http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 9 (7): 1403-1408, 2010
ISSN 1812-5638

An Energy Equilibrium Routing Algorithm Based on Cluster-head Prediction for Wireless Sensor Networks

Lejiang Guo, Bingwen Wang, Zhuo Liu and Wei Wang
Department of Control Science and Engineering,
Huazhong University of Science and Technology, Wuhan Hubei, 430074, China

Abstract: Because of the limited energy in wireless sensor networks, the research on routing technology on the network layer is pivotal in the architecture of wireless sensor networks. Aiming at the defect problem of the clustering network, this study presents an energy equilibrium routing algorithm based on Cluster-head Prediction for Wireless Sensor Networks (CP-EERP). The algorithm uses the cluster-head prediction mechanism which improves cluster-head lifetime, balances the energy between nodes and prolongs the network lifetime. CP-EERP includes network initialization phase, cluster building phase, data transmission phase and cluster-head prediction phase. The simulation result shows that the algorithm performs better, in terms of power efficiency and the number of communication neighbors, than the classic routing algorithm.

Key words: WSN, prediction algorithm, energy efficiency, network lifetime, CP-EERP

INTRODUCTION

Wireless Sensor Networks (WSN) have restrictions due to energy, memory and communication ability. The network nodes usually work in unattended environments, such as mine, wetlands, craters, etc. The node does not work when node's battery falls to a certain power. Therefore, it needs to design the reasonable energy-efficient protocol and network topology for WSN, such as Leach and HEED (Boukerche and Samarah, 2008). As shown in Fig. 1, cluster network is the most classic hierarchical network structure in WSN. In the cluster topology, all cluster-member nodes firstly communicate with the cluster-head node. This mechanism avoids all the nodes communicate directly with the sink and greatly saves the network energy. The cluster-head node takes on a variety of tasks and communicates with the sink. It consumes much more energy than other cluster-member nodes (Dai *et al*., 2009). The routing algorithm implements cluster-head cycle-switching mechanism to keep energy equilibrium of all nodes by adjusting cluster-head nodes and cluster-member nodes.

In spite of the extraordinary progress in the cluster topology, there are still some problems:

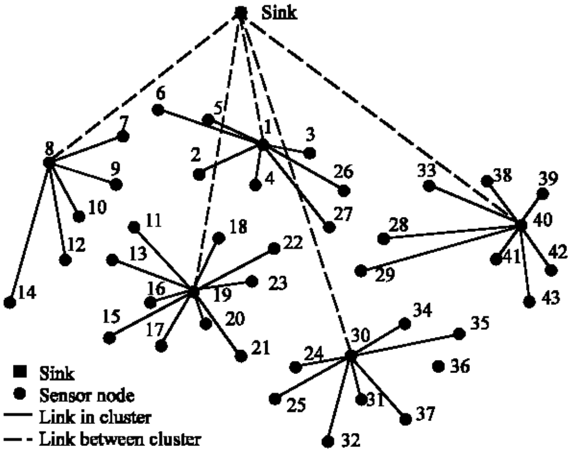

Fig. 1: The cluster topology

- **The cluster-head selection:**In the cluster topology, it needs to focus on cluster-head selection. If the cluster-head selection is unreasonable, it will lead to excessive energy consumption of cluster-head node and then quickly exhaust its energy, finally the network lifetime will become short
- **Data transmission:** In the data communication phase, all cluster-member nodes directly communicate with cluster-head node and then the cluster-head node communicates with the sink. If the communication distance is long, it will need much energy and cause serious waste of the node resources (Lin and Tsai, 2006)
- **The network topology change:** In the network topology, when the cluster-head node consumes much energy and has less residual energy, the

Corresponding Author: Lejiang Guo, Department of Control Science and Engineering,
Huazhong University of Science and Technology, Wuhan Hubei,430074, China

network need to select a new cluster-head node by cluster-head cycle-switching mechanism to maintain energy equilibrium (Dai and Wu, 2004). However, the cluster-head selection needs to consume some energy of the network topology. At the same time, the frequent cluster-head selection and the network construction will waster much energy, so the algorithm must control the change frequency of network topology

This study analyzes the problems existing in the cluster-head network and presents an energy equilibrium routing algorithm based on cluster-head prediction for wireless sensor networks. The algorithm improves the cluster-head selection and lifetime, balances the node energy and prolongs the network lifetime.

NETWORK MODEL

Definition 1: The network structure G(V, E), G is network topology diagram, V is network node set, E is the set of network Two-way edge.

Definition 2: S_N^i is the set of physical neighbor around i, $i \in V$. If $i \in V$ and $i \in S_N^i$, it shows that i and j cover each other, $e(i, j) \in E$. If $i \in S_N^i$ and $j \notin S_N^i$, it shows that j covers i, i does not cover, $\vec{e}(i, j) \notin E$.

Definition 3: If there is a two-way path p(i, j) = {i,...,j} between $i \in V$ and $j \in V$, G is a two-way connectivity network. In wireless sensor networks, the node location information is important to the design of network topology and routing algorithm. The previous design of network protocol considers more RSSI-based algorithm design, but RSSI only reflects the connection relationship between two nodes through Signal strength. It can not reflect the distance and direction between the nodes. With the development of GPS and the node localization algorithm, the position between nodes can be easily obtained. Therefore, the topology control algorithm and routing protocol design process will make full use of node location message (Alzoubi *et al.*, 2002). The nodes are distributed in an ideal two-dimensional or three-dimensional space, the distance between nodes reflects the nodes adjacent tightness instead of RSSI signal strength.

Definition 4: R_i is the transmission range of node I, R_j, is the transmission range of node j, (x_i, y_i, z_i) is i location coordinates, (x_j, y_j, z_j) is j location coordinates. The distance of i and j is:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

$$\text{If } d_{ij} < R_i,\ d_{ij} < R_j, \text{ then } i \in S_N^j \text{ and } i \in S_N^i,\ e(i, j) \in E$$

Because the failure of any node may result in the failure of collection task, it may lead to serious network disconnected. Therefore, the network lifetime is the first node death time in this study.

Definition 5: T_i is the run time-length of the network node i, the network lifetime can be expressed $\min\{T_1, T_2,...,T_n\}$. From definition 5, if extending the network lifetime, it must extend the first node's death time. Suppose that the whole network energy is E_0, the network node number is N. The first death node is T_i, the death time is T_i. When the first node dead, the node j consumes the energy E_j. In order to maximize the network life time extension, it must abide by Eq. 1:

$$E_0 = \sum_{j=1}^{N} E_j \qquad (1)$$

Form Eq. 1, it means that when the first node death, the energy consumption of all the nodes is equal to the network energy as a whole. Only in this way, the network does not appear to waste energy, all nodes are dead when the first node consumes its own energy. Therefore, the balance of the energy between network nodes is an important constraint indicator to extend the network life-time maximum. In the network model, the nodes need to perform different roles and undertake different functions. From Eq. 1, it needs to constantly change network node and the role of the function to balance the energy between network nodes. In wireless sensor networks, data collection is the core task of the network for different applications, the data collection types is also varied, such as temperature, humidity, light and so on. Meanwhile, according to network data collection modes, the data divides into the original data and processed data. When the network collects the original data, the numerous transferring data packets consume much energy. With the data processed collection, the transferring data packet capacity can be greatly reduced. The corresponding energy consumption will be reduced a lot because the sensor nodes implement data fusion with it own limited resources.

In this study, all the sensor nodes can be data fusion which has very important significance for some specific applications, such as acquisition and monitoring area maximum temperature, minimum temperature and average temperature statistics.

THE CLUSTER-HEAD PREDICTION MECHANISM

The network operation mode is that each round has usually a cluster-head node. Each cluster-head node has its cluster-member nodes. The cluster-head prediction mechanism is that a cluster-head node selects from its own neighbor nodes which has much energy in the next rounds when a cluster-head node run (Zhou *et al*., 2008). In the cluster-head prediction mechanism, the assumption that the transmission range of nodes can cover the entire network. Then, the neighbor cluster-head node will be all other nodes in the network.

Definition 6: r is the number of current network rounds execution. The cluster-head predicts the cluster-head node in the next L round. When mod(r, L) = 0, it requires the cluster-head node to implement the cluster-head prediction mechanism.

Definition 7: The set $R = \{r_0, r_1,..., r_n\}$ and mod(ri, L) = 0, $0 \le i \le n$, $r \in R$, the cluster-head node executes the cluster-head prediction algorithm, it calculates the next L cluster-head node, the cluster-head prediction node add to the set $S_{LCH} = \{Ch_i | 0 \le i \le L\}$, CH_0 is current cluster-head node in S_{LCH}.

Definition 8: $E = \phi$ is the residual energy set of predicted node. The number of non-cluster-head node is k. The number of the cluster-head node is n'. The set of the Ordinary nodes is $S_N = \{N_i | 1 \le i \le n'\}$. The cluster-head prediction mechanism includes the following steps:

- **Step 1:** Take a common node $N_i \in S_N$ out, the initial energy is E_i
- **Step 2:** Take a cluster-head node $CH_m (0 \le m \le L)$ out from S_{LCH}
- **Step 3:** Calculate E_{ci} which is the energy consumption of N_i when CH_m is the cluster-head node. Residual energy of N_i is $E_{ri} = E_i - E_{ci}$, then $E_i = E_{ri}$
- **Step 4:** Check the collection of all cluster-head nodes which have been traversed. If the traversal completion, it puts residual energy E_i of the node N_i to E, then enters step5. If the operation is not complete, it returns step 2 and continues to execute when all cluster-head node is traversed in the set S_{LCH}
- **Step 5:** Check whether the traversal completion in S_N, if completion, it enters step 6. Otherwise, it returns step1 and continues to traverse the node in S_N
- **Step 6:** Take E_{max} out from E, a new cluster-head node $N_{max}$is added to S_{LCH}, $S_{LCH} = S_{LCH} \cup \{N_{max}\}$. N_{max} is deleted from the set S_N, it empties the set E, then $S_N = S_N \backslash \{N_{max}\}$, $E = \phi$
- **Step 7:** Judge whether the set S_{LCH} has already L+1 cluster-head node. If it is true, the algorithm stops the cluster-head prediction. Otherwise, CH_0 is deleted from S_{LCH}, $S_{LCH} = S_{LCH} \backslash \{CH_0\}$, it return to step 1
- **Step 8:** The cluster-head predictable node L is put to S_{LCH}. Each predictable node is stored and severed as cluster-head according to the sequence in S_{LCH}. All predictable nodes will serve as cluster-head in the next L round

The energy consumption of the predictable cluster-head node can be described in Eq. 2:

$$E_{ci} = \begin{cases} DE_{elec} + D\varepsilon_{fs} d_{im}^2 + CE_{elec} + CE_{elec} + C\varepsilon_{fs} d_{im}^2 & d_{im} < d_0 \\ DE_{elec} + D\varepsilon_{mp} d_{im}^4 + CE_{elec} + CE_{elec} + C\varepsilon_{mp} d_{im}^2 & d_{im} > d_0 \end{cases} \quad (2)$$

D is the size of data package size, C is the size of control package, d_{im} is the communication distance between N_i and Ch_m. When the common nodes communicate with the cluster-head node in Eq. 2. It firstly must establish a connection and transmit data between nodes. The connection part can be expressed:

$$CE_{elec} + CE_{elec} + C\varepsilon_{fs} d_{im}^2 \text{ or } CE_{elec} + CE_{elec} + C\varepsilon_{mp} d_{im}^4$$

This part includes the energy consumption of requesting or confirming the connection between the common node and the cluster-head node. Energy transmission consumption can be expressed by:

$$DE_{elec} + D\varepsilon_{fs} d_{im}^2 \text{ or } DE_{elec} + D\varepsilon_{mp} d_{im}^4$$

It includes the consumption of sending data package from N_i to CH_m.

THE ENERGY EQUILIBRIUM ROUTING ALGORITHM BASED ON CLUSTER-HEAD PREDICTION FOR WIRELESS SENSOR NETWORKS

This section will present the Energy Equilibrium Routing Protocol based on Cluster-head Prediction for wireless sensor networks (CP-EERP). Particular implementing scheme of the five phases is included in CP-EERP, network initialization phase, cluster building phase, data transmission phase, cluster-head prediction phase. Prediction phase is running only a few rounds to meet certain conditions.

Network initialization phase: In the network initialization phase, the base stations randomly select k cluster-head node from all nodes, these message of cluster-head node is packed and broadcast in the network (Yingshu *et al*., 2005). At the same time, all nodes broadcast a message and wait for the reply message in order to obtain message from their neighbor nodes. In this algorithm, the node transmission range can cover the entire network, so the set of neighbor node is equivalent to the set of all nodes in the network.

Cluster building phase: The steps of cluster building phase are related with the node role and the current round. In the next, the node N_i is used to illustrate the process of cluster building phase. When the round r = 0, the node N_i receives a base station broadcast messages, this message contains k cluster-head node selected randomly and then it determines whether it is a cluster-head node (Guha and Khuller, 1998). If N_i finds it is a cluster-head node, N_i Labels itself as the initial cluster-head node. In a period of time T, it waits for the message JoinMsg from non-cluster-head node to join the cluster. Moreover, the cluster-head node is generated at the time r = 0, when N_i receives the message JoinMsg from non-cluster-head node, the cluster-member nodes are recorded and stored in the set $S_{ini\text{-}imen}$ which is the initial run of the cluster-member list. It is important to the network failure. Conversely, if N_i find that it is not cluster-head node, according to sources in the other cluster-head node location, N_i calculates the distance from their nearest cluster-head node CH and then it send a message JoinMsg and waits for the request message ACK from CH. The message ACK contains the corresponding time slot. The node has successfully joined the cluster when N_i receives ACK. Each new node will join the cluster allocation through a time slot, the time slot is used for data transmission phase. When r>0, the node N_i acquires the cluster-head message not from the message of the base station, but rather from the broadcast message of all cluster-head nodes.

When the network runs, all current cluster-head node execute a cluster-head prediction algorithm in certain conditions, then all predictable cluster-head node broadcast the message in the network. Take current round r for an example, if r meets the condition, it makes cluster-head node execute a cluster-head prediction algorithm, each cluster-head node will predict the cluster-head node in the next L round and broadcast the message. The cluster-head information and each cluster-head rounds are shown in Table 1.

The number of the cluster-head node is k when the network is initialized, the number of the predictable cluster-head node is k and then each cluster-head can predict the cluster-head in the next L round. The number of the cluster-head node is kL after the prediction. All nodes will receive the predictable cluster-head message including ID location. From Fig. 2, the node i will become the cluster-head node in r+1 round. When the network run in r+1 round, the node calculates the distance with the node i, the result determines whether the node i can become its cluster-head node. If it is suitable for the cluster-head node, it sends JoinMsg to the node i. In r+1 round, there is k-1 cluster-head node. When r>0, the cluster-head nodes are acquired by the prediction mechanism, not by competition. And it only needs to wait for other JoinMsg to complete the building cluster process. When all nodes have become a member of the cluster, cluster building phase is finished.

Table 1: The predictable cluster-head node and round

ID	Round
i	r+1
j	r+2
k	r+3
…	…
m	r+L

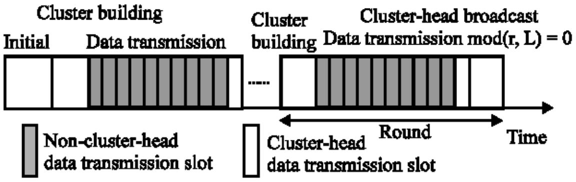

Fig. 2: Time sequence of the algorithm

Data transmission phase: In data transmission phase, the node N_i completes the task of receiving and transmitting data. If N_i is the cluster-head node, N_i will wait for a period of time to receive data packets from the other cluster-member nodes. Each packet is included with the residual energy in the cluster. The node N_i extracts the residual energy from the data package. It extracts other data from the data package for data fusion and sends the package to the base station. If N_i is a common node, the node will send data packets to CH in its own cluster. In data transmission phase, all non-cluster-head nodes have been allocated time slots, each node can send data package in its own slot. This method can avoid all nodes to transmit data which cause network congestion and data conflict.

Cluster-head prediction phase: Cluster-head node prediction phase is a rather exceptional network phase, this phase is performed when the network run to a specific round. It performs this phase when current round r is a multiple of L. By taking a cluster-head node CH_i as an example, it explains the cluster-head predictable phase in detail. When mod(r, L) = 0, CH_i, need to determine weather the number of non-cluster-head node is greater than L in this cluster. If the number of non-cluster-head node is greater than L, Ch_i predicts the cluster-head node in the next L round, puts the predictable cluster-head node into the list and broadcasts the list in the whole network. All other cluster-head nodes or non-cluster-head nodes will receive the broadcast predictable message. They store the predictable message in the next L round (Jie *et al*., 2008). If the number of non-cluster-head is less than L, the cluster-head node will not complete the predictable cluster-head node in the next L round. The cluster-head node will broadcast a restart message. When the initial cluster-head node receives the message, it will broadcast a message, acquire the initial Energy message from the member of $S_{ini\text{-}i\text{-}men}$ and then predicts the cluster-head node in the next L round. The message of the predictable node will broadcast in the network. When mod(r, L)≠0, CH$_i$ does not operate.

Time sequence of the algorithm: The algorithm is divided into five phases, the network initialization phase is needed to configure. When the network run, the whole network usually use cluster building phase, data transmission phase and cluster-head predictable phase. Cluster building phase is that the non-cluster-head sends a request message to join the cluster, the cluster-head node receives a message and sends a confirmation back message. The data transmission phase is that all non cluster-head nodes send packets to the cluster-head node in each time slot, the cluster-head node integrates and sends data packets to the base station. When mod(r, L) = 0, the cluster-head node predicts the next L cluster-head and broadcasts the cluster-head message. Time sequence of the algorithm is shown in Fig. 2.

SIMULATION

Experimental environment is Matlab7.1 (Hahn and Valentine, 2010), simulation parameter is shown in Table 2. We made the experiments from January 2010 to March 2010.We completed this project under the guidance of Professor Wang in control theory laboratory. The Network covers an area of size 100 * 100 m, the node number is 200, the position of the base station is (50,175).

From Fig. 3, with L increases, the energy consumption of average round is downward trend. Therefore, the change of L will not affect the optimal number of cluster-head node which is also consistent with the theoretical analysis.

Set the number of cluster-head to 2, 3, 4, 5. L increases from 1 to 10. The relationship between the average round energy consumption and L is shown in Fig. 4. When the number of cluster-head is fixed, L increases at the same time, the average energy consumption of the network decreases each round. In the circumstances of different number of cluster-head, with each round of L change, the average energy consumption of the network shows downtrend.

Table 2: Simulation parameter

Parameters	Value
E_0	0.5J
E_{elec}	50e-009J/bit
ε_{fs}	10e-012J/bit/m^2
ε_{mp}	1.184e-018J/bit/m^4
E_{DA}	5e-009J/bit/signal
C	200 bits
D	4000 bits

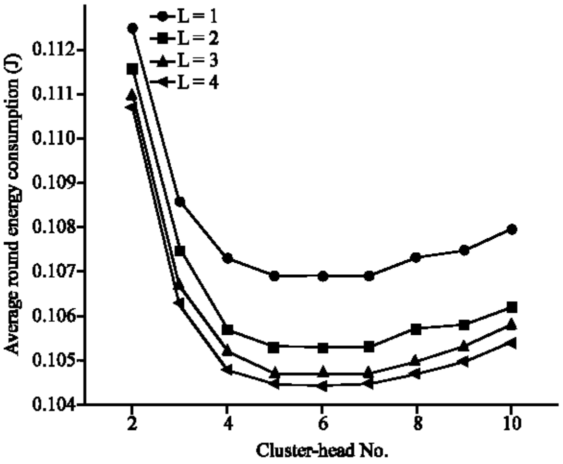

Fig. 3: Simulation of optimal cluster-head number

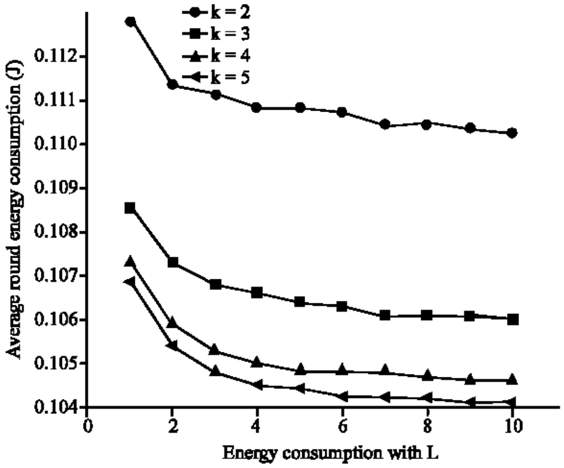

Fig. 4: Energy consumption with L

Compared with CP-EERP algorithm, this paper proposes a complementary method TMP-L.TMP-L and CP-EERP are different in the mechanism of cluster-head prediction. When mod(r, L) = 0, cluster-head prediction mechanism of CP-EERP uses traversal prediction method. If a node serves as a cluster-head node, other nodes predicts the residual energy, the largest residual energy of node is selected as a cluster-head node in the next round. The cluster-head selection mechanism of TMP-L also uses the cluster-head prediction mechanism. However, the

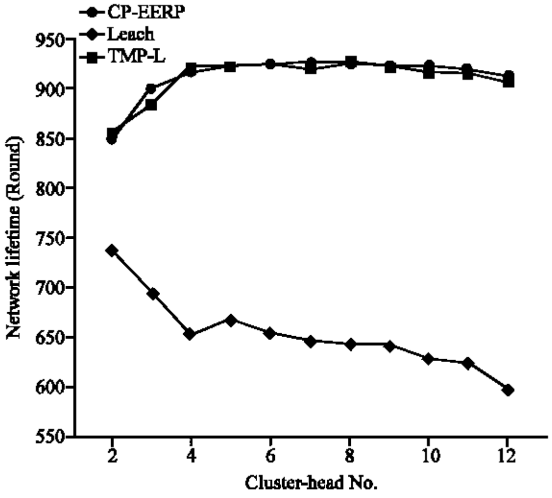

Fig. 5: The simulation of the network lifetime

Table 3: Performance comparison

MSG complexity	Time complexity	Reference
$o(n^2)$	o(log n)	Boukerche and Samarah (2008)
$o(n^2)$	o(n log n)	Zhou *et al.* (2008)
o(n)	o(n)	Li *et al.* (2005)
<o(n)	o(n)	CP-EERP

residual energy is the residual energy in own cluster under the current round. It also select L node as the cluster-head node in the next L round. Exception to the cluster-head selection algorithm CP-EERP, TMP-L is the same as CP-EERP. When the cluster-head number increases from 2 to 12, L is set to 3, the simulation is show in Fig. 5.

The network lifetime of CP-EERP is about 50% higher than the Leach, the highest value is up to 70% to 80%.The lifetime of TMP-L is are similar to CP-EERP, it illustrates that TMP-L is also much better than Leach. CP-EERP is slightly better than TMP-L because CP-EERP is the cluster-head node selection based on residual energy in the next round. TMP-L is the cluster-head node selection based on the current residual energy.

A simulation has been conducted to compare our proposed algorithm with (Boukerche and Samarah, 2008; Zhou *et al.*, 2008; Li *et al.*, 2005). The performance comparison of some reviews algorithms is listed in Table 3.

Message complexity and Time complexity are two important factors to measure the quality of the algorithm. So we can conclude from Table 3 that our approach outperforms the list approach.

CONCLUSION

This study analyzes the characteristics of clustering network structure, proposes a cluster-head prediction algorithm for the network energy management. The mechanism considers the cluster-head selection, data transmission and network topology. It adopts a reasonable method to select the cluster-head node in the network. The algorithm greatly improves the energy balance of the network and prolongs the network lifetime.

ACKNOWLEDGMENTS

This research is supported by two grants from the National Natural Science Foundation of China (No. 60773190, No. 60802002).

REFERENCES

Alzoubi, K., P.J. Wan and O. Frieder, 2002. New distributed algorithm for connected dominating set in wireless ad hoc networks. Proceedings of the 35th Annual Hawaii International Conference on System Sciences, Jan. 7-10, Washington, DC., USA., pp: 297-297.

Boukerche, A. and S. Samarah, 2008. A novel algorithm for mining association rules in wireless ad hoc sensor networks. IEEE Trans. Parallel Distrib. Syst., 19: 865-877.

Dai, F. and J. Wu, 2004. An extended localized algorithms for connected dominating set formation in ad hoc wireless networks. IEEE Trans. Parallel Distributed Syst., 15: 908-920.

Dai, Z., Z. Li, B. Wang and Q. Tang, 2009. An energy-aware cluster-based routing protocol for wireless sensor and actor network. Inform. Technol. J., 8: 1044-1048.

Guha, S. and S. Khuller, 1998. Approximation algorithms for connected dominating sets. Algorithmica, 20: 374-387.

Hahn, B.H. and D.T. Valentine, 2010. Essential MATLAB for Engineers and Scientists. 4th Edn., Elsevier, USA.

Jie, W., D. Fei and Y. Shuhui, 2008. Iterative local solutions for connected dominating sets in ad hoc wireless networks. IEEE Trans. Comput., 57: 702-715.

Li, Y., M.T. Thai, W. Feng, C.W. Yi, P.J. Wan and D.Z. Du, 2005. On greedy construction of connected dominating sets in wireless networks. Wireless Commun. Mobile Comput., 5: 927-932.

Lin, C.H. and M.J. Tsai, 2006. A comment on HEED: A hybrid energy efficient distributed clustering approach for ad hoc sensor networks. IEEE Trans. Mobile Comput., 5003: 1471-1472.

Zhou, K., L. Meng, Z. Xu, G. Li and J. Hua, 2008. A dynamic clustering-based routing algorithm for wireless senor networks. Inform. Technol. J., 7: 694-697.