

An energy-aware routing protocol for wireless sensor network based on genetic algorithm

Lingping Kong¹ · Jeng-Shyang Pan^{1,2} · Václav Snášel³ ·
Pei-Wei Tsai⁴ · Tien-Wen Sung²

Published online: 22 June 2017
© Springer Science+Business Media, LLC 2017

Abstract Energy saving and effective utilization are an essential issue for wireless sensor network. Most previous cluster based routing protocols only care the relationship of cluster heads and sensor nodes but ignore the huge difference costs between them. In this paper, we present a routing protocol based on genetic algorithm for a middle layer oriented network in which the network consists of several stations that are responsible for receiving data and forwarding the data to the sink. The amount of stations should be not too many and not too few. Both cases will cause either too much construction cost or extra transmission energy consumption. We implement five methods to compare the performance and test the stability of our presented methods. Experimental results demonstrate that our proposed scheme reduces the amount of stations by 36.8 and 20% compared with FF and HL in 100-node network. Furthermore, three methods are introduced to improve our proposed scheme for effective cope with the expansion of network scale problem.

Keywords Wireless sensor network · Genetic algorithm · Energy-aware routing protocol

✉ Jeng-Shyang Pan
jengshyangpan@fjut.edu.cn

¹ Innovative Information Industry Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China

² Fujian Provincial Key Lab of Big Data Mining and Applications, Fujian University of Technology, Fuzhou, China

³ Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava, Ostrava, Czech Republic

⁴ Department of Computer Science and Software Engineering, Swinburne University of Technology, Melbourne, Australia

Mathematics Subject Classification 00-01 · 99-00

1 Introduction

A wireless sensor network (WSN) comprises hundreds or even thousands of sensor nodes that are equipped with low-cost and low power devices. With the low property constrains, a single sensor node can only sense and communicate within a limited range [1–4]. Therefore, in order to detect and gather those important information from environment, the information must be gathered after the shared efforts of many nodes. During this period of gathering time, much energy would be consumed by sensor nodes while sensing, receiving and sending data [5]. In most of network architecture, those battery-powered sensor nodes are hard to replace or recharge. this case leads to the paralysis of the network once there are nodes run out of their power. So power saving and effective utilization are vital issue for wireless sensor network research [6]. The (WSN) gradual gets involved into our daily life in activities and services. over the recent few years, with the potential use of sensor networks in civil and military applications, the design of the network has attracted more and more importance, such as combat field management, disaster surveillance [7,8]. The (WSN) is also used in water pollution monitoring area, where the (WSN) is scattered in several several countries or remote environment to detect the dangerous gases and bacteria. Not just in the above applications, (WSNs) also may use on healthcare, which assist those elder and disabled or people with diseases living alone in the house. Moreover, (WSNs) also may apply to agriculture, which will benefit farmers in charging the industry and free them from endless maintenance. Current researches [9,10] in sensor routing problem mostly concerned about protocols that are energy-aware and toler-

ant to sensor loss or energy exhaustion, thus to prolong the lifetime of sensor networks [11–13]. However the requirement to the consideration of quality of service is urgent for the development of wireless sensor network, and the energy limitation makes the difficulties prominent [14]. Consider the following scenario: In a surveillance environment, some sensor nodes are deployed to detect the useful information and track the change of the data. After collecting the data, the data would be forwarded to controlling center by sensor nodes. The controlling center only take the proper actions if the data is under a real-time exchange between sensors. However the data transferring process needs the progressive transmission among the sensor nodes, outer sensor nodes connect with inner nodes then report events to controller center. Those inner nodes always deplete their energy earlier than outer nodes because of the heavy traffic relayed by them. What's worse, once the inner node is dead due to the insufficient power, then those outer nodes went out of work due to the break of connection. Finally the surveillance environment system will be damaged even a lot of sensor nodes alive. In those traditional two-tier networks, the imbalanced relaying mission that is assigned automatically based on the distance to the sink is the vital reason affecting the network running. In order to find a solution to the imbalanced workload [2, 15–17]. In this paper, we build a middle oriented network layer. In this network, the middle layer is composed of several stations, which are in work as a intermediary by gathering the data from sensors and forward them to sink node. The goal of our design is try to lower the amount of the relay stations while guarantee the full coverage.

The rest of paper is organized as below: In Sect. 2, the survey of routing protocols for (WSNs) and clustering techniques are presented. In Sect. 3, our design of middle oriented network layer and routing protocols will be introduced in detail. In Sect. 4, some improved strategies for the extension of network scale problem are proposed. The next section shows the simulation of several schemes for building routing protocol and their energy cost comparison results. Finally, we concludes our paper in the last section.

2 Related works

During the past years, wireless sensor network has applied into many applications such as military, healthcare, surveillance and agriculture as a powerful technology. Many researchers have done a lot of studies for the use of WSNs and take the best advantage of it. Most of the research [18, 19] proposals on WSNs have been developed to keep the high efficiency of energy during the process of data transmission from the source to destination. Of course, the network establishment costs will be different with varied network

application. EADC (energy-aware clustering algorithm) is an even sizes cluster constructing strategy by using competition range characteristic and in the area those nodes are distributed non-uniform in network. This algorithm was proposed by Yu et al. [2]. Yu et al. also puts forward a cluster-based routing algorithm which gives more forwarding tasks to those available nodes. The strategy is to force cluster heads to pick nodes with more power or with fewer neighbors as next stops. (EAUCF) (energy-aware unequal clustering algorithm) was presented by Bagci [5] for solving hot spots problem which is also the bottleneck problem. it happens because those cluster-heads with closer distance to base station are easily die quickly than the others due to the heavy workload. So prolong the lifetime of those nodes who are either close to sink or have low remaining power are the key aim for (EAUCF) algorithm. The strategy is to decrease the workload of them. Artificial bee colony algorithm is often used by authors. One of the utilization is to apply it into energy efficient clustering mechanism by Karaboga [20]. By increasing the number of sensing signals the algorithm can take a full advantage of benefit from WSN. Thus extends the lifetime of network. An algorithm Geographic and Energy-Aware Routing (GEAR) was proposed by Yu et al. [21], the author used energy aware mechanism to plan a route for packet from the source to the target region and design Recursive Geographic Forwarding and Restricted Flooding algorithm to propagate the data within the range of destination region. Most previous cluster based routing protocol algorithms only cares the relationship of cluster heads and sensing nodes but ignore the huge difference costs for a cluster head than a common sensor node. In this paper, In this paper, we present an routing protocol that is energy aware one based on genetic algorithm for a middle oriented network layer. The middle layer consists of several stations, which are in charge of gathering the data from sensing nodes and re-forward it to sink node directly. The amount of stations should be not too many that will cause too much costs. However the amount should not be too few that will spend extra transmission consumption of sensor nodes. The following parts introduce several location selection strategies [22–25]. Firstly the farthest first traversal (FF) [26] was proposed by Gonzalez. The first point is to choose arbitrarily and each successive point is as far as possible from all previously chosen points. This is the brief version of farthest first traversal. Another related strategy is a variant of FF called Harel method that is used by Harel and Koren [27]. His method tries to find M Pivots for high-dimensional graph embedding task. The process of Harel Method starts with a random point, and then each following point is as far as possible from the last chosen point instead of the previously chosen points. There are two kinds of *Harel* methods HL and HD. They are different with the distance computing equation. HL uses the Euclidean distance and HD uses Dijkstra distance. A popular way to solve loca-

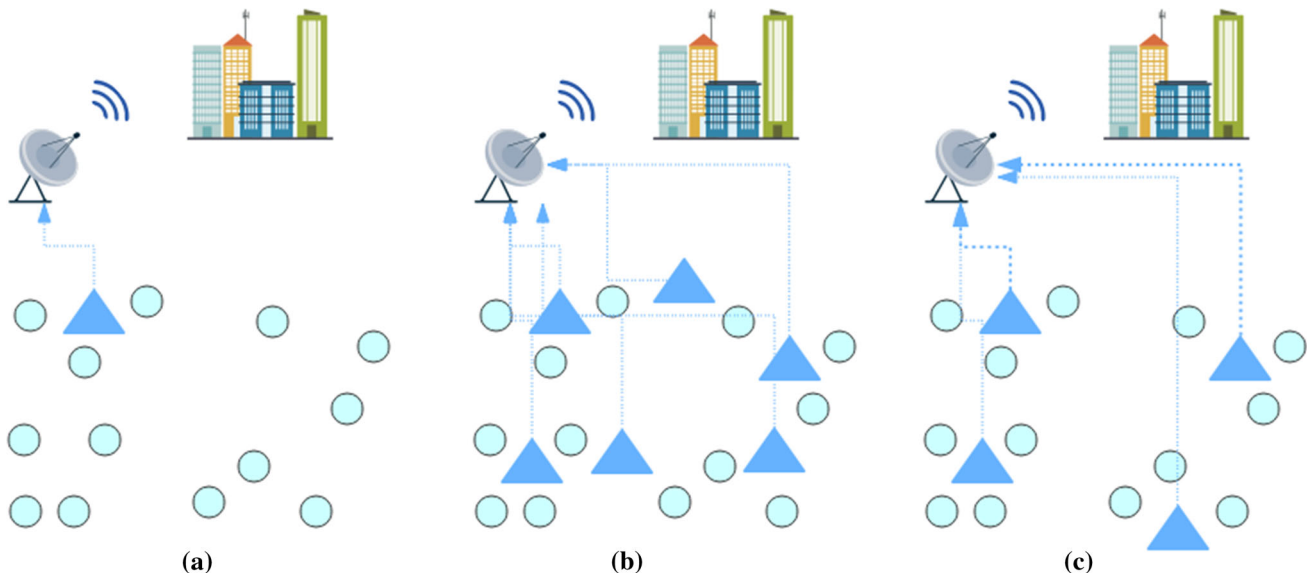


Fig. 1 A middle oriented layer network

tion selection problem is dominating set method, which is to find a minimum set individuals that could cover the rest of members [28]. The paper presents a heuristic algorithm to get a dominating set by using *CovCnt* value (the neighbor amount that also means the vertex could be contacted by the remaining neighbors) and *Score* value (estimate the possibility of becoming a center). The dominating set starts from empty then grows during the algorithm based on lazy principle, which is to add new point into set as late as possible.

In this paper, the pattern of the network is similar with 2-tier network with the sink in the center area. It connects and communicates to all the relay nodes directly. Those common sensors only sense the interesting events from environment and reporting them to relay nodes. We make the following assumptions regarding the network model:

- We consider a periodic event collecting happens where data is sensed and m data bits are transmitted in one slot.
- There is no constrain for the distribution of sensor nodes, they all randomly deployed in the area. The locations of relay nodes are chosen from the original place of sensors. The total amount of sensor nodes and relay nodes are constant.
- When a periodic event happens, the closest sensor near the source event will response and restrain the act of other nodes. If two sensor nodes synchronous sense the event, both the data will be transmitted to relay nodes.
- A relay node will forward the collected data to the sink every four slots.
- The sink node has no constrain of using power, thus we will ignore the energy consumption of it.

In this paper, we use a similar network model with paper [29], the energy model includes sensing, receiving and sending three parts of consumption. This paper will follow it to compute the energy of sense, transmit and receive data.

$$\begin{aligned}
 E_{sense} &= \alpha m, \quad \alpha = 60 \times 10^{-9} \text{ J/bit} \\
 E_{rx} &= (\beta_1 + \beta_2 r^2)m, \quad \beta_1 = 45 \times 10^{-9} \text{ J/bit}, \\
 &\quad \beta_2 = 10 \times 10^{-12} \text{ J/bit/m}^2 \\
 E_{Rx} &= \gamma m, \quad \gamma = 135 \times 10^{-9} \text{ J/bit}.
 \end{aligned}
 \tag{1}$$

3 The proposed routing protocol implement

The theory of evolution inspires the application of computational models family. Genetic algorithm (GA) is one of them. GA are usually used to solve optimization and search problems for getting high-quality solutions with bio-inspired operations. A simply implementation of GA evolution process begins with generating several initialization solutions [30–33]. And store the best individual by comparing their fitness values or other parameters. Then put those individuals into crossover or mutation mode for further evolving. After a lot of iterations, output the best individual, which is the final solution. In this section, a three-tier network routing protocol construction method is presented in detail. Our design has two phases, *Selecting candidate middle layer phase* and *genetic algorithm phase*.

Selecting candidate middle layer phase In our protocol, we build a middle oriented network layer. we illustrate the working style of this middle layer by using an example. Assume

Table 1 Pseudo code for middle layer candidate selecting

```

For it from 1 to M
  Set G, set S= {}
  While G! = null
     $\forall z \in G, S = \{z\}$ 
    Remove m where  $m \in G_{\square m \leftrightarrow z}$ 
    Remove z from G,
  Output S
End While
End For
    
```

the Fig. 1 is the model of network. The sink receiver is arranged on the edge of this field. It is responsible for collecting the data from the middle layer and then transmitting those data to outside internet. Those nodes in triangle are stations which work for receiving important data from the lower layer and forwarding the data to sink node. There are several stations in the network, they all communicate with sink node directly. The lowest layer deployed with a lot of common sensor nodes which are in charge of sensing the interesting packages from environment and transmit these events to middle oriented layer. It is a simple three-tier network. However, the number of stations plays an important role in balancing the energy consumption and operating in an effective and stable way for the network [34]. The large number of stations as in Fig. 1b will cause too much construction cost. On the contrary, for a small quantity of stations network as in Fig. 1a, there will be some isolate nodes that could not connect with a station. The perfect state in Fig. 1c for this kind of network is to minimize the amount of stations and guarantee a full coverage surveillance environment.

The next step is to form a middle layer by selecting several appropriate stations. And each position of station is chosen from the sites of sensor nodes located in. We will select *M* sets of solution. Each of the solutions may construct a middle layer and ensure a full network coverage. The process of picking stations is shown as the pseudo codes in Table 1. In the simulation phase, stations and common sensor nodes are equipped with the same sensing and communicating device. The following symbols express as: Set *G* contains the whole sensor nodes in the network. And initially the set *S* is empty. The result of *S* will be a set of candidate stations. The first step starts by randomly choosing a node *z* from *G* and adding it into *S*. Every different initial node *z* generates a different set of solutions. In the while body, node *m* is the one that belongs to the set *G* and it can sensed by node *z* directly.

Table 2 The configuration of a chromosome

0	1	2	3	4	5	...	n-1
0	0	1	1	1	0	...	0

Genetic algorithm phase Since each set of solutions from above meets the full coverage request. In this phase, we integrate the genetic algorithm to obtain the most appropriate set of stations with the best fitness value.

- *Initialization* Table 2 is one of the examples for showing the structure of one chromosome. Assume each sensor node with a serial number from 0 to (*n* - 1), here, label *n* is the quantity of sensor nodes in our field. The binary value '0' and '1' represent that the corresponding node is a common node or a station. The configuration of one chromosome is as follows.

Step 1 Pick one item from *M* sets of solutions that are generated from *Phase one* as MLS.

Step 2 Create a chromosome with all bits value '0'.

Step 3 Set those bits value '1' that those bits represent node coming from MLS.

The *M* set of solutions will evolve *M* chromosomes.

- *Evaluation* Evaluate the population with *M* chromosomes by computing their fitness value using Eq. 1. Store the best fitness value as *g_{best}*.
- *Fitness function* During the whole process, each chromosome represents a solution of middle oriented network layer model. In order to evaluate the quality of each solution, a proper fitness function *f_m* is designed here which is shown in Eq. 2.

In Eq. 2, *S* is the set of all nodes in area, and the size of *S* is *n*. the set *C* collects the bits with values 1 in a chromosome, *d(C, S_{sink})* count all the distance from sink node to each relay node, and *d(x_i, S_{sink})* computes the distance between node to sink. The coefficient *α, β* are distance independent parameters (Table 3).

$$f_m = \alpha \cdot |S - C| - \beta \cdot \left(\frac{\sum_{1 \leq i \leq C} d(C_i, S_{sink})}{\sum_{1 \leq i \leq n} d(x_i, S_{sink})} \right) \tag{2}$$

- *Crossover* Before crossover process, separate the population into two parts. Not all of the population operates this crossover phase. Only those chromosomes that with better than a half population in fitness value run this operation. After the crossover operation, the new child will replace the one that is worse than a half population in fitness values. We only use one-point fixed operators in crossover process. The crossover point is set to the middle position of the chromosome. For example, Let *M(mother)* and *F(father)* be the parent Chromosome, *M_{i,1}, ..., M_{i,n}* and

Table 3 Crossover operation

0	1	2	k	k+1	5	n-1	
0	0	1	...	1	0	...	0
0	1	2	k	k+1	5	n-1	
0	1	0	...	0	1	...	1
0	1	2	k	k+1	5	n-1	
0	0	1	...	0	1	...	1

Table 4 Crossover operation

0	1	k	k+i	n-1			
0	1	...	0	...	1	...	1
0	1	k	k+i	n-1			
0	1	...	1	...	0	...	1

$F_{j,1}, \dots, F_{j,n}$ respectively. A middle point is chosen and then the child C string is as follows.

$$C = M_{i,1}, \dots, M_{i,\lceil n/2 \rceil}, F_{j,\lceil n/2 \rceil+1}, F_{j,\lceil n/2 \rceil+2} \dots, F_{j,n} \tag{3}$$

- **Mutation** This evolution process works by mutate a bit value with a defined probability of a chromosome. In our paper, we use a constant mutation rate of 0.02, as the bits k and $(k+1)$ are shown in Table 4. Back [31] suggests a value $1/l$ rate for mutation as the lower rate limitation, where l is the length of a chromosome. This operation provides a attempt in a lower probability for random search resulting in helping us to avoid the loss of valuable station candidates information and widen the search space.
- **Remedy** As mentioned above, every chromosome represents a solution of our middle oriented network layer model, whose validity or infeasibility is very important. That model satisfies that all the common sensor nodes with value ‘0’ must connect to any one station with value ‘1’. If not, the chromosome either should be re-initialized or remedied. The remedy strategy includes reverse those value ‘0’ bit to value ‘1’ or pick some of them to reverse the value.

An overview to sketch middle oriented network layer construction process:

- (1) Select M set of candidate middle layer solution as MLSs.
- (2) Create the population based on MLSs.
- (3) Evaluate each chromosome using fitness function, store the one with the best fitness value as g_{best} . Set $t := 0$.
- (4) Execute the crossover and mutation operation. Remedy each damaged.
- (5) Repeat steps 3 and 4. Until $t = iteration$. Output g_{best} .

4 Improved strategy

Some methods may be good in the final results, but it may cost too much computation time and memory space. On the contrary, an algorithm may spend only several steps resulting with an acceptable solution. As the genetic algorithm, it is like a kind of exhaustion testing and it represents all the possible combined solutions. If we want to get an expectation solution, it just needs to set a proper iteration times for genetic algorithm. The above two phases based on genetic algorithm gets a good scheme to build the middle oriented layer, but there is a flaw for apply it for expanding the network size. An increasing number of sensor nodes will raise a higher consumption of running time and memory space. The main issue about high consumption on time and space comes from the configuration of the chromosome in genetic algorithm phase. The length of one initial chromosome becomes larger as the growth of the number of sensor nodes in network. Hence cutting down the binary string length is the effective way to solve this problem. The follows are some strategies to improve the proposed strategy.

Dividing zone In the above *genetic algorithm phase*, all the sensor nodes are taken into consideration for optimizing the most suitable station location. Actually, we may stop taking the whole area of network as a unit and divide the field into several small pieces. For each piece of land, operate the *selecting candidate of the middle layer process* and *genetic algorithm process*. In the end, combine each part solution together as the final resolution. It is different from the farthest first traversal results, there is no station located at the boundary of field. Therefore, the combination outcome may make a good performance.

In the simulation of experiment, we divide the network into four sectors. And this improved strategy is labeled as *SDg* (sector divide ga).

Sampling nodes Sampling [35,36] is concerned with the selection of a subset of individuals from a population to estimate characteristics of the whole population. usually, the process of sampling comprises the following: One, choose elements as candidates. Two, specify a sampling method, third, define the sample size and sampling plan. Here we use two frequently used sampling methods to do

the improvement: simple random sampling (RSg) and systematic sampling (SSg).

- (1) *Systematic sampling* Systematic sampling (also known as interval sampling) starts selecting the elements randomly and then proceeds with the picking of every k_{th} element from the set. $k = (\text{population size/sample size})$ in this case. If sampling size is the half size of the whole nodes, then k is 2. Firstly sample the sensor nodes into two sets, set A and set B . Operate the *Selecting candidate middle layer phase* and *genetic algorithm phase* for set A , get the temporary result $t1$. Remove those nodes in set B covered by the nodes from $t1$. Then do the *Selecting candidate middle layer phase* and *genetic algorithm phase* again for the rest of nodes in set B , then get temporary result $t2$. The final station result is the combination of $t1$ and $t2$.
- (2) *Simple random sampling* Simple random sampling is to select without replacement a random amount of n records out of a set of N . In experiment, we pick a random sample of n records out of a pool of N sensor nodes, every node in network could be chosen in an equal probability. In our simulation, the size of sampling nodes is half-size of the total sensor nodes or one third of total nodes, which depends on the number of total sensor nodes. After sampling the nodes, the rest of process follows the SSg.

5 Simulations

The simulation results include the two parts, the amount number of relay stations and the energy cost. We implemented farthest first traversal (FF), dominating set (DO), Harel line (HL), harel Dijkstra (HD) and our proposed scheme, we use the five methods run the experiments in the same condition. In our simulation model, 100–600 sensor nodes are scattered uniformly and randomly in 200×200 units of 2D space. The communication radius R for a sensor node is different along with the different size of network. 40, 30, 25, 20, 15 and 10 units are set for 100-, 200-, 300-, 400-, 500- and 600-node network respectively. The location of sink node is the center of the our area.

In Table 5, it shows the results of five methods about the amount of stations in four different size of network, 100/40 describes the amount of sensor nodes is 100 and their communication radius is 40 units in the simulation environment. From the table, it shows that for 100-node with 40 units communication radius network, it only uses 12 stations to build a middle oriented layer, where those stations can cover all sensor nodes in the network, and 18 stations are needed for dominating set method, Farthest first traversal algorithm need more stations than that. Our proposed scheme reduces the

Table 5 The results of stations number

NN/CR	Methods				
	Our proposed scheme	FF	DO	HL	HD
100/40	12	19	18	15	17
200/30	21	29	29	25	29
300/25	33	44	40	36	43
400/20	51	67	64	55	64

Table 6 The results of stations number

NN/CR	Methods						
	SDg	RSg	SSg	FF	DO	HL	HD
400/20	52	48	53	67	64	55	64
500/15	88	86	87	110	106	91	106
600/10	175	172	176	209	198	186	189

Table 7 The result of stability values for our proposed scheme

Node number	StD	Avg	Max	Min
100	0.25	12.9375	13	12
200	0.793200	23.3125	24	21
300	0.543905	33.8125	35	33
400	0.619139	52.625	53	51

Table 8 The result of stability values for improved strategies

Node number	Methods	StD	Avg	Max	Min
400	SDg	1.1672617	54.8125	56	52
	SSg	0.683130	54.25	55	53
	RSg	1.062622	50.9375	52	48
500	SDg	1.064581	90.25	91	88
	SSg	0.946484	89.6875	91	87
	RSg	1.087811	87.875	89	86
600	SDg	0.910585	177.1875	178	175
	SSg	1.5	179.375	181	176
	RSg	1.327591	174.8125	176	172

amount of stations by 36.8 and 20% compared with FF and HL in 100-node network.

The results of 400- to 600-node networks are shown in Table 6. Seven methods include three improved strategies are shown in the table. In the simulation, we divide the area into four sectors in *SDg* method for 400-, 500- and 600-node network. The random sampling size is one half or one third of the total number for *RSg*. In *SSg*, the interval of the sampling is set two for 400 and 500, and three for 600-node network. From the table, we see the three improved strategies work

Table 9 Data sensing effective ratio

Node number	Our proposed (%)	FF (%)	DO (%)	HL (%)	HD (%)
100	95.62	89.22	87.13	90.77	95.02
200	54.98	48.27	50.39	54.92	52.87
300	69.72	63.00	65.54	69.63	67.91
400	65.95	62.46	61.96	65.26	63.48
500	63.72	57.84	60.60	61.13	59.22
600	58.55	55.40	58.06	56.94	58.20

Table 10 The result of total energy consumption

NNN	Methods	Total E	NN	Methods	Total E
100	Ours	1870981.739	400	Ours	1.0759E7
	FF	2044118.65		FF	1.14437E7
	DO	2068256.169		DO	1.149413E7
	HL	2019272.7		HL	1.09127832E7
	HD	1934544.28		HD	1.12540559E7
200	Ours	6502476.049	500	Ours	1.37097E7
	FF	7475331.86		FF	1.50898E7
	DO	7068458.789		DO	1.43577E7
	HL	6597517.77		HL	1.43196867E7
	HD	6807352.430		HD	1.479805E7
300	Ours	7595938.23	600	Ours	1.78788E7
	FF	8420092.98		FF	1.895575E7
	DO	7983957.849		DO	1.8045359E7
	HL	7568729.20		HL	1.84030683E7
	HD	7850107.18		HD	1.799684E7

with good performance. *RSg* strategy reduces the amount of stations by 25% compared with than *DO* and *HD* two method in 400-node network. *SDg* spends 20% less stations than *FF* method in 500-node network.

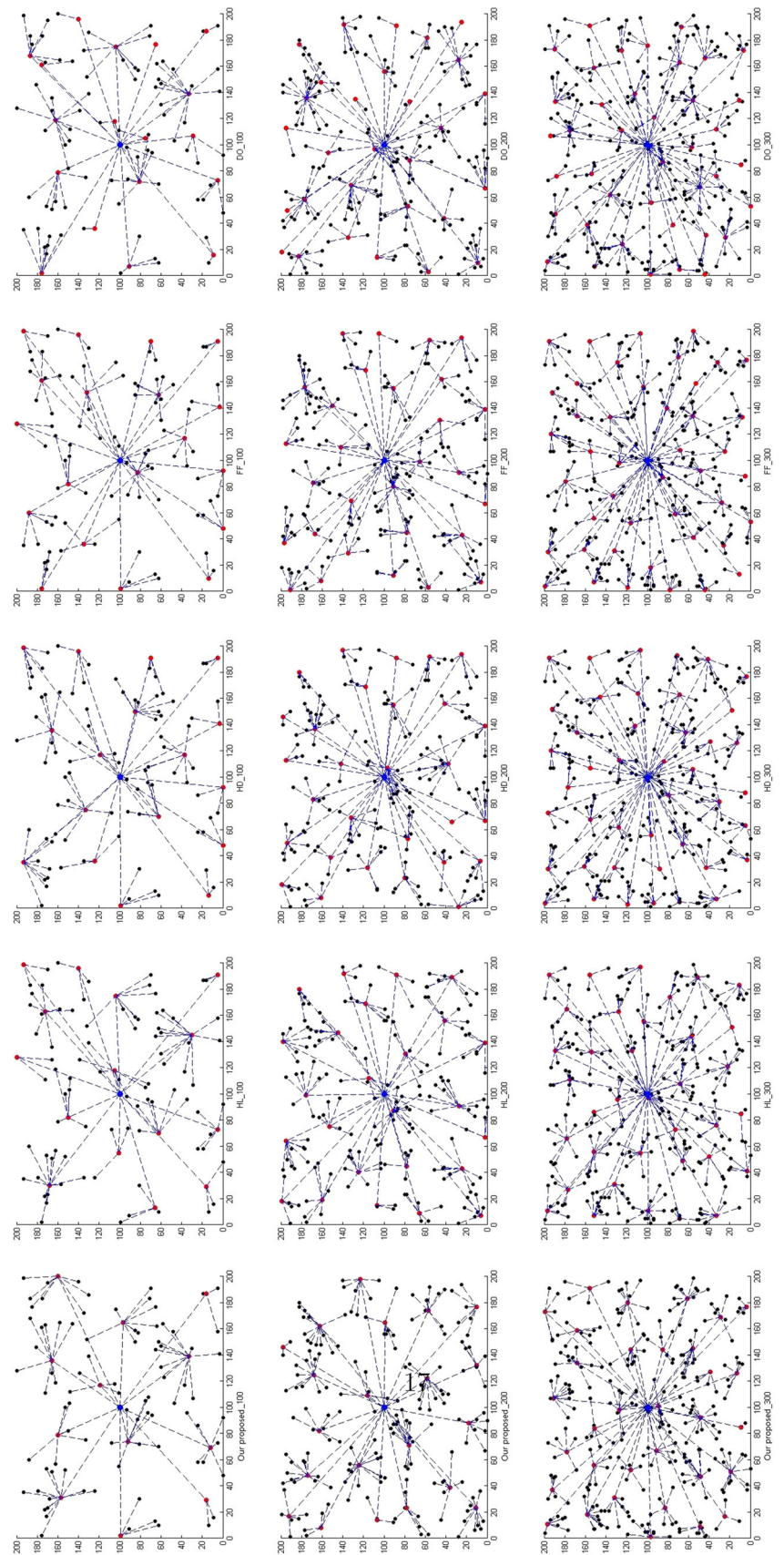
In order to test the stability and robustness of our proposed scheme. Our proposed scheme and three improved strategies run sixteen times under the same condition except for the random seeds used. Table 7 and 8 list the standard deviation (*StD*), average amount (*Avg*), maximum amount (*Max*) and minimum amount (*Min*), where each value evaluates the amount of stations outcome. From the table, we see the size of network for 100- to 400-node networks, the *StD* value stays below than one. Even the maximum value is still better than other methods in Table 1. Table 8 shows the improved strategies that the biggest value in *StD* is 1.5 in 600-node, but the maximum amount value 181 is still smaller than other methods.

Figures 2, 3 and 4 depict the configuration graph of the protocol. The sink node in blue deployed in the center of area. Those nodes in red are stations that are founded out by each method. The rest of nodes in black are common sensor nodes, which are responsible for sensing and collecting data

from environment and transmit the data to stations. For 400- to 600-node network, it is too complex to show the link of the node to the sink node.

Table 9 lists the data sensing effective ratio value of the five different methods. The data sensing effective ratio is defined as the value of the practical number of important events occurred (*Rn*) in area to the value of the real number of packages to the events sent by sensor nodes to relay nodes. The lower ratio of the data, the more duplication packages were sent by the sensor nodes, and the more energy consumes, the worst. The *Rn* values are fifty times of the number of sensor nodes amount in network. Each events will be discovered by nearby sensors within radius *r* circles, if there is no one sensor in this area, then the sensors within radius *2r* circles will be called to work, then the sensors within *3r* circles, it will stopped after the radius go out of the largest sensing radius with *R*. Once there are sensors that response to the events, other farther sensors will be suppressed, and then the sensing sensors transfer the collected data to relay nodes. From the table, it shows our proposed method gets the higher sensing effective ratio than the others.

Fig. 2 Network configurations graph 100–300 nodes



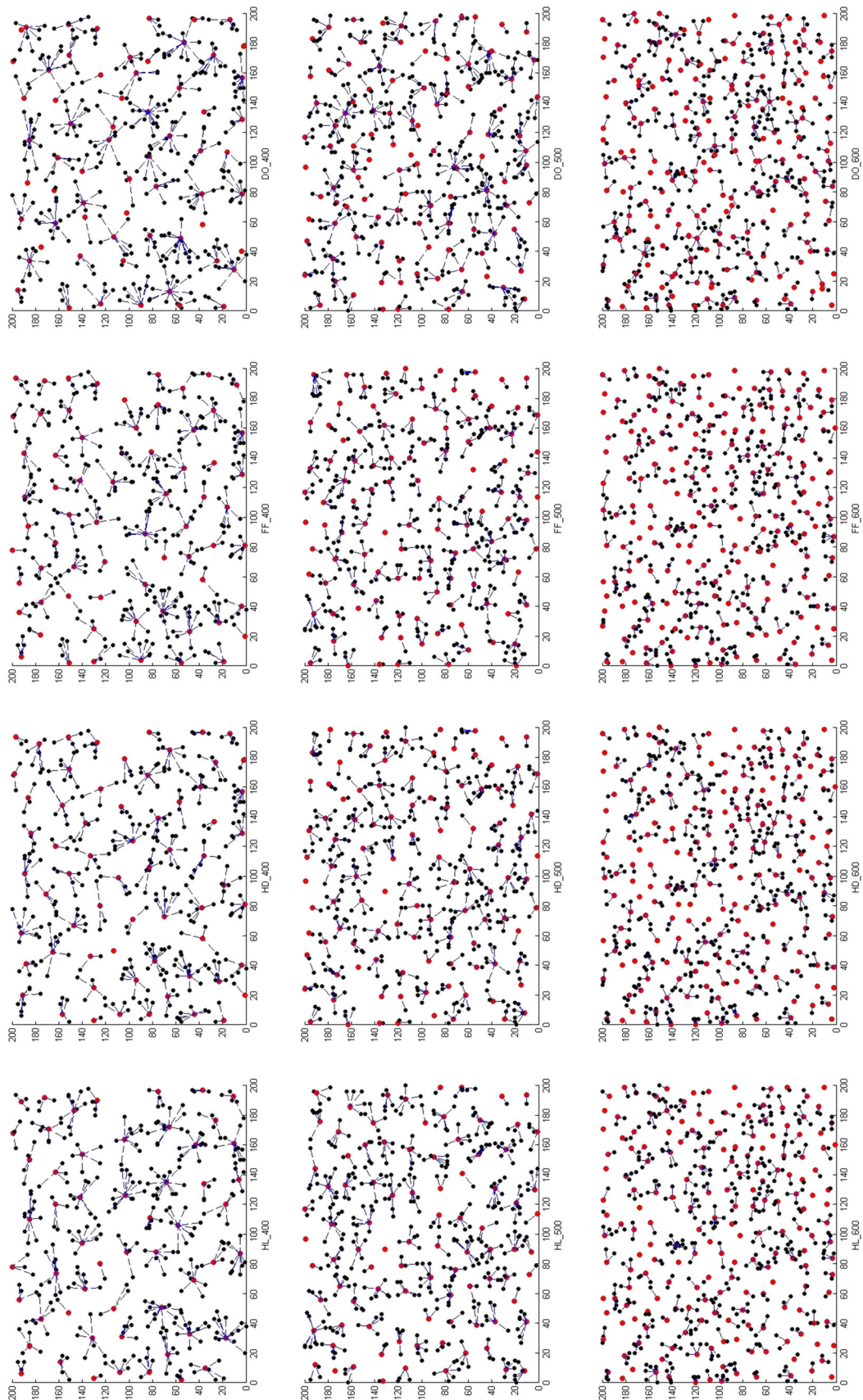


Fig. 3 Network configurations graph 400–600 nodes (HL, HD, FF, DO)

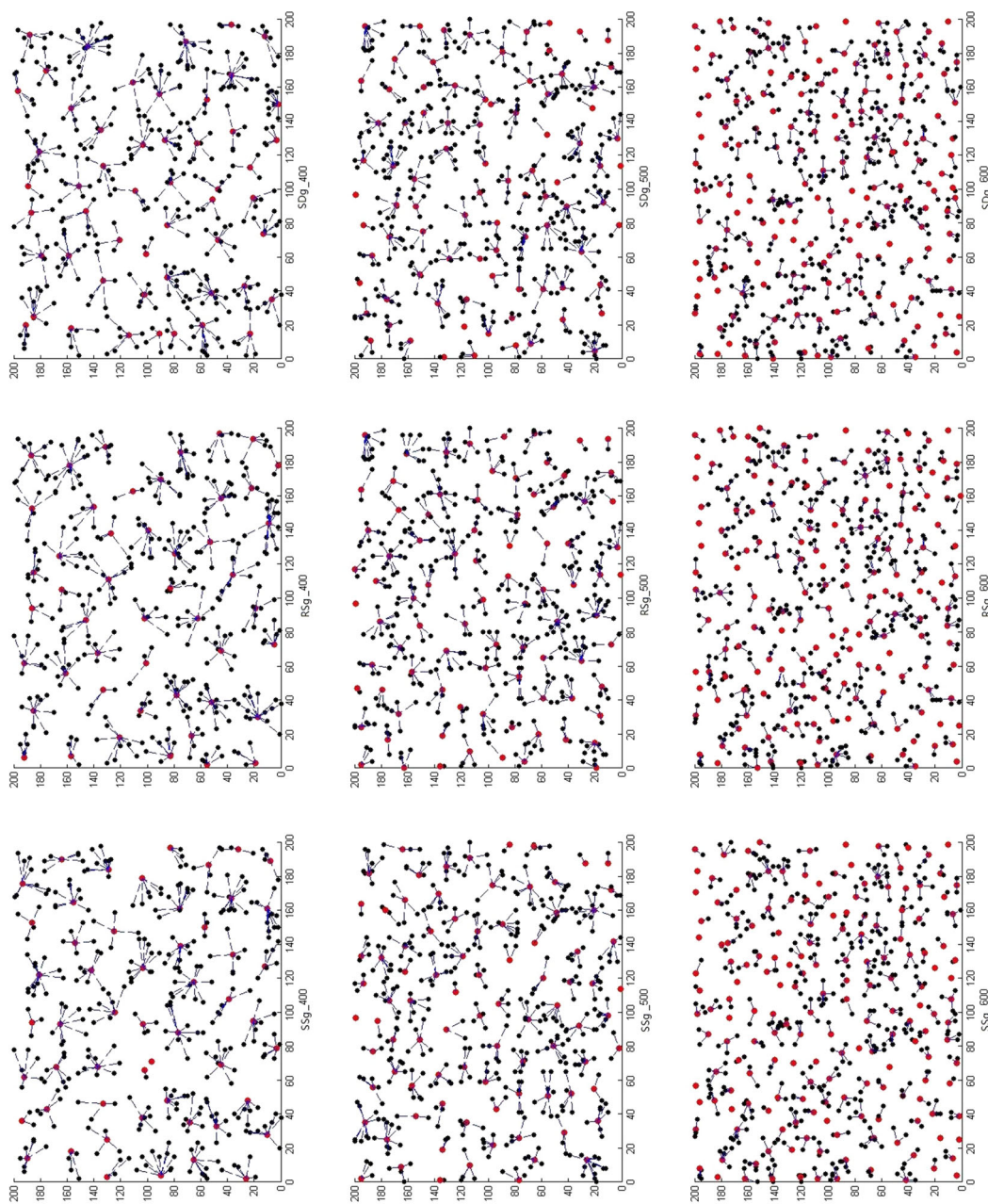


Fig. 4 Network configurations graph 400–600 nodes (SSg, RSg, SDg)

Table 10 shows the total energy consumption (*Total E*) for 100- to 600-node networks for forwarding 5000–30,000 packages. Figure 5 compares the exactly each item energy consumption of sensing activities, receiving and transmitting activities. *SE* sums the energy cost for sensing process by sensor nodes, and *RE* records the receiving cost by relay nodes. The other *SS* and *RS* are the cost for sending message by sensor nodes and relay nodes. *Total E* account the total energy cost for all the operations showed in table. Our proposed scheme spends the less energy than the other methods.

6 Conclusion

In this paper, we introduce *selecting candidate middle layer phase* and *genetic algorithm phase* two parts to construct an energy-aware middle layer oriented routing network, where the middle layer are consist of the node stations and their locations are chosen by special strategy. This strategy needs to consider the distance between common node and the station, the amount of the stations should be as less as possible while the network stays full coverage. In the experiment, we implement four various strategies and compare their perfor-

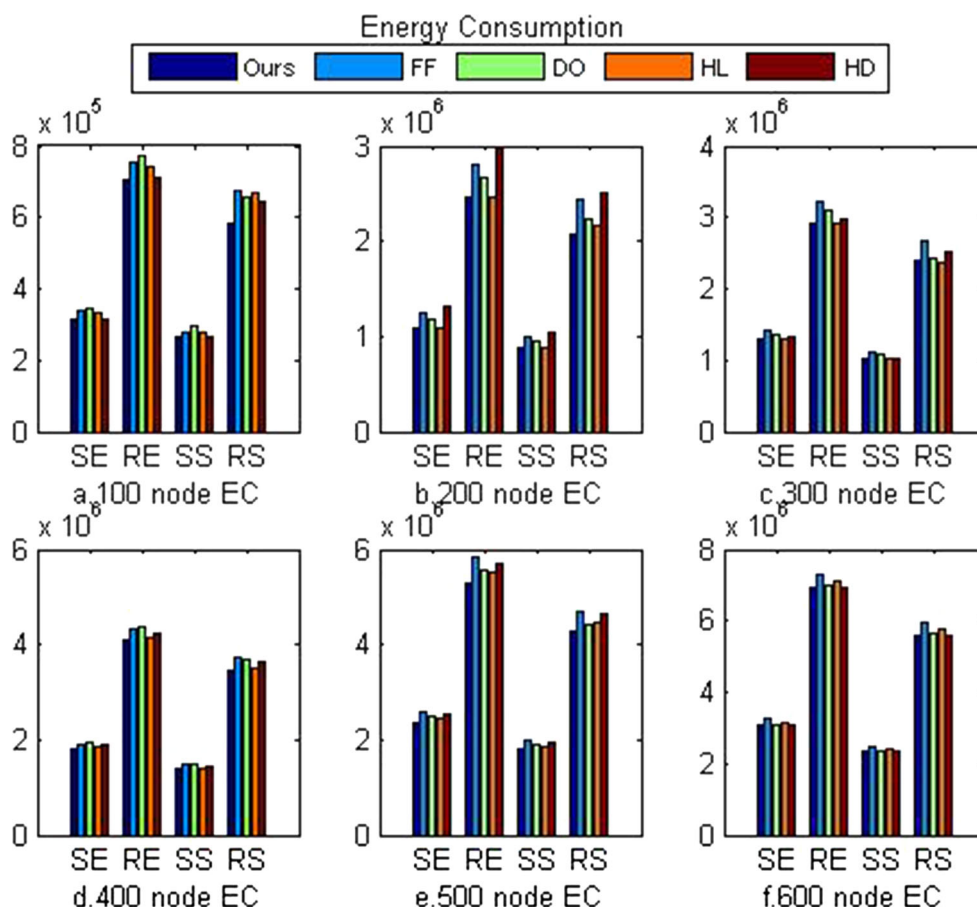


Fig. 5 Energy consumption

mance with our proposed scheme. The experimental results show that our proposed scheme spends fewer stations than the other ways, which decrease a lot of construction fee. Furthermore we also present three methods to improve our proposed scheme, which can be effectively applied to the expansion of network scale problem. Whats more, our proposed scheme gets the higher Data sensing effective ratio of the events data, the fewer duplication packages were sent by the sensor nodes, and the less energy consumes, the better. The experiment results also show that our schemes consume less energy than the other methods and thus prolong the network lifetime.

Acknowledgements This work is partially supported by the Fujian Provincial Natural Science Foundation, China, under Grant No. 2017J01730; and partially supported by the Key Project of Fujian Education Department Funds (JA15323), Shenzhen Innovation and Entrepreneurship Project with the Project Number: GRCK2016 0826105935160.

References

1. Yan, R., Sun, H., & Qian, Y. (2013). Energy-aware sensor node design with its application in wireless sensor networks. *IEEE Transactions on Instrumentation and Measurement*, 62(5), 1183–1191. doi:10.1109/TIM.2013.2245181.
2. Yu, J., Qi, Y., Wang, G., & Gu, X. (2012). A cluster-based routing protocol for wireless sensor networks with nonuniform node distribution. *AEU-International Journal of Electronics and Communications*, 66(1), 54–61.
3. C, Y.-H., Chen, C.-M., Lin, Y.-H., & Sun, H.-M. (2013). Sashimi: Secure aggregation via successively hierarchical inspecting of message integrity on WSN. *Journal of Information Hiding and Multimedia Signal Processing*, 4(1), 57–72.
4. Chang, F.-C., & Huang, H.-C. (2016). A survey on intelligent sensor network and its applications. *J. Netw. Intell*, 1(1), 1–15.
5. Bagci, H., & Yazici, A. (2013). An energy aware fuzzy approach to unequal clustering in wireless sensor networks. *Applied Soft Computing*, 13(4), 1741–1749. doi:10.1016/j.asoc.2012.12.029.
6. Tuna, G., Gungor, V. C., Gulez, K., Hancke, G., & Gungor, V. (2013). Energy harvesting techniques for industrial wireless sensor networks. In G. P. Hancke & V. C. Gungor (Eds.), *Industrial wireless sensor networks: Applications, protocols, standards, and products* (pp. 119–136). New York: CRC Press.
7. Pantazis, N. A., Nikolidakis, S. A., & Vergados, D. D. (2013). Energy-efficient routing protocols in wireless sensor networks: A survey. *IEEE Communications Surveys and Tutorials*, 15(2), 551–591. doi:10.1109/SURV.2012.062612.00084.
8. Nguyen, T.-T., Dao, T.-K., Horng, M.-F., & Shieh, C.-S. (2016). An energy-based cluster head selection algorithm to support long-lifetime in wireless sensor networks. *J. Netw. Intell*, 1(1), 23–37.
9. Goyal, D., & Tripathy, M. R. (2012). Routing protocols in wireless sensor networks: A survey. In *2012 second international confer-*

- ence on advanced computing and communication technologies (pp. 474–480). IEEE.
10. Zhang, D., Li, G., Zheng, K., Ming, X., & Pan, Z.-H. (2014). An energy-balanced routing method based on forward-aware factor for wireless sensor networks. *IEEE Transactions on Industrial Informatics*, *10*(1), 766–773. doi:10.1109/TII.2013.2250910.
 11. Vazirani, V. V. (2013). *Approximation algorithms*. Berlin: Springer.
 12. Du, H., Xu, Y., & Zhu, B. (2015). An incremental version of the k-center problem on boundary of a convex polygon. *Journal of Combinatorial Optimization*, *30*(4), 1219–1227. doi:10.1007/s10878-015-9933-3.
 13. Liang, D., Mei, L., Willson, J., & Wang, W. (2016). A simple greedy approximation algorithm for the minimum connected k-center problem. *Journal of Combinatorial Optimization*, *31*(4), 1417–1429. doi:10.1007/s10878-015-9831-8.
 14. Akkaya, K., & Younis, M. (2003). An energy-aware QoS routing protocol for wireless sensor networks. In *Distributed computing systems workshops, 2003. Proceedings 23rd international conference on IEEE* (pp. 710–715). <http://csdl.computer.org/comp/proceedings/icdcs/2003/1921/00/19210710abs.htm>.
 15. Anker, T., Bickson, D., Dolev, D., & Hod, B. (2008). Efficient clustering for improving network performance in wireless sensor networks. In R. Verdone (Ed.), *Wireless sensor networks, Vol. 4913 of Lecture Notes in Computer Science* (pp. 221–236). Springer. doi:10.1007/978-3-540-77690-1_14.
 16. Banerjee, S., & Khuller, S. (2001). A clustering scheme for hierarchical control in multi-hop wireless networks. In *INFOCOM* (pp. 1028–1037). www.ieee-infocom.org/2001/paper/676.ps.
 17. W, H., Ni, M.-M., & Zhong, Z.-D. (2010). A novel energy efficient clustering algorithm for dynamic wireless sensor network. *Journal of Internet Technology*, *11*(1), 103–107.
 18. Forero, P. A., Cano, A., & Giannakis, G. B. (2011). Distributed clustering using wireless sensor networks. *IEEE Journal of Selected Topics in Signal Processing*, *5*(4), 707–724. doi:10.1109/JSTSP.2011.2114324.
 19. Ge, R., Ester, M., Gao, B. J., Hu, Z., Bhattacharya, B., & Ben-Moshe, B. (2008). Joint cluster analysis of attribute data and relationship data: The connected k-center problem, algorithms and applications. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, *2*(2), 7. doi:10.1145/1376815.1376816.
 20. Karaboga, D., Okdem, S., & Ozturk, C. (2012). Cluster based wireless sensor network routing using artificial bee colony algorithm. *Wireless Networks*, *18*(7), 847–860. doi:10.1007/s11276-012-0438-z.
 21. Yu, Y., Govindan, R., & Estrin, D. (2001). Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks.
 22. Ramesh, K., & Somasundaram, D. K. A comparative study of clusterhead selection algorithms in wireless sensor networks. ArXiv preprint [arXiv:1205.1673](https://arxiv.org/abs/1205.1673).
 23. Feldmann, A. E. (2015). Fixed parameter approximations for k-center problems in low highway dimension graphs. In *International colloquium on automata, languages, and programming, Vol. 9135 of Lecture Notes in Computer Science, Springer* (pp. 588–600). Springer. doi:10.1007/978-3-662-47666-6_47.
 24. Chechik, S., & Peleg, D. (2015). The fault-tolerant capacitated k-center problem. *Theoretical Computer Science*, *566*, 12–25. doi:10.1016/j.tcs.2014.11.017.
 25. Elloumi, S., Labbé, M., & Pochet, Y. (2004). A new formulation and resolution method for the p-center problem. *INFORMS Journal on Computing*, *16*(1), 84–94. doi:10.1287/ijoc.1030.0028.
 26. Gonzalez, T. F. (1985). Clustering to minimize the maximum inter-cluster distance. *Theoretical Computer Science*, *38*, 293–306.
 27. Harel, D., & Koren, Y. (2002). Graph drawing by high-dimensional embedding. In S. G. Kobourov, & M. T. Goodrich (Eds.), *International symposium on graph drawing, Vol. 2528 of Lecture Notes in Computer Science, Springer* (pp. 207–219). Springer. <http://link.springer.de/link/service/series/0558/bibs/2528/25280207.htm>.
 28. Robič, B., & Mihelič, J. (2005). Solving the k-center problem efficiently with a dominating set algorithm. *CIT. Journal of computing and information technology*, *13*(3), 225–234.
 29. Marta, M., & Cardei, M. (2009). Improved sensor network lifetime with multiple mobile sinks. *Pervasive and Mobile Computing*, *5*(5), 542–555. doi:10.1016/j.pmcj.2009.01.001.
 30. Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and computing*, *4*(2), 65–85.
 31. Back, T. (1993). Optimal mutation rates in genetic search. In S. Forrest (Ed.), *Proceedings of the 5th international conference on genetic algorithms* (pp. 2–8). Morgan Kaufmann.
 32. Beasley, J. E., & Chu, P. C. (1996). A genetic algorithm for the set covering problem. *European Journal of Operational Research*, *94*(2), 392–404.
 33. Thi-Kien Dao, T.-S. P., & Nguyen, T.-T. (2015). A compact artificial bee colony optimization for topology control scheme in wireless sensor networks. *Journal of Information Hiding and Multimedia Signal Processing*, *6*(2), 297–310.
 34. T, Y.-C., & Huang, C.-F. (2005). A survey of solutions for the coverage problems in wireless sensor networks. *Journal of Internet Technology*, *6*(1), 1–8.
 35. Levy, P. S., & Lemeshow, S. (2013). *Sampling of populations: Methods and applications*. New York: Wiley.
 36. Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. In *Flow and the foundations of positive psychology* (pp. 35–54). Springer.



Lingping Kong is a Ph.D. student in Harbin Institute of Technology Shenzhen Graduate School. Her current research interests include wireless sensor network, artificial intelligence etc.



Jeng-Shyang Pan received the Ph.D. degree in Electrical Engineering from the University of Edinburgh, U.K. in 1996. Currently, he is a Professor in Harbin Institute of Technology Shenzhen and Fujian University of Technology. His current research interests include soft computing, information security, big data mining and signal processing.



Václav Snášel is a researcher and university professor at the VSB-Technical University of Ostrava. His research and development experience includes over 25 years in the Industry and Academia. He works in a multi-disciplinary environment involving artificial intelligence, conceptual lattice, information retrieval, semantic web, neural network, web intelligence etc.



Tien-Wen Sung received his Ph.D. degree in Computer and Communication Engineering from Department of Electrical Engineering at Cheng Kung University, Taiwan in 2014. He is currently an Associate Professor in College of Information Science and Engineering, Fujian University of Technology, China. His research interests include wireless sensor networks and e-learning.



Pei-Wei Tsai received his Ph.D. in Electronic Engineering in 2012. He is the Executive Editor of Journal of Network Intelligence. He is a research fellow at the Swinburne University of Technology in Australia. His research interests include swarm intelligence, information security, big data analysis, intelligent transportation, and wireless sensor network.