proposed two algorithms, the first one is RNCA, which is used to generate the initial set of possible positions and the other one is GA-H-SA which is used to optimize the number of relay nodes generated from RNCA.

III. GENERATION STRATEGY DESIGN

In this section we should do is to generating possible positions of relay nodes according to RNCA, then establish connections between possible positions and sensor nodes using a novel method. In particular, our algorithm can make sensor nodes connect with relay nodes relative uniformly.

We firstly describe the algorithm of RNCA as follows:

We assume that the total number of relay nodes is defined as R_N , the communication range of sensor nodes is **r**. For every sensor node, we set a flag bit f_i with a initial value namely 0. The points which we find used following algorithm are called possible positions denoted by P-position.

Step 1. For every sensor node s_i , whose flag bit f_i equals 0, find one other sensor node s_{i+1} whose flag bit f_{i+1} also equals 0, measure their distance called d, if failed, return to Step 2.

If d = 2r, we can find one P-position p, satisfies the condition that $distance(s_i, p) = distance(s_{i+1}, p) = r$,

make a circle of center p and radius r, as is shown in Fig. 2.(a), check all the sensor nodes belongs to this circle, turn their flag bits from 0 to 1. Then let $R_N = R_N + 1$, return to Step 1.

If d < 2r, we can find two P-positions p_i , satisfy the condition that $distance(s_i, p) = distance(s_{i+1}, p) = r$, then make circles of center p_i and radius r, as is shownin Fig. 2.(b), check all the sensor nodes belongs to this circle separately, turn their flag bits from 0 to 1. Then let $R_N = R_N + 2$, return to Step 1.

If d > 2r, for each of the two, we take any two distinct points on the circle of center s_i or s_{i+1} and radius **r** as P-positions as is shown in Fig. 2.(c). Then make a circle of center P-positions and radius r, check all the sensor nodes belongs to this circle, turn their flag bits from 0 to 1. Then let $R_N = R_N + 4$, return to Step 1.

Step 2. For every sensor node which is not connected with relay nodes, we measure the distance d_{1i} from all the relay nodes, select one which is the minimum distance d_{1i-m} of the sensor nodes, note that, all the available distances should within the communication range of the sensor nodes, that is r, then check how many other sensors have connected with this relay node. If the number is no more than $\frac{mS_N}{R_N} (1 < m < 2)$,

we connect the link. In our simulation we set m = 1.4, which is derived from experimental method. Otherwise, discard it, repeat this step. If it fails, return to Step 3.

Step 3. We take any point on the circle of center of this senor node and radius r in order to create a new relay node, then return to Step 2.



Fig. 2. Three situations to generate P-positions.

IV. OPTIMIZATION STRATEGY DESIGN

A. Overview of GA-H-SA

Generating all the P-positions used RNCA, we should optimize them.

GA simulates the genetic and evolutionary process of organisms in the natural environment and form a kind of adaptive global optimization probability search algorithm. It is a powerful tool especially suitable for processing complex optimization problems such as NP-hard problems, combinatorial optimization, automatic control, neural networks, image processing, etc. [13], [14]. Though its global search ability is very strong, local search ability is insufficient.

SA is an extension of the local search algorithm. It is a random combination optimization method, which simulated the physical annealing in the thermodynamics process, and widely used in optimization problem. It has strong local search ability, but it is inefficient in terms of global search.



Fig. 3. Flow chart of GA-H-SA.

Aiming at the shortcomings of GA in local search ability and SA in the global search ability [15], [16], this paper combines them, put forward a new algorithm called GA-H-SA. It plays the global search ability of GA and local search ability of SA, overcome poor local search ability or premature phenomenon of GA and poor global search ability or low efficient of SA [17]. In this algorithm, GA finds generally good global solutions, SA searches local optimization of population generated by GA. The procedure of the algorithm is depicted as follows and the flow chart is shown in Fig. 3.

Step 1. Initialize: population size N, crossover probability P_c , mutation probability P_m , temperature coefficient of cooling C, initial temperature of annealing T_0 ,

the current temperature T, count of iterations of Metropolis algorithm in one temperature phase L, the current count of iterations of Metropolis algorithm in one temperature phase j, the maximum iterations of temperature cooling K.

Step 2. Calculate the fitness f_i of each individual in the population.

Step 3. At the current temperature T, repeat Step 4 while the number of iterations is less than L.

Step 4. Selection, crossover and mutation on the current population, calculate the fitness f_{i+1} of new individuals.

Let $\Delta f = f_{i+1} - f_i$, accept it with probability Δf ?1: exp $(-\Delta f / T)$.

Step 5. T = CT, k = k+1, if k < K, go to Step 3, otherwise terminate and output the optimal solution.

B. Chromosome Representation

To apply GA-H-SA, we use P-positions to represent chromosomes which are defined bitwise. After P-positions generating by RNCA, we store them used directed graph, then give them ID numbers. If gene value of a bit is one, it implies that the corresponding P-position is chosen for the placement of a relay node, otherwise it is zero. Its length is the number of initial P-positions, then use GA-H-SA to optimize them in order to output a improved chromosome. Table I depicts an example of a WSN with 10 initial P-positions. So the length of the chromosome is 10, the number of relay nodes is 6, that is $R_N = 6$.

TABLE I: CHROMOSOME REPRESENTATION

ID NO.	1	2	3	4	5	6	7	8	9	10	
Gene value	1	0	1	0	1	1	0	1	0	1	

C. Initial Population

The initial population is a set of randomly generated chromosomes, note that, we should guarantee that all the relay nodes are chosen from the P-positions which are generated from RNCA. Each chromosome represents a possible solution to the problem. There will generate infinite number of chromosomes, in this paper, we use no more than 500 chromosomes for convenient analysis, and no less than 100 for making the result reliable.

D. Fitness Function

The three goals of this paper are to minimize the number of relay nodes, minimize the energy consumption, and maximize the connectivity between relay nodes and sensor nodes to employ fault tolerance of WSNs. Firstly we put them in mathematical models respectively, then use the weight sum approach for the construction of the multi-objective fitness function.

For convenient analysis, we only consider the connectivity function and the number of relay nodes in our fitness function. We use weights w and (1-w), where 0 < w < 1 to multiply with the objectives respectively and then convert objective functions into a single objective function. In our experiment, we let w = 0.2, 0.4, 0.6, 0.8, and find their

optimal solutions respectively, then we use our energy function for every sensor node with these solutions to choose the best chromosome, that is the placement of relay nodes which we want. The three objective functions are as follows.

We use connectivity to measure redundancy of relay nodes. For every sensor node if the available relay nodes are no morethan two, use the number of available path of the sensor nodes denotes connectivity directly, otherwise reduce proportion of the redundancy in order to avoid distribution of redundancy too dense.

$$\begin{cases} N_i = n_i & n_i \le 2\\ N_i = 2 + (n_i - 2) * \alpha & n_i \ge 3 \end{cases}$$
(1)

where n_i denotes the number of available path of the sensor nodes, the factor α is tuned to be 15% in this paper, which is derived from experimental method.

The next function which denotes the number of relay nodes is given by

$$f(R_N) = R_N \tag{2}$$

where the variable RN denotes the number of relay nodes.



We use a model that is popular in literatures for the radio hardware energy dissipation where the transmitter dissipates energy to run the radio electronics and the power amplifier, and the receiver dissipates energy to run the radio electronics [1], as shown in Fig. 4. In this paper, we only consider the transmitter, as we only consider energy dissipation of sensor nodes. We used both free space fading and multiple channel fading depending upon distance between transmitter and receivers. We select d_{Thres} as a threshold distance between free space and multi path fading. If a distance between a sensor node and its relay node is lager than the threshold, multi path fading will be used. It is directly proportional to d_{li}^4 where d_{li} is the distance between relay node and sensor node, otherwise, free space energy consumption formula will be applied. It is directly proportional to d_{li}^2 . Thus, to transmit a k-bit packet, the consumption of energy of sensor nodes is as follows.

$$E_{T}(k, d_{li}) = E_{T-elec}(k) + E_{T-am}(k, d_{li})$$
(3)

$$E_T(k, d_{li}) = \begin{cases} k \times E_{elec} + k \times \varepsilon_s \times d_{li}^2 & d_{li} < d_{Thres} \\ k \times E_{elec} + k \times \varepsilon_m \times d_{li}^4 & d_{li} \ge d_{Thres} \end{cases}$$

$$d_{Thres} = \sqrt{\frac{\varepsilon_s}{\varepsilon_m}} \tag{4}$$

It should be noted that this function is applied only for a transmitter node. In this paper, we use the energy model to calculate the consumption of energy of sensor nodes in order to make a fair comparison to select the appropriate value of W.

We can make the fitness function as follows:

$$Min\{w\frac{1}{C(\mathbf{n}_{i})}+(1-w)f(R_{N})\}$$
(5)

Subject to: $\forall d_{li} \leq r, 0 < w < 1, n_i \geq 1$.

E. Selection

For the selection process, we choose chromosomes in an existing population with lower fitness value in order to produce new offspring used genetic operations described in the next paragraph. The chromosomes with lower values have higher chances of selection in this paper. There are several different selection methods, such as roulette wheel selection, rank selection, tournament selection, elitist selection and so on. We select roulette wheel selection in our simulations.

F. Genetic Operations

There are two steps of genetic operations, they are crossover operation and mutation operation.

The crossover operation is applied on two randomly selected parent chromosomes to generate two child chromosomes. The simplest and most popular method of crossover is a single-point crossover, which is shown in Fig. 5. In this method, two randomly selected parent chromosomes exchange part of their information after a random point. Note that, we have already defined a crossover rate P_c in the beginning of GA-H-SA.

P1:	10011 01011
P2:	11010 <mark>11010</mark>
C1:	10011 11010
C2:	11010 <mark>01011</mark>
Fig. 5	. Single-point crossover.

Mutation operation is applied on each chromosome to diversify population as is shown in Fig. 6. This happens according to a probability called mutation rate denote by P_m which is given in the start of the algorithm.

Original chromosome:1001111010 Mutated chromosome:1001110010 Fig. 6. Mutation.

G. Simulated Annealing

Simulated annealing is a random combinatorial

optimization method developed in the early 1980s. It simulates the thermodynamics process of high temperature metal cooling and widely used in applications of combinatorial optimization problem. In the optimization we set an initial temperature firstly, then generate an initial state and calculate the value of fitness function. Change the current state and calculate the latest fitness function f_{i+1} , set $\Delta f = f_{i+1} - f_i$, if $\Delta f < 0$, then accept it as a new one. Otherwise, accept it with probability $\exp(\frac{-\Delta f}{T})$. Repeat the above steps until the system is refrigerant.

V. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of our algorithm via simulations. There are 200 sensor nodes dispersed randomly in a $200 \times 200m^2$ square area and the BS is taken at the center of the region with coordinate (100, 100). All nodes have 0.5J energy initially. The communication range of each sensor node is 50m, we consider an initial population of 100 chromosomes. The values of the parameters used in the simulation are shown in Table II.

TABLE II: SIMULATION PARAMETERS					
Para.	Value	Para.	Value		
area	200×200	r	50m		
The BS	(100,100)	S_N	150		
Ν	100	P_m	0.4		
W	0.6	L	30		
k	1000 <i>bits</i>	T_0	40		
\mathcal{E}_{s}	$10 pJ / (bit.m^4)$	P_{C}	0.6		
\mathcal{E}_m	$0.0013 pJ/(bit.m^4)$	Κ	10		
E_{elec}	50 <i>nj / bit</i>	С	0.6		

The simulation result of the proposed RNCA is shown in Fig. 7 for 200 sensor nodes. Use this algorithm we can get the initial set of P-positions of relay nodes which are yellow circles. Fig. 8 shows the selected P-positions which are pink circles for w = 0.6 after execution of GA-H-SA.



Fig. 7. Initial set of P-positions of relay nodes.

We also compare the performance of the proposed GA-H-SA with GA and Greedy through simulations in terms of the number of relay nodes as is shown in Fig. 9. The following performance results runs of the same initializations, specifically, we set w = 0.4, which is a appropriate value obtained from the simulations that ensure most of the sensor nodes consume the least amount of energy.



Fig. 8. Selected P-positions of relay nodes.



Fig. 9 illustrates that, GA-H-SA gives better result than the other two algorithms as fewer relay nodes are used for the same number of sensor nodes. The number of relay nodes increases with the increasing of the number of sensor nodes, but when the number of sensor nodes reaches a certain number, the rate of growth decreases. The reason is that as the network becomes dense, there is no need to add additional relay nodes.

VI. CONCLUSION

In this paper, we present a relay node cover algorithm to generate P-positions of relay nodes and a genetic simulated annealing hybrid algorithm to optimize them. Our goal is to find a good trade-off among the three objectives, that is to optimize the placement of relay nodes in order to maximize lifetime and achieve better performance in hierarchical wireless sensor networks. Then we present the experimental results of the proposed algorithms for w = 0.6, and we choose a appropriate value of w through comparing the results. However, we don't consider the connectivity of the relay nodes, which will be discussed in the further study.

REFERENCES

- W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "Anapplication-specific protocol architecture for wireless microsensornetworks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660-670, 2002.
- [2] R. Liu, I. Wassell, and K. Soga, "Relay node placement for wireless sensor networks deployed in tunnels," *International Conference on Wireless and Mobile Computing, Networking and Communications* (WiMob), 2010, pp. 144-150.
- [3] X. Cheng, D. Du, L. Wang, and B. Xu, "Relay sensor placement in wireless sensor networks," *Wireless Networks*, vol. 14, no. 3, pp. 347-355, 2009.
- [4] L. Sitanayah, K. N. Brown, and C. J. Sreenan, "Multiple sink and relay placement in wireless sensor networks," *International Proceedings of the 1st Workshop Artificial Intelligence for Telecommunications and Sensor Networks (WAITS12), 20th European Conference on Artificial Intelligence (ECAI12), 2012.*
- [5] E. Lloyd and G. Xue, "Relay node placement in wireless sensor networks," *IEEE Transactions on Compututers*, vol. 56, no. 1, pp. 134-138, 2007.
- [6] G. Marco, "Mogamesh: A multi-objective algorithm for node placement in wireless mesh networks based on genetic algorithms," in *Proc. 2009 Wireless Communication Systems, 2009. ISWCS 2009. 6th International Symposium*, pp. 388-392.
- [7] J. Tang, B. Hao, and A. Sen, "Relay node placement in large scale wireless sensor networks," *Computer Communications*, vol. 29, pp. 490-501, 2006.
- [8] E. Farahmand, S. Sheikhpour, A. Mahani, and N. Taheri, "Load balanced energy-aware genetic algorithm clustering in wireless sensor networks," in *Proc. 2016 1st Swarm Intelligence and Evolutionary Computation Conf.*, 2016, pp. 119-124.
- [9] C. Park and I. Jung, "Traffic-Aware routing protocol for wireless sensor networks," in *Proc. 2010 Information Science and Applications International Conf.*, 2010, pp. 1-8.
- [10] S. Misra, S. Hong, G. Xue, and J. Tang, "Constranied relay node placement in wireless sensor networks: Formulation and approximations," *IEEE/ACM Transactions on Networking*, vol. 18, no. 2, pp. 343-447, 2010.
- [11] S. H. Lee, S. Lee, H. Song, and H. S. Lee, "Gradual cluster head election for high network connectivity in large-scale sensor networks," in *Proc. 2011 Advanced Communication Technology (ICACT) International Conf.*, 2011, pp. 168-172.
- [12] A. D. Mafuta1 and T. Walingo, "Spatial relay node placement in wireless sensor networks," in *Proc. 2016 IEEE 83rd Vehicular Technology Conf. (VTC Spring)*, 2016, pp. 1–5.
- [13] J. R. Liu and C. Leach, "Genetic algorithm-based energy-efficient adaptive clustering protocol for wireless sensor networks," *Journal of Sensor and Actuator Networks, Int. J. Mach. Learn. Comput.*, vol. 1, 2011.
- [14] M. Azharuddin and P. K. Jana, "A GA-based approach for fault tolerant relay node placement in wirless sensor networks," in *Proc.* 2015 Computer, Communication, Control and Information Technology Third International Conf., 2015, pp. 1-6.

- [15] G. Rudolph, "Convergence analysis of canonical genetic algorithms," *IEEE Transactions on Neural Networks*, vol. 5, pp. 102-11, 1994.
- [16] Q. Xiaofeng and F. Palmieri, "Theoretical analysis of envolutionary algorithms with an infinite population size in continuous space. Part I: Basic properties of selection and mutation," *IEEE/ACM Transactions* on Networking, vol. 5, pp. 102-119, 1994.
- [17] M. Yoshikawa and H. Yamauchi, "Hybrid architecture of genetic algorithm and simulated annealing," *Engineering Letters*, vol. 16, no. 3, 2008.



Yan Wang was born in Henan in 1992. She received the B.S degree in computer science and technology from Henan Normal University, Henan, China, in 2015. She is a M.E student at the Department of Computer Science in Central China Normal University now. Her current research interests include wireless sensor networks and congestion control in computer communication networks.



Liansheng Tan is now a professor at the Department of Computer Science in Central China Normal University. He received his Ph.D degree from Loughborough University in the UK in 1999. He was a research fellow in Research School of Information Science and Engineering, the Australian National University, Australia, from 2006 to 2009. He was a postdoctoral research fellow in the School of Information Technology and Engineering at University

of Ottawa, Canada in 2001. He also held a number of visiting research positions at Loughborough University, University of Tsukuba, City University of Hong Kong, and University of Melbourne. Liansheng Tan is currently an editor of international journal of computer networks and communications. He was an associate editor of dynamics of continuous, discrete & impulsive systems and an editor of international journal of communications systems from 2003 to 2008. Tan has published more than 100 referred papers widely in international journals and conferences. His research interests include graph theory, performance evaluation and architecture in telecommunications networks, wireless networks and sensor networks and wireless communication and mobile computing.



Yanjiao Zhang was born in Henan 1992. She received the B.S degree in computer science and technology from Henan Normal University, Henan, China, in 2015. She is now a M.E student at the Department of Computer Science in Central China Normal University. Her current research interests include wireless networks and sensor networks and wireless communication and mobile computing.