International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 16 (2017) pp. 5690-5696

Path Cost Based Load and Energy Balanced Clustering & Routing Algorithm

Mohit Kumar
Department of Computer Science and Engineering
Cambridge Institute of Technology
Ranchi, India.

Dr. Dinesh Kumar
School of Computer and Systems Sciences
Jaipur National University
Jaipur, India.

Abstract

Design of efficient clustering and routing algorithm are most important issues for any WSN applications. Most of the existing clustering and routing algorithm suffers from uneven load distribution. In order to prolong the network lifetime energy consumption in routing must also be considered. While deciding the routing path existing load of the gateways should also be considered apart from other factors like residual energy, distance and back transmission. In this paper we have proposed PBLBC(Path cost based load balanced clustering and routing algorithm). we first present the clustering algorithm which is both energy aware and load balanced. In the proposed algorithm the gateway selection is made on the basis of cost function. We have next proposed a routing algorithm in which next hop gateway is selected using path cost function. Simulation result shows the efficacy of the proposed algorithm.

Keywords: Wireless sensor networks; clustering; energy efficiency; load balanced;PBLBC;

INTRODUCTION

Wireless Sensor Networks (WSN's) have widespread use in various fields such as disaster warning, environmental monitoring systems, health sector, safety and surveillance in military & forces, many home applications etc. Due to their enormous usage and widespread potential, WSN's have attracted the attention of many researchers.

A WSN mainly comprises of a large number of tiny nodes/ sensor nodes deployed I selected target area randomly. The nodes are provided with a power unit and components for communication, data sensing & processing. After sensing the local information the sensor node sends the data to base station also known as sink. But the major setback with WSN's is their energy. WSN's operate on low power sources and the energy gets exhausted easily. It is then very difficult or almost impossible to replace the exhausted nodes with new ones. These issues have been studied continuously by the researchers in order to design an efficient WSN.

Clustering Algorithms have been found as a promising are in designing energy efficient WSN's. In a cluster based algorithm the nodes are divided into groups known as clusters, with each cluster having a cluster head (CH). The Cluster Heads from each cluster formed, aggregates the local data and sends it to the base station or sink.

However in many of the WSN's, cluster head selection in done among the normal deployed nodes which as mentioned earlier have low energy sources and die quickly. In order to overcome these issues many researchers have proposed the use of gateways or relay nodes as Cluster Heads [13,14,15,17]. Gateway or Relay nodes are provisioned with extra energy as compared to normal sensor nodes. Since gateways are also battery operated hence there is power constraint with them also. Lives of the gateways are also very crucial in order for long term operation of the network. It is also very important to note that improper cluster formation may overburden some gateways reducing the network lifetime. Hence the cluster formation should be both energy balanced as well as load balanced in order to improve the lifetime of the network. Many researchers have proposed load balancing algorithm based on the number of nodes assigned to each gateways (CHs). However they missed out the fact that residual energy of the gateways should also be considered with load calculation.

In this paper, we propose 'Path Cost based Load Balanced CRA' called PBLBC. The proposed algorithm addresses all the above issues.

The proposed algorithm is a distributed algorithm consisting of two phases namely cluster formation and data routing. The cluster formation of the algorithm is very efficient as it addresses issues like energy consumption, load balancing as well as hot spot problem. The algorithm requires local information like residual energy of gateways, load of the gateways and distance between sensor nodes and gateways. For multi-hop routing the gateways are used as relay nodes for transferring the data to the base station. Therefore, a gateway needs to find best neighbor gateway for data forwarding in-order to minimize the energy consumption. In the proposed algorithm best neighbor gateway is calculated on the basis of path cost which is calculated using factors like residual energy, distance and load of gateways. The simulation result of proposed algorithm is compared with EELBCA[15] and LBC[17]. Result shows that proposed algorithm outperforms other two algorithms in energy consumption and network lifetime.

RELATED WORK

A number of clustering algorithm have been developed for WSN [2],[3]. LEACH [6] is the most popular clustering algorithm for WSNs. It works on the principle of the rotation on CHs among sensor nodes which is useful in balancing of load. However the main drawback of LEACH is CH selection, as it is based on probability, so a node having very less energy might get elected as CH. The other problem with LEACH is that the communication between CHS and base station is done via single hop which is not feasible in a large network. Over the years many improved algorithm have been developed over LEACH like HEED [5], PHGASIS [7], TEEN [12] etc. The PHGASIS algorithm outperforms LEACH in network lifetime but it is also not suitable for large networks. On the other hand TEEN reduces energy consumption but the main drawback with TEEN is that if threshold will not reach the nodes will never communicate and in the event of node failure the user will never know about it.

HEED is a distributed clustering scheme which selects CHS based on residual energy and intra-cluster distance. P.Kuila et al[14] proposed distributed clustering and routing algorithm for sensor networks. It worked well for multi-hop routing. The proposed algorithm was energy balanced but not load balanced. Mounir Arioua et al[10]proposed multi-hop cluster based routing approach based on combination of LEACH and MTE protocols but they also did not considered the load balancing aspect. K.Biswas et al[18] used heuristic measure to establish both energy sufficiency and energy efficiency. Sometimes, the delay in this algorithm is increased, because there is no restriction on next hop selection and many times a node which is in opposite direction of base station are selected as next hop which leads to increased delay. Zhu Yong et al [19] proposed DECSA(Distance energy structure algorithm) which considers residual energy information as well as distance of nodes. But while data forwarding in multi-path routing is doesn't consider the existing load of the CHs.

Most of the above algorithms doesn't consider clustering and routing issues together. Some researchers have addressed there issues together but their algorithm are not load balanced one. The proposed algorithm in this article addresses clustering and routing issues together. Apart from that it is also load balanced.

NETWORK MODEL AND TERMINOLOGIES

We assume that all the sensor nodes and gateways are deployed randomly. The sensor nodes and gateways becomes stationary after their deployment. The sensor nodes can only be assigned to that gateway that falls within the communication range of the sensor node. Hence there are certain gateways to which a specified sensor node can be assigned. Therefore each sensor node has a set of gateways to which it can be assigned and it can be assigned to only one gateway amongst them. Each round [6] can be defined as complete period of data gathering and transmission of aggregated data from all the relay nodes to the sink.. The lifetime of the network is measured as number of rounds until the death of first gateway. All the communication is performed over wireless link and a link is set up between sensor nodes only if they are within each other's communication range. We use following terminologies and notations in the proposed algorithm.

- The set of sensor nodes is denoted by S = $\{S_1, S_2 \ldots . S_n\}$
- $\Psi = \{g_1, g_2 \ldots \ldots . g_m\}$ denotes the set of gateways where n>m
- $\mathrm{dist}(g_j, s_i)$ denotes the Euclidian distance between sensor node s_i and gateway g_j.
- $E_{Residual}(s_i)$ denotes the residual energy of sensor node s_i
- dist (g_j, BS) denotes the Euclidian distance between the gateway g_j and the BS.
- $¥_i$ denotes the traffic load contributed by nodes S_i to gateway G_s, $S_i \in$ S, $G_s \in \Psi$ and $¥_i \in$ Z where Z is the set of rational number.
- Comm(s_i) is the set of all the gateways, which are within the communication range of s_i and s_i can be assigned to any one of them, where $s_i \in$ S.
- Comm(g_i) is the set of all gateways including sink that lies within the communication range of g_i. g_i can relay data to only that gateway which are within the communication range of g_i where $g_i \in \Psi$.
- Let $Ł_s$ be the total load assigned to gateway G_s. In other words

  $Ł_s = \sum_{i=1}^{n} ¥_i \times \beta_{ij}$

  where β_{ij} is the Boolean variable

  $$\beta_{ij} = \begin{matrix} 1 \; if\ sensor\ node\ s_i\ is\ assigned\ to\ gateway\ g_j \\ 0 \; otherwise \end{matrix}$$

- $dist(G_r, G_s)$ denotes the Euclidian distance between gateway G_r and G_s.
- $E_{Resi}(G_s)$ denotes the residual energy of the gateway (G_s)
- dist (g_j, BS) denotes the Euclidian distance between gateway g_j and BS.

ENERGY MODEL

We have used the first order radio model [1] for estimating the energy consumption by the sensor nodes. According to this model if the distance between transmitter and receiver is smaller than the crossover distance (d_0) then free-space model is used otherwise multi-path model is used. The energy consumed to send *l*- bit message by node i over a distance d is calculated using eq(1)

$$E_{tx}(l,d) = \begin{cases} l\alpha_{tx} + l\varepsilon_{fs}d^2 & for\ d < d_0 \\ l\alpha_{tx} + l\varepsilon_{mp}d^4 & for\ d \geq d_0 \end{cases} \quad (1)$$

Where α_{tx} and α_{rx} are per bit energy dissipation in transmitting or receiving circuitary and ε_{fs} and ε_{mp} are the energy expended for free-space and multi-path respectively. The energy expended to receive *l*- bit message is given by

$$E_{RX}(l)=l\alpha_{rx} \quad (2)$$

PROPOSED PROTOCOL

The basic idea of the proposed algorithm is to improve the network lifetime through an efficient clustering algorithm. The proposed clustering algorithm is both energy aware and load balanced. The routing algorithm used is also energy balanced as well as load balanced. The proposed algorithm takes care of both issues energy consumption and load balancing together. Before clustering, all the sensor nodes as well as gateways are assigned unique ids. This process is known as bootstrapping. The sensor nodes then broadcast their Ids using CSMA/CA MAC layer protocol. The respective gateway that are within the communication of these sensors on receiving the sensor Ids transfer it to the base station along with other local network information. The base station then performs the clustering.

The sensor nodes using CSMA/CA MAC layer protocol, broadcast their Ids. The gateways that lie within the communication range of these sensors receive the Ids. It then sends the Ids along with other local network information to the base station. The base station then performs the clustering based on the received information. The network setup is performed in two phases namely

- Cluster formation
- Multi hop routing

Cluster formation is a process in which sensor nodes are assigned to the gateways. In our algorithm, cluster formation process is both energy balanced and load balanced. Each sensor node in a cluster senses and captures the local information and sends it to respective gateways. The gateways then aggregates all the information received by the sensors that are within its range and sends it to the base station via a single hop or multi hop depending on the distance between the base station and gateways.

In multi hop communication the best relay node is chosen. This decision of choosing the best node is based on certain factors. In our proposed algorithm we have taken into account two factors energy balancing and load balancing

A : Cluster Formation

In order to form clusters each sensor node decides its cluster membership on the basis of some cost functions. Let GS_Cost(g_j,s_i) be the cost of gateway selection g_j for sensor node s_i. The cost function is calculated based on few factors which are as follows.

1) A sensor node should be assigned to gateway that is within its communication range and have higher residual energy than other gateways in the set comm.(s_i). Hence

$$GS\_Cost(g_j,s_i) \propto E_{Residual}(s_i) \quad (5.1)$$

2) A sensor node should join that gateway which is nearest to it, so that it can save its energy which is consumed during communication.

$$GS\_Cost(g_j,s_i) \propto \frac{1}{dist(g_j,s_i)} \quad (5.2)$$

3) The gateways which are near to sink always acts as a relay node in routing data to the sink. Thus increased workload of such gateways may lead to their death quickly compared to other gateways leading to hot-spot problem. So, in order to counter such problem the number of gateways must be less than the number of nodes assigned to any other gateway which are far from sink. Therefore the cost of selection of gateways near to base station should be higher than the cost of selection of gateways which are far away from base station.

$$GS\_Cost(g_j,s_i) \propto dist\ (g_j, BS) \quad (5.3)$$

4) Before joining any gateway a sensor node must also consider the load of the gateways. A sensor node should join the gateway which has minimum load from all the gateways within its communication range.

$$GS\_Cost(g_j,s_i) \propto \frac{1}{Ł_s} \quad (5.4)$$

Combining all the above conditions , implies that

$$GS\_Cost(g_j,s_i) \propto \frac{E_{Residual}(s_i) \times dist\ (g_j,\boldsymbol{BS})}{dist(g_j,s_i) \times Ł_s}$$

ie GS_Cost(g_j,s_i) $= C_1 \times \frac{E_{Residual}(s_i)\times \text{dist}\,(g_j,\boldsymbol{BS})}{dist(g_j,s_i)\times Ł_s}$ where C_1 is a proportionality constant. Without loss of generality we assume that $C_1 = 1$. Therefore,

$$\text{GS\_Cost}(g_j,s_i) = \frac{E_{Residual}(s_i)\times \text{dist}\,(g_j,\boldsymbol{BS})}{dist(g_j,s_i)\times Ł_s} \quad (5.5)$$

Each sensor node will select the gateway which has maximum cost. The gateway selection cost for each gateway that is within the communication range of $s_i$will be calculated using equation 4.5. After that the gateway with maximum cost will be selected and sensor node s_i will join that gateway. The sensor node sends request to join message to that gateway by using non-persistent CSMA MAC protocol. All the communication that takes place inside a cluster is of single hop. In order to avoid collision each gateway sends TDMA schedule to its group members which ensures collision free communication between sensor nodes and gateways.

Algorithm PBLBC

Input: (1) A set of sensor nodes S={S_1,S_2....................S_n} with traffic load $Load_i$ ∀i, 1≤i≤n and set of gateways (cluster heads) ψ = {g_1,g_2..................g_m}

(2) Residual energy ie $E_{Residual}(g_i)$

(3) Euclidian Distance ie $\text{dist}(\text{g}_\text{j}, \text{s}_\text{i})$

Output: An assignment S→ψ such that energy consumption and load are balanced during clustering.

Step 1 while (R_{set}≠ NULL) do

Assign successive s_i to their corresponding g_j such that $s_i \in R_{set}$ then delete s_i from R_{set} and S.

End while

Step 2 Sort the sensor nodes from U_{set} in non-decreasing order on the number of possible gateways upon which a sensor node can be assigned.

Step 3 While (U_{set}≠ NULL) do

3.1 Pick the first node from U_{set} say s_i

3.2 Assign s_i to g_j such that

GS_Cost(g_j,s_i) = MAX{ GS_Cost(g_j,s_i) | ∀g_j ∊ Comm(s_i)

3.3 Delete s_i from U_{set} and S

End while

Step 4 Stop

B: Next Hop Gateway Selection

After sensing the local information, the gateway aggregates this data and sends it to the base station either directly or through other gateways by using them as relay nodes. In the proposed algorithm the next hop gateway is selected on the basis of calculated path cost. The gateway G_r calculates the path costs from itself of all the gateways that falls within its communication range. The path cost is calculated on some factors which are as follows

1) In order to minimize communication energy the next hop gateway should be at nearest distance.Hence

$$\text{Path\_Cost}(G_r,G_s) \propto \frac{1}{dist(G_r,G_s)} \quad (5.6)$$

2) The next hop gateway must have sufficient residual energy.

$$\text{Path\_Cost}(G_r,G_s) \propto E_{Resi}(G_s) \quad (5.7)$$

3) Ultimately the data needs to be transferred to the base station, so the distance of the next hop G_s with the BS should also be considered while evaluating the path cost.

$$\text{Path\_Cost}(G_r,G_s) \propto \frac{1}{dist(G_{s,BS})} \quad (5.8)$$

4) The next hop gateway should be selected in such a way that its most lightly loaded gateway. Therefore,

$$\text{Path\_Cost}(G_r,G_s) \propto \frac{1}{Ł_s} \quad (5.9)$$

Combining the above conditions,

$$\text{Path\_Cost}(G_r,G_s) \propto \frac{E_{Resi}(G_s)}{dist(G_r,G_s)\times dist(G_{s,BS})\times Ł_s}$$

ie Path_Cost(G_r,G_s) $= C_2 \times \frac{E_{Resi}(G_s)}{dist(G_r,G_s)\times dist(G_{s,BS})\times Ł_s}$ where C_2 is a proportionality constant. Without loss of generality we assume that $C_2 = 1$. Therefore,

$$\text{Path\_Cost}(G_r,G_s) = \frac{E_{Resi}(G_s)}{dist(G_r,G_s)\times dist(G_{s,BS})\times Ł_s} \quad (5.10)$$

The gateway with maximum path cost is selected as the best candidate and all the data is forwarded through that gateway. The path costs are calculated using equation 4.10. The highest path cost will be selected as next hop leading to energy balanced network. Beside being energy efficient the proposed algorithm also considers the load balancing issues. In case if the path costs of all the gateways from G_r is smaller than the path cost of sink from G_r then the data is forwarded directly from G_r to the base station.

Algorithm: Next_hop Selection

Input (1): Residual energy of all the Gateways in the communication range of Gateway(G_r)

(2) Load of all the Gateways in the communication range of Gateway(G_r)

(3) Distance of all the Gateways in the communication range of Gateway(G_r) and distance of G_r to sink

Output : Next_hop Gateway of G_r

Step(1) Next_hop = Sink

Step(2) R = Neighbour(G_i)

Step(3) While(R≠ NULL)

Select Gateway G_j from R

If(Path_Cost(G_r,G_s) > Path_Cost(G_r,Sink) then

Next_hop Relay = G_s

Endif

R=R-G_s

Endwhile

Step(4) Stop

EXPERIMENTAL RESULT

For simulation purpose we have used MATLAB(R2016b). In the simulation environment we have deployed 100-400 sensor nodes and 10-25 gateways in the field size of 500×500. The base station is positioned at the centre ie 250×250.The initial energy of sensor nodes and gateways is 2J and 5J respectively. The radio characteristics and parameters used in the simulation along with their values are given in Table 1. We have used two metric to compare our algorithm with other algorithms.

Standard deviation of load per cluster- We measure the deviation of load after completion of each round. Standard deviation is a very good metric for evaluating the distribution of load per cluster. In our experimental set-up we have run our simulation by varying the number of sensor nodes rom 100-400 and the number of gateways from 10-25. In this case ,smaller variance signifies even distribution of load among gateways. The result shown in Fig 1 clearly shows that our approach clearly outperforms the other algorithms.

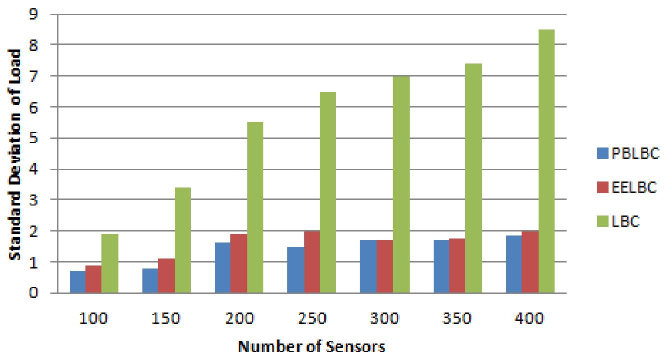

Figure 1: Comparison in terms of standard deviation of load

Network lifetime – The network lifetime here is assumed as the number of rounds till the first gateway dies. We have compared our algorithm with LBC and EELBC on the basis of network lifetime. It can be observed from Fig-3 that PBLBC outperforms LBC and EELBC.

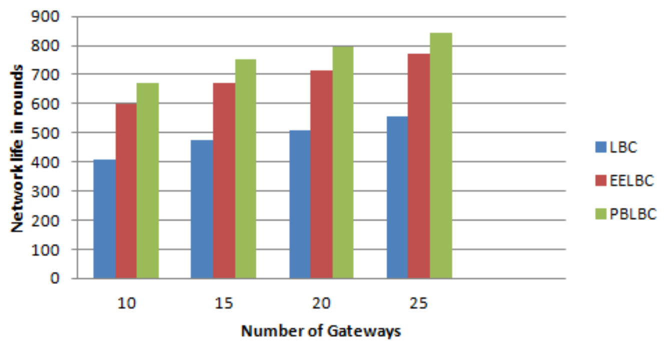

Figure 2: Comparison based on network lifetime for 10-25 gateways

Parameters	Values
Network Area	500×500
Position of BS	(250,250)
No. of Senor Nodes	100
No. Of Gateways	10
Communication Range (sensor nodes)	100m
Communication Range (Gateways)	125m

Initial Energy of sensor nodes	2J
Initial Energy of Gateways	5J
Node Distribution	Random
Transmission Energy	50nJ/bit
Energy amplification(ε_{fs}) (for free space)	100pJ/bit/m^2
Energy amplification(ε_{mp}) (for multi-path)	0.0013pJ/bit/m^4
Data Aggregation Energy	5nJ/bit
d_0(Crossover distance)	87.0m
Packet size	4000 bits
Message size	200 bits

CONCLUSION

In this paper we have presented an energy and load balanced clustering and routing algorithm for wireless sensor networks. The algorithm uses less energy constrained nodes, gateways as Cluster heads. In cluster formation phase all non cluster heads are assigned to gateways within their communication range, considering gateway selection cost. The proposed algorithm is energy balanced as well as load balanced. In multi hop routing phase the best gateway is selected based on path cost. Each gateway calculates the path cost of only those gateways that fall in its communication range. Experimentation result shows that the proposed algorithm effectively balances the energy consumption and it is also more efficient in respect of network lifetime.

Our algorithm works well with static environment i.e. node with no mobility. Our future endeavour will be to design algorithm with consideration of node mobility. The proposed algorithm effectively balances the energy consumption.

REFERENCES

[1] Heinzelman et al,"Application specific protocol architecture for wireless microsensor networks", IEEE Transaction on wireless communications, pp 660-670,2002

[2] I.F.Akyildiz, W.Su, Y.Sankarasubramanium and E.Cayirci,"Wireless sensor networks:survey",Computer Network, vol 38, pp393-422, March 2002.

[3] AA.Abbasi, M.Younis."A survey on clustering algorithms for wireless sensor networks",Computer Communiaction,pp 26-41,2007

[4] W.B.Heinzelman, A Chandrakasan and H.Balakrishnan,"Energy efficient communication protocol for wireless microsensor networks",In proceeding of Hawaii International conference on System Sciences,Vol 8,January 2000.

[5] O.Younis and S.Fahmy,"HEED: A hybrid energy-efficient distributed clustering approach for ad-hoc sensor networks",IEEE Transaction on Mobile Computing,Vol 3,pp 366-379,November 2004.

[6] W.B.Heinzelman, A Chandrakasan and H.Balakrishnan,"Energy efficient communication protocol for wireless microsensor networks",In proceeding of Hawaii International conference on System Sciences,Vol 8,January 2000.

[7] Stephanie Lindsey et al,"PEGASIS:power efficient gathering in sensor information systems", IEEE Tranactions on Parallel and Distributed Systems, Vol.13, No. 9,pp 924-1130,2002.

[8] P.Kuila and Prasanta K.Jana," Heap and Parameter based load balanced clustering algorithm for wireless sensor networks",International Journal of Communication Networks and Distributed Systems,Vol-14 No-4,2015

[9] N.Dimokas,D.Katsaros and Y.Manolopoulos," Energy efficient distributed clustering in wireless sensor networks",Journal of Parallel and Distributed Computing, Vol.70, pp 371-383 April 2010.

[10] M.Ariouna,Y.Assari,I.Ez-zazi,A.Oualkadi,"Multi-hop cluster based routing approach for wireless sensor networks", International conference on Ambient Systems,Networks and Technologies,Vol.83.pp.584-591,May,2016

[11] S.Lindsay and C.Ragavaendra,"PEGASIS: Power-efficient gathering in sensor information system",In proceeding of IEEE Aerospace Conference Vol.3 pp 1125-1130, March 2002.

[12] A.Majeshwar and D.Agarwal,"TEEN: A routing protocol for enhanced efficiency in wireless sensor networks",In proceeding of International Parallel and Distributed Processing Symposium.pp 2009-2015,April 2001.

[13] A.Bari, S.Wazed, A.Jaekel, S.Bandhopadhyay." A genetic algorithm based approach for energy efficient routing in two-tiered sensor networks",Ad-hoc Networks,pp.665-676,2009.

[14] K.Pratyay, S.K.Gupta , P.K.Jana," A novel evolutionary approch for load balanced clustering problem for wireless sensor networks",Swarm Evol.Computing,pp.48-56,2013.

[15] K.Pratyay, P.K.Jana," Heap and parameter based load balanced clustering algorithm for wireless sensor networks",International journal of communication networks and distributed systems,Vol-14.No-4,Jan 2015

[16] P.Kuila and Prasanta K.Jana,"An energy balanced

distributed clustering and routing algorithm for wireless sensor networks", IEEE International Conference on Parallel, Distributed and Grid Computing,2012

[17] G.Gupta and M.Younis,"Load balanced clustering of wireless sensor networks",ICC-IEEE International Conference,Vol.3,pp.1848-1852,2002

[18] Biswas, K., Muthukkumarasamy, V., Sithirasenan, E., & Singh, K. . An energy aware heuristic-based routing protocol in Wireless Sensor Networks. In Computer and Information Technology (ICCIT), 2014 17th International Conference *on* pp. 53-58,Dec 2014.

[19] Z.Yong and Q.Pei," A Energy-Efficient Clustering Routing Algorithm Based on distance and residual energy for wireless sensor networks", International Workshop on Information and Electronics Engineering (IWIEE),2012.