Virtual grid based clustering algorithm for wireless sensor networks

Anaram Yaghoobi
Islamic Azad University, Rasht Branch
Faculty of Engineering
Rasht, Iran
AnaramYaghoobi@iaurasht.ac.ir

Farnaz Hoseini
Young Researchers and Elite Club, Ardabil Branch
Islamic Azad University, Ardabil, IRAN
Farnazhoseini@iaurasht.ac.ir

Gholamhossein Ekbatanifard
Assistant Professor of Computer Engineering
Faculty of Engineering
Islamic Azad University, Lahijan Branch
ekbatanifard@liau.ac.ir

Abstract—**clustering is a fundamental improvement technique in Wireless Sensor Networks where the bandwidth and energy are significantly more limited compared to other networks. These constraints need to *be considered and they require the innovative designs.* We propose an algorithm to collect data from farm land and report them to the sink node for automatic irrigation system which works based on the soil moisture and weather temperature in each part of the farm. This system will help to increase the amount and quality of products. Due to the ultra large size of the farm it is cheaper and simpler to distribute nodes by plane in there. On the other hand collecting node's data will be a big trouble. If we use small and cheap nodes, then we will not need to collect them or the collection process can be done every five or ten years. We have designed a clustering protocol for this purpose because this system should work for several months. We used the virtual grid and a very simple localization algorithm. The results prove the efficiency of many cases like energy-efficiency (55.48% less than LEACH), Normalized routing load (64.73% less than LEACH), Throughput (51.49% less than LEACH) and delay (17.69% better than LEACH).**

Keywords-component: Irrigation system, Sensor node, Virtual grid, Localization.

I. Introduction of WSN

In a Wireless Sensor Network, the clustering is a fundamental improvement technique in which the bandwidth and energy are significantly more limited compared to other networks. These constraints need to be considered and they require the innovative designs. Clustering algorithms appeared from adhoc networks, which is inspired by wired networks, such as the Internet. LEACH[1] is the first typical clustering protocol designed for wireless sensor networks. Because it is the first mature algorithm for cluster formation, LEACH becomes a baseline for successors. Since the birth of LEACH, more and more clustering algorithms has been proposed and show their advantages in energy efficiency and node management compared with flat routing protocols, such as directed diffusion[2]. HEED[3] is another classical clustering approach proposed in 2004, to achieve better load balance, HEED considers the residual energy as the criterion to choose cluster-heads. On the other hand, HEED is too strict to destroy the randomness of the algorithm, which might lead to worse energy efficiency. In order to realize more load balancing, Cao presents a new adaptive back off strategy in[4] to not only realize load balance among sensor node, but also ensure that the elected cluster-heads are evenly distributed. The Max-Min d-cluster algorithm proposed in[5] generates d-hop clusters which can achieve better load balance and generate fewer clusters than early algorithms. But this algorithm does not ensure that the total energy consumption is minimized. Moreover, the Max-Min d-cluster algorithm suffers high implement complexity, making it unsuitable for resource limited sensor networks. Assuming that sensors are distributed according to a homogeneous spatial Poisson process, Bandyopadhyay et al[3] first introduced a multi-hop hierarchical clustering algorithm for wireless sensor networks, they prove that the algorithm can work better than previous algorithms in multi-hop networks. But the random property of the algorithm, a lot of forced cluster-heads will appear in many cases, i.e., the node has no cluster to join and has to communicate directly with BS, which is often located far away from the node's vicinity. The resulting long-range transmission to the BS is very energy demanding. Also the residual battery energy has not been considered in their algorithm. The rest of this paper is organized as follows. Section 2 includes a detailed survey of the related research. The proposed algorithm is discussed in section 3. Section 4 discusses the simulation and its results Finally, Section 5 concludes the paper.

II. Related Works

The various existing clustering algorithms can be divided into following categories:

A. Aggregating Data

Data aggregation basically means that the data is collected from different sensors; some fusion or computation is done on the data, and communicate the computed data to the sink node, thus reducing the possibility of transmission of supererogatory data[6]. Usually, the main purpose of the data aggregation approach is to minimize the consumption of the node's energy as far as possible and thus prolonging the lifetime of the network. There are 2 methods defined of data conglomeration for WSN. Cluster-heads gather sensed data from their candidate members; perform operations (if required) before transmitting to the base station.

B. *Rotating the CH's role*

As the cluster-head's responsibility is to collect data from its candidate members, and communicates it to the sink node, thus energy expenditure of the cluster-head is much more as compared to non-cluster-head nodes. Thus the amount of energy consumption per the cluster-head can be reduced by circulating the charge from one node to another, time to time and thereby causing the premature death of the cluster-heads

C. *Localization of cluster-head at the cluster center*

Besides equilibrating the cluster size[7], circulating the CH responsibility[8-9], another fashion of energy efficiency is employing the cluster-heads in the middle of the cluster[10-11]. In n-hop cluster, the candidate node transmits the data to cluster-head via nodes in the path. If the cluster-head will be situated in the mediate of the cluster, then sending the data to the head will require fewer hops. Thus, energy preservation can be accomplished

D. *Designating node with lowest power needed*

For ceasing energy dissipation, the fashion that can be adapted is to depute the smallest energy required for communication with the candidate nodes in the cluster, and gateways with the energy required for communication with other cluster's gateways.

E. *Equalizing the size of cluster*

Energy expenditure can be equilibrated by circulation the charge of the cluster-head. Another way of equilibrating the energy consumption is by dividing the load among clusters itself. I.e. attempt is made sectionals the network into clusters consisting of equal number of nodes.

F. *Optimizing path based on cluster's power level*

Among the existing clustering algorithms, most of them presume uniform deployment of the sensors in WSN. Nonetheless, it's not always true. It is possible that the concentration of nodes varies in different areas of the network. Within an area, a node can communicate with another node[12] using the low power of transmission. However, using that (the low power) of transmission, they may not communicate with the other area nodes. Thus the nodes can be unionized in clusters based on power level. This leads to energy preservation

III. Proposed Algorithm

Our proposed algorithm is designed for automatic irrigation of large-scale farms. In this environment, nodes can measure soil moisture and temperature of the weather and report them to the sink. The Sink node evaluates the reported values and irrigate according to soil moisture and temperature in each region automatically. These nodes only have a few months of work and distributed in the farm randomly. The nodes are not collected. So that have to be very cheap. For this reason, they need to energy efficient protocols. We use four secondary sinks and a main sink. Nodes report their sensed values to the nearest sink.

Main sink needs the coordinates of each reign to exact irrigation. Thus, nodes should have their coordinates. The nodes should have a GPS or a localization algorithm. Due to the low cost of nodes, the nodes cannot be equipped with GPS. But by placing the sink in the grid and specific distances, we can do localization simply. First, we explain the localization algorithm and then clustering algorithm.

A. *Localization*

The main sink node is in the middle of the farm and origin of a coordinate system. The coordinate axes are in the geographical directions. Secondary sinks are on the bisectors of the coordinate axes. Then sinks broadcast a message to the entire network. Nodes received these messages and calculate distance to each sink. Nodes to calculate their coordinates only need distances to three specific points. We use the (Equation 1) to find the coordinates of nodes.

$$\begin{aligned} D_1^2 &= (X - X_1)^2 + (Y - Y_1)^2 \\ D_2^2 &= (X - X_2)^2 + (Y - Y_2)^2 \\ D_3^2 &= (X - X_3)^2 + (Y - Y_3)^2 \end{aligned} \qquad (1)$$

In (Formula 1) D1, D2 and D3 are distances from sinks. (x1,y1), (x2,y2) and (x3,y3) are coordinates of sinks. (x, y) is the coordinates of node.

B. *Clustering protocol*

We use LEACH clustering protocol that is optimized. First, we broaden virtual grid on the network. This is done by the main sink node. The main sink broadcast a message that contains an upper left angle, length and width of the farm and the length and width of each cell. The nodes calculate the cell that they are placed in it. In each cell there is just one cluster. Every node transmits hello message to the nodes in its cell and other nodes buffer the Id of transmitter if it is placed in its cell. Nodes search in their buffer for the lowest Id. The node with the lowest Id is the cluster-head of the first round and other node sort the list in the buffer in ascending order. The Id of the node itself is included in the list too. Nodes schedule themselves according to the sorted list to send their data to cluster-head. The residual energy of the node is also included in sent data. Cluster-head buffer the reported values and at the end of each round calculate the average of reported soil moisture and temperature of the weather. Cluster-head elects the node with the most residual energy as cluster-head of next round. We make our algorithm fault most tolerant without introducing any overhead to the network. Nodes do not clear their sensed data until the cluster-head sends next round cluster-head election packet. If a cluster-head does not send this packet, it means that the cluster-head has died. So nodes again search buffer for new cluster-heads with lowest Id. Compare between LEACH algorithm (in table1) and improved method are listed in table 2.

Table1: LEACH algorithm

LEACH

In each cell there is just one cluster.

- ✓ Set-up phase: each node in the whole base, generates a random number between [0,1].

*Compares this number with the threshold value.

*If this random number is less than threshold, the node is selected as a cluster-head.

- ✓ Steady-state phase: all members of the cluster send their data to own cluster-head in single-hop method.

Table2: Improved method

IMPROVED

- ✓ Use localization algorithm for decrease cost of nodes.
- ✓ Every node transmits hello message (include id's node) to the other sensors.

*Other nodes buffer the Id of transmitter.

*Nodes search in their buffer for the lowest Id.

- ✓ The node with the lowest Id is the cluster-head of the first round.

*Nodes schedule themselves according to the sorted list to send their data to cluster-head.

*Cluster-head buffer the reported values and at the end of each round calculate the data.

- ✓ Cluster-head elects the node with the most residual energy as cluster-head of next round.

* Nodes do not clear their sensed data until the cluster-head sends next round cluster-head election packet.

*If a cluster-head does not send this packet, it means that the cluster-head has died.

So nodes again search buffer for new cluster-heads with lowest Id.

IV. SIMULATION AND RESULTS

A. *Scenarios*

We have evaluated the proposed algorithm by using the NS2 simulator and have compared it with leach and qleach[13]. Three different scenarios have been done. Scenarios are different in number of nodes. Simulations conditions are listed in table 3.

Table 3: Simulation conditions

Parameters	**Scenarios**		
	First	**Second**	**Third**
Scenery Size	100×100	100×100	100×100
Antenna	Omni Directional	Omni Directional	Omni Directional
Simulation Time	1500	1500	1500
Sink Number	1	1	1
Sub sink Number	4	4	4
Common Node Number	100	150	200
Common Node Location	Random	Random	Random
Initial Energy of Common Nodes	1.0 Joule	1.0 Joule	1.0 Joule

B. *Energy consumption*

The energy consumption of each scenario is calculated. Fig 1, fig 2 and fig 3 show the energy consumption for scenarios with 100, 150 and 200 common nodes respectively. The results prove the efficiency of the proposed algorithm. This algorithm consumes 55.48% less energy averagely.

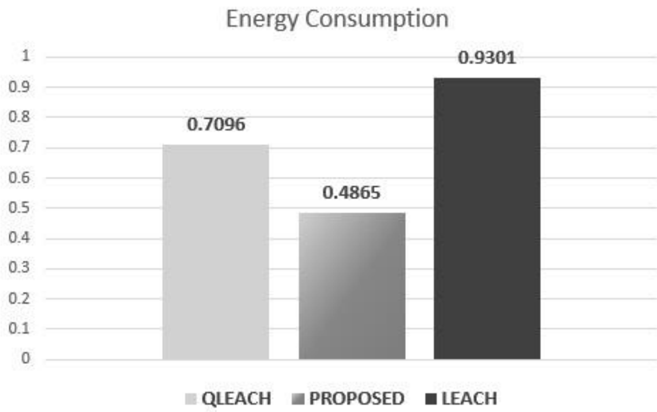

Fig 1: Energy consumption of scenario with 100 common nodes

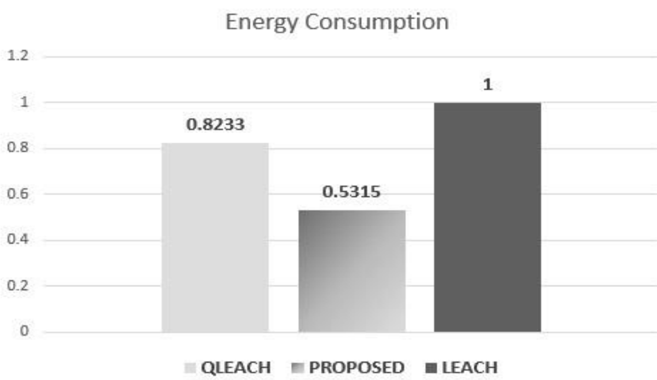

Fig 2: Energy consumption of scenario with 150 common nodes

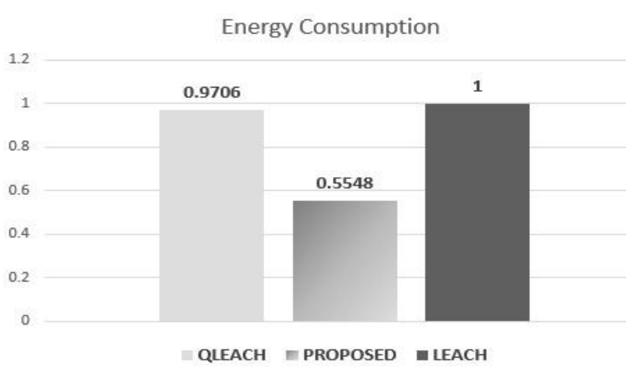

Fig 3: Energy consumption of scenario with 200 common nodes

C. Normalized routing load

The NRL of each scenario is calculated. Fig 4, fig 5 and fig 6 show the NRL for scenarios with 100, 150 and 200 common nodes respectively. The results prove the efficiency of the proposed algorithm. This algorithm has 64.73% less NRL averagely.

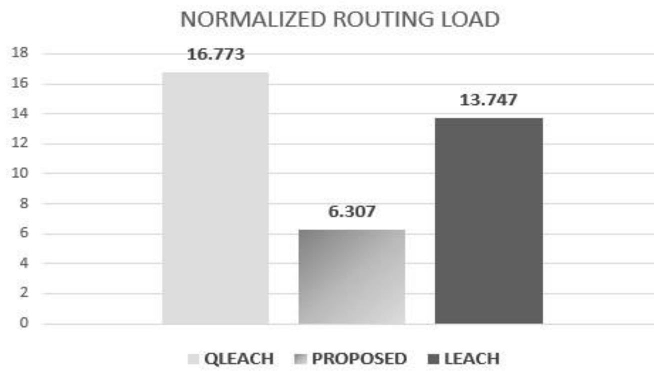

Fig 4: NRL of scenario with 100 common nodes

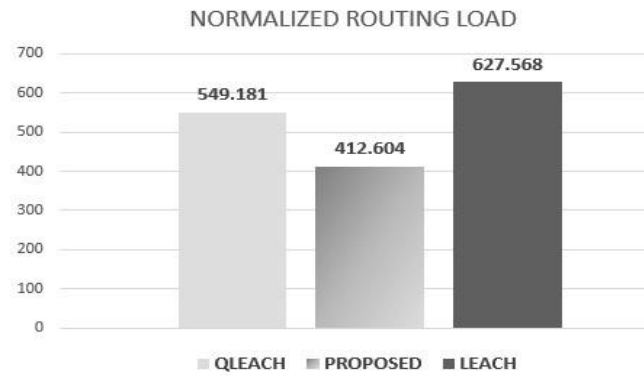

Fig 5: NRL of scenario with 150 common nodes

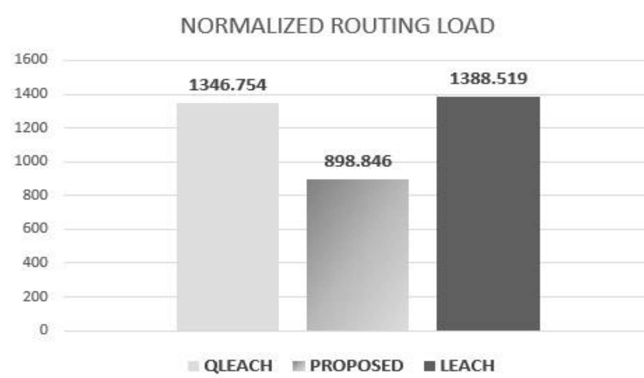

Fig 6: NRL of scenario with 200 common nodes

D. Throughput

The throughput of each scenario is calculated. Fig 7, fig 8 and fig 9 show the throughput for scenarios with 100, 150 and 200 common nodes respectively. The results prove the efficiency of the proposed algorithm. This algorithm has 51.49% less throughput averagely.

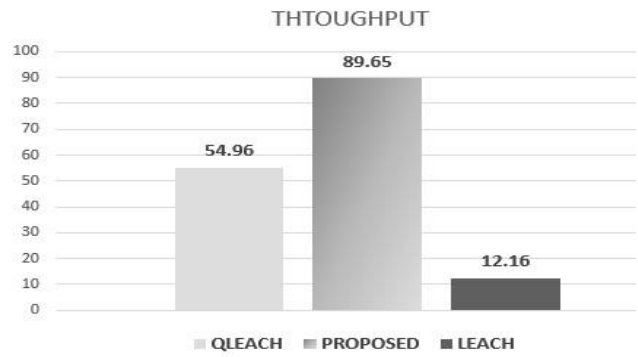

Fig 7: throughput of scenario with 100 common nodes

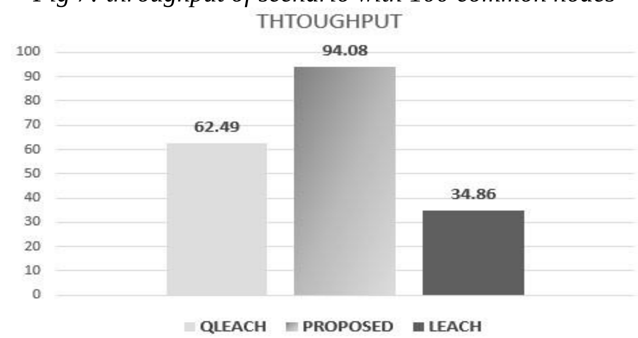

Fig 8: throughput of scenario with 150 common nodes

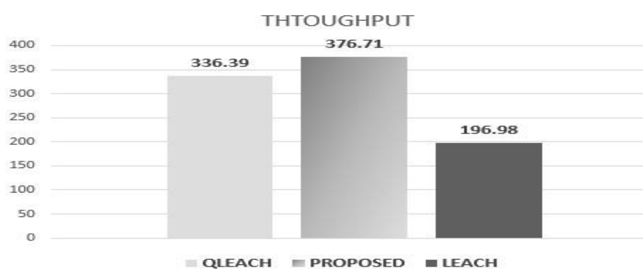

Fig 9: throughput of scenario with 200 common nodes

E. Delay

The delay of each scenario is calculated. Fig 10, fig 11 and fig 12 show the delay for scenarios with 100, 150, and 200 common nodes respectively. The results prove the efficiency of the proposed algorithm. This algorithm has 17.69% less delay averagely..

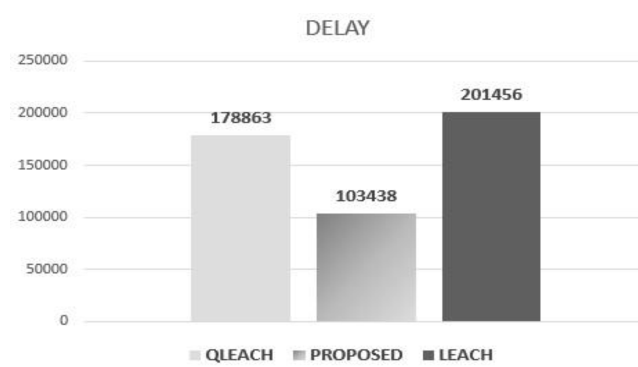

Fig 10: delay of scenario with 100 common nodes

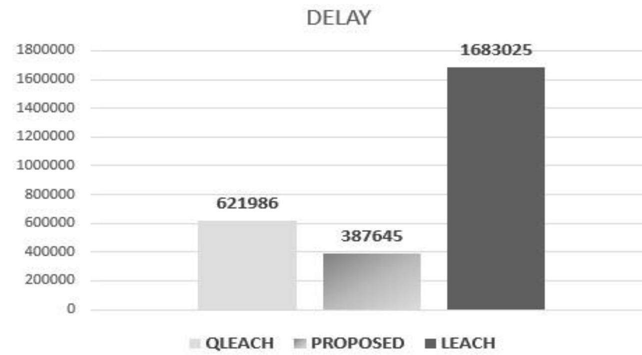

Fig 11: delay of scenario with 150 common nodes

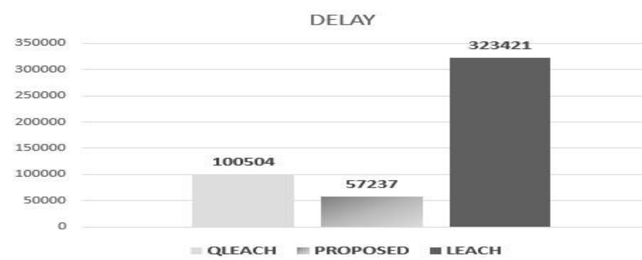

Fig 12: delay of scenario with 200 common nodes

V. CONCLUSION

We provide a plan to irrigate farms with the lowest cost possible and according to the environmental conditions and the type of plants. Our goal is to reduce costs and increase agricultural products. On the other hand, we had insisted on the correct use of water on farms. Because there is no need to irrigate plants that there is good soil moisture around them. On the other hand, if the nodes are equipped with GPS we provide a clustering protocol can be used in all scenarios. We could even consider the density of nodes in each area for determining the boundary of cells. The smallest cells in dense areas and in low-density area larger cells are formed. For this purpose, nodes should send a hello message to sink that contains their coordinates just at the beginning of the networking. The sink will calculate the density of nodes in each region and will create the cells in each region with respect to its calculations. Our proposed algorithm is a distributed algorithm. Sink just creates the cells. In other words sink determines the cluster size.

We designed our algorithm for large scale networks. We could use the second level of clustering for better performance. The overhead of second level clustering is low too. Because the clustering approach of this level is licking the first level with a little difference in criterion for electing the cluster-head.

VI. REFRENCES

1. CAO Y, H.C., *A DISTRIBUTED CLUSTERING ALGORITHM WITH AN ADAPTIVE BACKOFF STRATEGY FOR WIRELESS SENSOR NETWORKS.* IEICE TRANSACTIONS ON COMMUNICATIONS, 2006. **E89B**(2): P. 609-613.

2. YOUNIS O, F.S., *DISTRIBUTED CLUSTERING IN AD-HOC SENSOR NETWORKS: A HYBRID, ENERGY-EFFICIENT APPROACH.* IEEE INFOCOM 2004: THE CONFERENCE ON COMPUTER COMMUNICATIONS, VOLS 1-4, PROCEEDINGS, 2004: P. 629-640.

3. BANDYOPADHYAY S, C.E.J., *AN ENERGY EFFICIENT HIERARCHICAL CLUSTERING ALGORITHM FOR WIRELESS SENSOR NETWORKS.* IEEE INFOCOM 2003: THE CONFERENCE ON COMPUTER COMMUNICATIONS, VOLS 1-3, PROCEEDINGS, 2003: P. 1713-1723.

4. HEINZELMAN WB, C.A., BALAKRISHNAN H,, *AN APPLICATION-SPECIFIC PROTOCOL ARCHITECTURE FOR WIRELESS MICROSENSOR NETWORKS.* IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, 2002. **1**(4): P. 660-670.

5. INTANAGONWIWAT CH, G.R., ESTRIN D,. *DIRECTED DIFFUSION: A SCALABLE AND ROBUST COMMUNICATION PARADIGM FOR SENSOR NETWORKS.* IN *THE 6TH ANNUAL INTERNATIONAL CONFERENCE ON MOBILE COMPUTING AND NETWORKING*. 2000. ACM NEW YORK, NY, USA: THE ACM DIGITAL LIBRARY.

6. MURUGANATHAN SD, M.D., BHASIN RI, FAPOJUWO A,, *A CENTRALIZED ENERGY-EFFICIENT ROUTING PROTOCOL FOR WIRELESS SENSOR NETWORKS.* COMMUNICATIONS MAGAZINE, IEEE 2005. **43**(3): P. S8 - 13.

7. CHEN H, M.S., *CLUSTER SIZING AND HEAD SELECTION FOR EFFICIENT DATA AGGREGATION AND ROUTING IN SENSOR NETWORKS*, IN *WIRELESS COMMUNICATIONS AND NETWORKING CONFERENCE, WCNC 2006.* 2006, IEEE XPLORE.

8. HEINZELMAN WR, C.A., BALAKRISHNAN H,. *ENERGY-EFFICIENT COMMUNICATION PROTOCOL FOR WIRELESS MICROSENSOR NETWORKS.* IN *THE 33RD HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES*. 2000. USA: IEEE.

9. LIUA J SH, L.C., *ENERGY-EFFICIENCY CLUSTERING PROTOCOL IN WIRELESS SENSOR NETWORKS.* AD HOC NETWORKS, 2005. **3**(3): P. 371–388.

10. AMIS AD, P.R., VUONG THP, HUYNH DT,. *MAX-MIN D-CLUSTER FORMATION IN WIRELESS AD HOC NETWORKS.* IN *NFOCOM 2000. NINETEENTH ANNUAL JOINT CONFERENCE OF THE IEEE COMPUTER AND COMMUNICATIONS SOCIETIES.* 2000. TEL AVIV: IEEE.

11. GHIASI S, S.A., YANG X. J, SARRAFZADEH M,, *OPTIMAL ENERGY AWARE CLUSTERING IN SENSOR NETWORKS.* SENSORS, 2002. **2**(7): P. 258-269.

12. KAWADIA V, K.P., *POWER CONTROL AND CLUSTERING IN AD HOC NETWORKS.* IEEE INFOCOM 2003: THE CONFERENCE ON COMPUTER COMMUNICATIONS, VOLS 1-3, PROCEEDINGS, 2003: P. 459-469.

13. MANZOOR B, J.N., REHMAN O, AKBAR A, NADEEM Q, IQBAL A, ISHFAQ M,, *Q-LEACH: A NEW ROUTING PROTOCOL FOR WSNS.* PROCEDIA COMPUTER SCIENCE 00 (2013) 1–6, 2013: P. 1-6.