Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **1**, January **2018**

Hierarchical Routing Strategy for Wireless Sensor Network

Jeng-Shyang Pan[1,2,3], Lingping Kong[2,*]
Tien-Wen Sung[1], Shu-Chuan Chu[4] and John F. Roddick[4]

[1]Fujian Provincial Key Laboratory of Big Data Mining and Applications,
College of Information Science and Engineering,
Fujian University of Technology, Fujian, China

[2]Innovative Information Industry Research Center,
Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China

[3]Chaoyang University of Technology, Taichung, Taiwan

[4]School of Computer Science, Engineering and Mathematics,
Flinders University, Australia

jspan@cc.kuas.edu.tw

Received June, 2017; revised December, 2017

Abstract. *Wireless sensor networks (*WSNs*) are designed for a large scale monitoring applications such as military surveillance, medical treatment, environmental monitoring and industry management. In this network, usually hundreds or thousands of low-cost sensor nodes are deployed. These sensing nodes detect the events in the environment and pass an upstream message towards a Sink node. The Sink node is responsible for processing data those are collected by sensor nodes. However, bottleneck exists for* WSN*: these inner nodes whose positions are close to the Sink node run out of power much earlier than those outer nodes. Because those inner nodes not only perform sensing operation, but also replay data originated from other nodes. In this paper,* GAHL *routing scheme based on genetic algorithm and Harel method is proposed to solve the problem. The scheme finds the small number of center heads in the network. Those center heads play a role in balancing the workload of inner nodes. With the involvement of center head, a three-tier network is built. From the simulation experiment, the scheme shows a good performance compared to other well-known algorithms. We observe a decrease in the number of center heads, as well as an improvement in the stability of the scheme.*
Keywords: Wireless sensor network, Genetic algorithm, Routing strategy

---

1. **Introduction.** A wireless sensor network (*WSN*) is a network comprising many wireless sensor nodes which equipped with sensing and communicating components [1, 2]. Usually those sensor nodes are scatted randomly or pre-defined across the whole area, and the number of sensor nodes could be hundreds or thousands. It has been the preferred choice for the design and deployment strategy in monitoring and controlling the next generation systems. The large numbers of low-cost sensors are used to monitor regions over the ground surface, underwater and atmosphere. *WSN* has attracted a lot of researcher attentions, it has been applied in many fields such as military surveillance, medical treatment, environmental monitoring and industry management [3, 4]. The most significant benefit of sensor network is that it extends the exploring capability to the physical environment where human beings cannot reach [5, 6]. Those sensor nodes can

be operated in environments that are hostile, challenging or ecologically where it is sensitive for people to visit. Then nodes deliver valuable data through the network to the processing center. A sensor node is a device that installed with sensing, computation and communication components. Sensing components are assigned to monitor phenomena such as motion, light, humidity, pressure and temperature et.al based on different application requirements. The processing component is used to compute the data and receive data from other companions. A sensor node reports sending or receiving packets to neighbor with the other nodes cooperative by the communication module. Due to the nature traffic in which all the sensor nodes deliver their sensed data toward the Sink in wireless sensor network. The situation around the Sink becomes more congested along with all the channel data flowing towards the Sink [7, 8]. More energy is demanded in these inner nodes compares to the outer nodes. Inner nodes must assist outer nodes to contact with the Sink. However, for most *WSNs*, sensor nodes are usually deployed by aircraft without re-placement and re-charge [5]. Those sensor nodes are constrained by battery level and computing capacity. Once the inner nodes run out of the power, the outer nodes cannot transfer their data to the Sink, hence the *WSN* is operating as expected. A series of problems on transmission errors, packet collision, interference, node failure or other unforeseeable reasons would happen after the inner nodes died. In this case, how to manage the network and optimize the lifetime of sensor nodes are the vital and challenging issue.

k-center problem is a combinatorial optimization problem. Given a set of nodes with fixed distances, one wants to find a set of vertices in which the largest distance between any node and its closest vertex is minimized. Since the k-center problem is NP-hard problem, thus there is no efficient strategy that always returns the right answer. There are many algorithms are proposed to solve the k-center problem [9], one of the most widely and popular used is farthest first traversal (FF) [10], which is introduced by Hochbaum, Shmoys and Gonzalez. The first point is chosen arbitrarily and each successive point is as far as possible from all previously chosen points. Those chosen points comprise the solution. D Harel [11] uses the idea to solve the high dimensional graph, but the new center is to choose the one which has the longest distance to the last added center instead of all chosen centers. Based on the difference of distance computing formulas, Harel methods have two forms, Harel_Line (HL), which uses Euclidean distance, Harel_Dijkstr (HD), which uses Dijkstra metrics distance. There are a series of minimum set cover problems [12, 13] also called minimum dominating set (DO) [10]for solving selecting pivots problem, which is a variant of k-center problem. Dominating defined as: if node a node b are neighbors, then node a dominates node b , or node b dominates node a . The brief overview of DO algorithms starts with an empty set T, then T grows based on lazy principle. That is, adding a new point into T as late as possible. Since sensing radius and receiving packet capacity of one sensor node is a constant value in our network, so the implementation DO scheme based on minimum dominating set works without Bottleneck Graph part [14].

The rest of the paper is organized as follow. Section 2 introduces the aim of the work and some related works. Section 3 presents *GAHL* routing scheme and the detail of its process. Simulation experiments and its results analysis will be shows in Section 4, followed by the conclusion in Section 5.

2. **Related Works.** In discussed in introduction section, bottleneck existing for *WSN*. However, when the network works, the bottleneck problem cannot be handled easily as long as the two layer network structure remains the same. The big difficulties of reconstruction for this kind of network are the selection of center heads and the construction of

third layer [15, 16]. A center head would cost much than a common sensor node consumes, so the number of center heads should not be too many. But if there are a few number of center heads it would make no difference to the original network structure [17]. In this paper, *GAHL* routing scheme is proposed which is trying to find center heads as few as possible. This scheme is derived from genetic algorithm and HL method.

Routing has always played a key issue in wireless sensor networks research. In the literature, researchers have devoted a lot of efforts on designing *WSN* routing protocol. Routing schemes may be diverse depending on the nature of user applications and the architecture of the underlying network [18]. Clustering is a popular technique for building routing strategy in a more tier structural network. In the hierarchical routing network [19, 20], nodes in the upper tier are called cluster head and acts as a routing backbone, while nodes in the lower tier perform surveillance tasks. Each cluster head connects with Sink node directly. A large of clustering algorithms have been investigated before. One of the first clustering based protocols is LEACH [21], it utilizes randomized rotation of the cluster head to balance the energy consuming workload. This method is well-suited for applications where scheduling monitoring and data collection. HEED [22] algorithm is proposed for sensor applications requiring efficient data aggregation. It utilizes residual energy and node degree to balance clusters with low message overhead. Most of the hierarchical routing strategies are built based on the assumption that the number of clusters is clear. It does not effectively take advantage of the environment factor and location information of sensor nodes for this assumption. In our paper, the number of cluster is not pre-defined. It tries to build a full coverage of sensor network with as fewer clusters as possible.

Genetic algorithm [11-14] imitates the adaptive evolution procedure present in nature and it is an effective stochastic optimization search algorithm. As with the genetic algorithm, there are many intelligence algorithm include Evolutionary algorithms (EA), particle swarm optimization (PSO), ant colony optimization (ACO) and their variants. Genetic algorithm generates solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, selection, mutation and crossover. It has been always a good tool for optimization task. M.V Ramana [21] designs a protocol based on genetic algorithm to test packet delivery ratio, end to end delay, number of hops to destination and jitter with various node mobility levels. Suneet K. [22] also puts it to application for relay node placement which provides k-connectivity of the sensor nodes. However, genetic algorithms do not scale well with complexity. The number of elements which are exposed to mutation may exponential increase in search space size. In addition, repeated fitness function evaluation for complex problems are often the most limiting factor of evolutionary algorithms.

In this paper, to successfully monitor the environment and reliable deliver the critical data, several center heads are chosen by the *GAHL* scheme and these heads play a role in balancing the workload of Sink node. For decreasing the construction fee cause by center heads, the scheme is designed to find the center heads as few as possible. *GAHL* scheme is designed to shorten the exploration space in the initialization part and work with a simple and effective fitness function. From the simulation experiments, the scheme shows a good performance compared to the other well-known algorithms.

3. *GAHL* Implement.

In our simulation, hierarchical routing runs in a three-tier structural network. A amount number of center heads are deployed which play a collecting and forwarding role in the network. Those heads are designed for balancing the workload of inner nodes. The important and challenging task for this routing is that the number of center heads is unclear. If there are too many heads, it will increase extra construction and deployment cost. A small number of heads would not satisfy the coverage ratio need. *GAHL* scheme is for selecting several vertices from a fixed number of sensor nodes, those sensor nodes can connect to at least one vertex in their communication range. The fewer vertices there are in area, the better. Those selected vertices work as the center head and center heads can communicate with the Sink directly. An implementation of a genetic algorithm begins with generate chromosome population. Then use the fitness function to evaluate each chromosome. Operate crossover and mutation to the population. Repeat the above steps until a good result is found.

3.1. **The annotation of symbol.** Table 1 summarizes the symbols used in the paper.

TABLE 1. Annotation of symbol

Symbol	Annotation
N	$N = N_1, N_2, ..., N_n$, sensor nodes set of the area
M	$M = M_1, M_2, ..., M_n$, the initial chromosomes
HL_r	The chosen nodes of HL method
r	Sensing radius of one node
$\alpha_1, \alpha_2, \alpha_3$	Constant parameters
k	The number of value 1 in one chromosome
C_n	The number of value 1 in one chromosome
$\|H, Sink\|$	The distance between a head to Sink node
$\|N, Sink\|$	The distance between a node to Sink node
$L,\ W$	The length and width of network area

3.2. **The process *GAHL*.** Coding each node in the network with a unique numbering from 1 to n. Randomly pick a digit from n, then use the node which the digit presented as the starting point of HL to generate one solution. In the HL operation, each successive chosen head is the one who has the longest distance with the last head-to-be. Mark the solution as *HL_r*. Repeat it m times with different starting digit, then we can get *HL_rs*.

The configuration of one chromosome is illustrated as in Fig. 1. Each chromosome is an n-bit binary string with value 0 or 1. A value 0 bit implies the node represented is a common node, and the value 1 bit represented node is a head node. Each *HL_r* generates a chromosome. Set the corresponding bit value 1 which bit represented node that node come from *HL_r*, and the rest of nodes corresponding bits note as value 0. For better evolving the population, several more value 1 bits should be extra added into one chromosome. In Fig. 1, bits number one, three and n-1 are head node candidates in this chromosome.

Code	1	2	3	4	5	...	n-1	n
Bit string	1	0	1	0	0	...	1	0

FIGURE 1. A chromosome structure

Evaluate the M chromosome population using fitness function. Store the best one with best fitness value. Before parents selection process, simply divide the chromosomes into two sets, Set A and Set B, Set A has those better chromosomes with better a half fitness values. The other half are stored in Set B. Randomly choose the father and mother chromosomes from the two sets. Let $C1$ (mother) $C_{i,1}, .., C_{i,n}$ and $C2$ (father) $C_{j,1}, .., C_{j,n}$

be the parent chromosome. Use one middle-point operator crossover operation, the child $C3$ strings of $C1$ and $C2$ is shown as Fig. 2:

$$C_2 := C_{i,1}, C_{i,2}, ..C_{i,[n/2]}, C_{j,[n/2]+1}, ..C_{j,n-1}, C_{j,n} \tag{1}$$

Code	1	2	...	k	k+1	...	n-1	n
C1	1	0	...	0	1	...	0	1

Code	1	2	...	k	k+1	...	n-1	n
C2	0	1	...	1	0	...	1	0

Code	1	2	...	k	k+1	...	n-1	n
C3	1	0	...	0	0	...	1	0

FIGURE 2. Crossover process

Mutation process works by inverting a bit value in the chromosome with a small probability, see Fig. 3, the mutation process of node number 3 from value 1 into value 0. Mutation rate value used in simulation is 0.02. The following figure shows the transformation.

Code	1	2	3	...	n-1	n
Bit string	1	0	1	...	0	1

Code	1	2	3	...	n-1	n
Bit string	1	0	0	...	1	0

FIGURE 3. Mutation process

Checkup process: Each chromosome presents a resolution of the k-center problem, so it must be validity and feasibility. It requires all the value 1 bits could dominate all the value 0 bits in one chromosome. If it is not, this one chromosome needs to be fixed. Change the uncovered nodes to be center heads.

3.3. **Fitness function.** A good fitness function could accelerate the coming of optimization resolution. Here presents a fitness function F, which considers three factors: the deployment area of heads, the distance between node to Sink and the number of heads. This parameter k is the number of heads in one chromosome. C_n denotes the number of nodes that dominated by heads without duplicate. The network area is divided into *Sectol* grids, which is a constant value $[W \times L/r^2]$. *GridN* records the grids number where those heads coming from. Coefficient $\alpha_1, \alpha_2, \alpha_3$ are constants.

$$F = \alpha_1 \times C_n + \alpha_2 \times \frac{GridN}{Sectol} + \alpha_3 \times \frac{\sum_{1\le i\le k} |H_i, Sink|}{\sum_{1\le j\le N} |N_j, Sink|} \tag{2}$$

An overview of *GAHL* scheme steps, table 2 shows the pseudo code of *GAHL* scheme.

Algorithm 1 *GAHL* implement process

function GAHL(N, M, HL_r, p)
Input $N = N_1, N_2, ..., N_n, M = M_1, M_2..., M_M$, HL_r={}
Output: M
  for $it = 1$ to threshold **do**
    for $it = 1$ to threshold **do**
      Copy N as $C = C_1, C_2, ..., C_n$
      Randomly pick z, $z \in C$, set $HL\_r = z$
      Remove node z from C. Let $p = z$
      Remove any node in C covered by p
      while $C \neq null$ **do**
        $z = C_i$, where $\max_{C_i \in C} C_i, p$
        $HL\_r = HL\_r \cup z$
        $p = z$, remove node z from C.
        Remove any node in C covered by p.
      end while
      Initialize M_{it} using HL_r
    end for
    Run crossover and mutation process
    Checkup process
  end for
  Output M
end function

4. **Experiment Analysis.** The simulation environment is 200 units network. All the sensor nodes are deployed randomly in the field. There is one Sink in the area, which is responsible for collecting and processing the data. The center head is in charge of receiving the data from sensing node then transfer the data to Sink node. All the sensor nodes are equipped with the same communication and sensing components. The amount of sensor nodes in the network varies from 100 to 400. In the simulation process, the communicating radius is the sensing radius. The sensing radius is 40-units, 30-units, 25-units and 20-units for 100-node, 200-node, 300-node and 400-node sensor network respectively. The mutation rate is 0.02, the chromosome population is 20. The experiment runs 42 times in the same condition for different random seed. The table 2 shows the result of the amount of center heads by five methods, farthest first traversal (FF) and dominating set (DO) and two Harel methods. For 400-node network, *GAHL* only uses 51 heads to cover the rest of nodes in network, and the other methods use at least 55 heads. And table 3 shows the fitness values of those methods. The larger the fitness value is, the better.

Table 2. The number of center results of five algorithms

Algorithm	Number of Nodes			
	100	200	300	400
GAHL	14	22	32	51
FF	19	29	45	67
DO	18	29	44	64
HL	15	25	38	55
HD	17	29	44	64

Table 3. The fitness values results of five algorithms

Algorithm	Number of Nodes			
	100	200	300	400
GAHL	128.6133	58.9531	88.6425	115.3778
FF	27.0415	56.6617	84.4375	110.1706
DO	27.3463	56.6602	84.7272	111.1261
HL	2 8.3029	57.9648	86.7009	114.0763
HD	27.6779	56.6622	84.7557	111.1405

The following table shows the robustness of the *GAHL* scheme, which runs 18 times in the same condition for different random seed. Table 4 lists the maximum value (Max), minimum value (Min) and standard deviation (StD) and average value (AVG) of our scheme. In the best case, our scheme only uses 51 heads to build a full coverage 100-sensor node network, and 53 heads for the worst case. However even for the worst case, it is still better than the other four methods, FF needs 67 and Dominating set needs 64. The StD value is bigger than one for 300-node network, but the Max value 34 heads are better than the others.

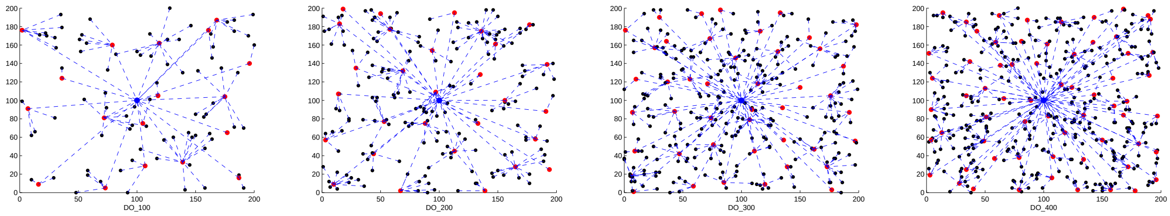
Table 4. The stability of GAHL algorithm

Value	Number of Nodes			
	100	200	300	400
Max	15	23	34	53
Min	14	22	32	51
AVG	14.7777	22.9444	33.1666	52.3333
StD	0.5522	0.2357	1.2287	0.6859

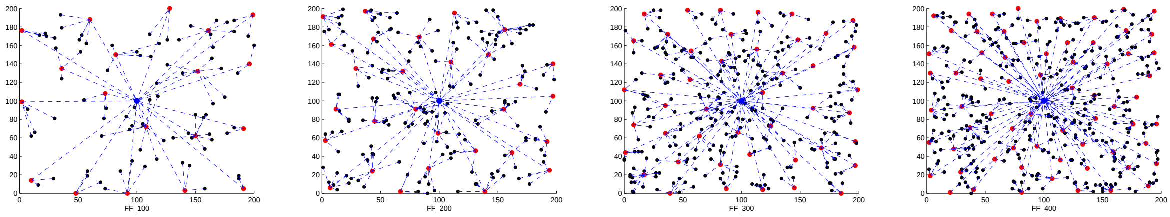
The following Fig. 4 simulate the configuration of network by five methods. The node in blue in the center is the Sink node, which is for processing the data. Those nodes in red are center heads and they can communicate with Sink node directly and collect the data from sensing node. The rest of nodes in black are common sensing nodes that are for collecting the data from environment. Each methods shows 100-node to 400-node four different size of networks.

5. **Conclusion.** The traditional two-tier sensor network suffers the energy bottleneck problem due to the unbalanced workload of inner nodes and outer nodes. In this paper, *GAHL* scheme based on genetic algorithm and Harel methods in Euclidean distance is proposed to solve the problem. *GAHL* scheme builds a middle tier for forwarding data from sensing nodes to Sink with the help of center heads that are chosen from sensor nodes. From the simulation experiment, the scheme shows a good performance compared to the other well-known algorithms. Not only the number of center heads is decreased, but the stability of the scheme is proved. In the end, the network configuration of routing structure of five methods are depicted.

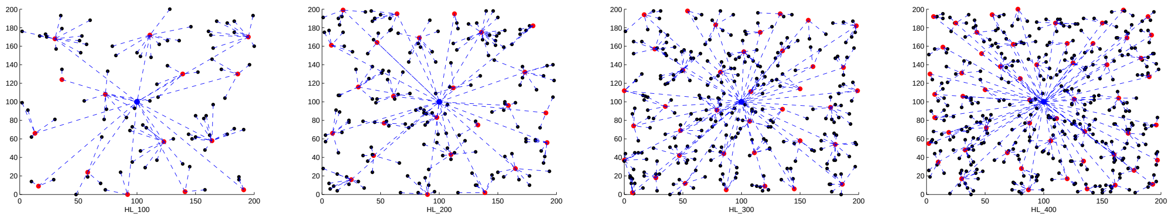
Acknowledgment. The authors would like to thank Dr. Ivan Lee and Dr. Pei-Wei Tsai for their valuable feedback to this paper. This work is partially supported by the Fujian Provincial Natural Science Foundation, China (2017J01730).



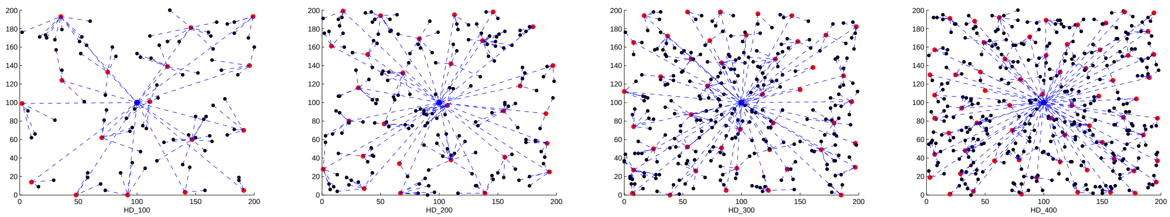
(a) DO



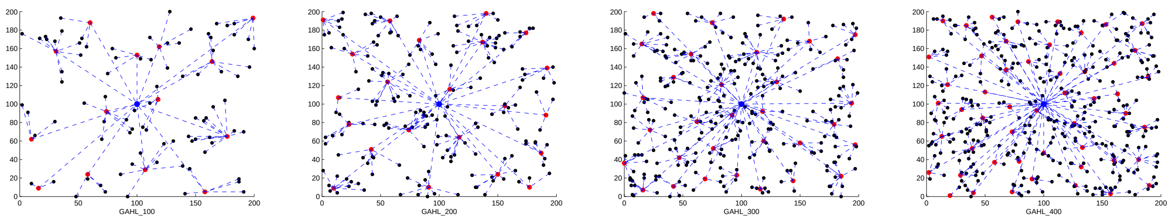
(b) FF



(c) HL



(d) HD



(e) GAHL

Figure 4. Network structure of 100-node to 400-node

References

[1] Mahmood, M. Adeel, W. K. G. Seah, and I. Welch , Reliability in wireless sensor networks: A survey and challenges ahead, *Computer Networks*, vol.79 , pp.166-187, 2015.

[2] F. C. Chang, H. C. Huang , A Survey on Intelligent Sensor Network and Its Applications, *Journal of Network Intelligence*, nol. 1, no. 1, pp. 1-15, 2016.

[3] Y. Y. Lin, A Heterogeneous Network Sharing System Design Based on Portable Devices, *Journal of Internet Technology*, vol. 18, No. 6 , pp. 1441-1449, 2017.

[4] Magno, Michele, A low cost, highly scalable wireless sensor network solution to achieve smart LED light control for green buildings, *IEEE Sensors*, vol.15, no.5, pp. 2963-2973, 2015.

[5] Lewis, Franck L , Wireless sensor networks, *Smart environments: technologies, protocols, and applications.* pp. 11-46, 2014

[6] T.T. Nguyen, J. S. Pan, S. C. Chu, J. F. Roddick, and T. K. Dao, Optimization Localization in Wireless Sensor Network Based on Multi-Objective Firefly Algorithm, *Journal of Network Intelligence*, vol. 1, no. 4, pp. 130-138, 2016

[7] Y. Yanjun, Q. Cao, and A. V. Vasilakos, EDAL: An energy-efficient, delay-aware, and lifetime-balancing data collection protocol for wireless sensor networks. *IEEE conference on Mobile ad-hoc and sensor systems*, 2013.

[8] Sran, Sukhwinder Singh , Energy aware chain based data aggregation scheme for wireless sensor network, *IEEE Conference on Energy Systems and Applications*, 2015.

[9] Hochbaum, Dorit S., and David B. Shmoys, A best possible heuristic for the k-center problem, *Mathematics of operations research,* vol. 10, no.2, pp. 180-184, 1985

[10] Gonzalez, F. Teofilo , Clustering to minimize the maximum inter-cluster distance, *Theoretical Computer Science*, vol. 38, pp. 293-306, 1985.

[11] H. David, and Y. Koren , Graph drawing by high-dimensional embedding, *International Symposium on Graph Drawing. Springer Berlin Heidelberg*, 2002.

[12] Feldmann, Andreas Emil , Fixed Parameter Approximations for k-center Problems in Low Highway Dimension Graphs, *Springer Automata, Languages, and Programming,* pp. 588-600, 2015

[13] 20. E. Sourour, M. Labb, and Yves Pochet , A new formulation and resolution method for the p-center problem, *INFORMS Journal on Computing*, vol. 16, no.1, pp. 84-94, 2004.

[14] R. Borut, and Jurij Mihelic , Solving the k-center problem efficiently with a dominating set algorithm, *CIT. Journal of computing and information technology*, vol. 13, no.3, pp. 225-234, 2005

[15] 15. Forero, Pedro A., Alfonso Cano, and Georgios B. Giannakis , Distributed clustering using wireless sensor networks, *IEEE Journal of Selected Topics in Signal Processing* vol. 5, no.4, pp. 707-724, 2011

[16] B. Suman, and Samir Khuller. , A clustering scheme for hierarchical control in multi-hop wireless networks, *IEEE Computer and Communications Societies,* Proceedings, 2001.

[17] A. Biswa Mohan, and S. V. Rao , Efficient coordination and routing protocol for wireless sensor and actor networks *Springer Recent Development in Wireless Sensor and Ad-hoc Network*, pp.55-67, 2015

[18] H. Mohammad, and Robert Newman , Adaptive routing in wireless sensor networks: QoS optimisation for enhanced application performance, *Information Fusion,* vol. 22, pp. 3-15, 2015

[19] Rajeshwari, P., B. Shanthini, and Mini Prince , Hierarchical Energy Efficient Clustering Algorithm for WSN, *Middle-East Journal of Scientific Research*, vol. 23, pp. 108-117, 2015

[20] L., Tiejun, et al , Space-time hierarchical-graph based cooperative localization in wireless sensor networks, *IEEE Transactions on Signal Processing,* vol. 64, no.2, pp. 322-334, 2016

[21] Razaque, Abdul , H-LEACH: Hybrid-low energy adaptive clustering hierarchy for wireless sensor networks, *IEEE Long Island Systems, Applications and Technology Conference*, 2016.

[22] O. Younis, S. Fahmy , Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks, *IEEE Transaction Mobile Computer*, vol. 3 ,pp. 366C379, 2004.

[23] R., Efrain, A biased random-key genetic algorithm for the capacitated minimum spanning tree problem *Computers and Operations Research,* vol. 57, pp. 95-108, 2015

[24] G. Pedram, and J. Atli Benediktsson , Feature selection based on hybridization of genetic algorithm and particle swarm optimization, *IEEE Geoscience and Remote Sensing Letters,* vol. 12, no.2, pp. 309-313, 2015

[25] 13. L. L. Cui, Y. F. Li, and L. Pan , Study on Coal Consumption Curve Fitting of the Thermal Power Based on Genetic Algorithm, *Journal of Power and Energy Engineering.* vol. 3 no.4, pp. 431, 2015

[26] J. X. Hao, , A genetic algorithm-based learning approach to understand customer satisfaction with OTA websites, *Tourism Management.* vol. 48, pp. 231-241, 2015

[27] Rao, MV Ramana, and T. Adilakshmi , Optimized Cluster with Genetic Swarm Technique for Wireless Sensor Networks, *Indian Journal of Science and Technology.* vol. 9, no.17, 2016

[28] Gupta, K. Suneet , P. Kuila, and Prasanta K. Jana, Genetic algorithm for k-connected relay node placement in wireless sensor networks, *Proceedings of the Second International Conference on Computer and Communication Technologies.* Springer India, 2016.