ANALYSIS OF ENERGY CONSUMING FOR CORRELATION BASED DATA AGGREGATION

R.Vijila, B.E (M.E).,
Department of communication system
PET Engineering College

C.S.Thayanandeeswari, M.E.,
Department of communication system
PET Engineering College

***Abstract* - In Wireless Sensor Networks, the sensor nodes are capable of collecting, sensing and gathering data from environment. In Data aggregation process it aggregates all sensor nodes data with minimum energy consumption and sends data to the sink node. The Pearson Correlation Coefficient method does not measure the correlation between sensor nodes in a complex environment. In this paper, a Data Density Correlation Degree (DDCD) is proposed to measure the correlation between sensor nodes data and its neighbouring sensor nodes data in communication uncertain conditions. In DDCD clustering method, all sensor nodes are divided into many clusters. Sensor nodes which are in same cluster have a high correlation degree. Every cluster has to select a representative sensor node which collects data from neighbouring sensor nodes and the data is sent to the sink node. In DDCD method, the communication radius is small and also it achieves lower distortion data. The simulation results show that the energy consumption of nodes is minimum compared with the existing methods.**

***Index terms*: Data Density Correlation Degree, Wireless Sensor Networks, Correlation, Representative sensor node**

I INTRODUCTION

Wireless sensor network consists of large number of sensor nodes deployed over an area. The recent developments of wireless sensor network (WSN) have enabled low cost, low power sensor nodes which are capable of collecting, sensing, processing and transmitting sensory data from sensing environment such as surveillance fields. The sensor nodes are cooperatively monitor the environmental conditions (for examples temperature, pressure etc.) They also have the ability to detect information such as temperature, vibration humidity and light. Then the information is processed and that the information is sent to the sink node

- In [1], J.Yick, and et al. described the important application areas of sensor networks are military areas, disaster and in health. Energy is the important resource for wireless sensor network. During data transmission process the sensor nodes consume more energy. In order to reduce the energy consumption the amount of data is transmitted to the sink node.
- In [1-2], described the energy is the main power source in a sensor node. Secondary power supply that harvests power from the environment such as solar panels may be added to the node depending on the appropriateness of the environment where the sensor will be deployed.

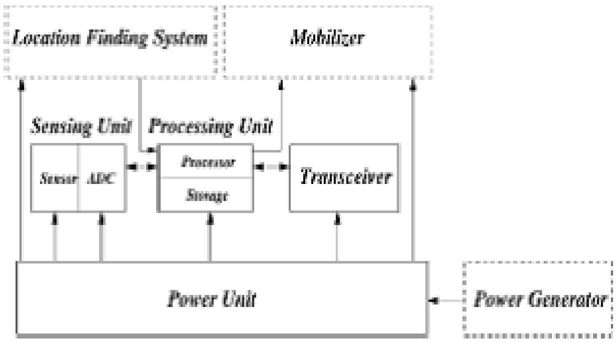

Fig 1. The components of a sensor node

- In this paper, [2] L. M. Oliveira, explained the performance of nodes deployed over an urban area. It means that the sensor network must possess self-organizing capabilities of all sensor nodes mainly broadcast communication whereas most ad hoc networks are based on point to point communication. In sensor network multi hop communication is expected to consume minimum energy than single hop communication.
- Data aggregation process was proposed by N.R Karthikeyan and et al. [3]. It is a process of aggregating the sensor data and that the data is transferred to the sink. Data aggregation process is used for improving energy efficiency in high density sensor network.
- M. C. Vuran and et al. [4] explained a spatio-temporal characteristic of point and field sources in wireless sensor network.
- C. Hua and T.S Yum [5], described by optimizing routing and data aggregation in sensor network. Maximum Lifetime Routing Algorithm (MLR) is used to avoid overwhelming bottle neck nodes, and increase the lifetime of the network. Network lifetime define the first node runs out of energy.
- [6] In this paper X. Guan and et al proposed an EECT algorithm. In this method the sensor nodes are grouped into clusters. To avoid the overlapping

the sensed zones are divided into virtual hexagons. The distance between the cluster head and sensor nodes are important factor for energy efficiency clustering methods. In this paper,

- [7] D. Kumar and et al. proposed Energy Efficient Clustering and Data Aggregation protocol for heterogeneous WSN. The main aim of EECDA protocol is to maintain the energy consumption of sensor nodes in single hop communication within a cluster.
- [8] In this paper, M/M/I queuing model was introduced by Reza Rasouli. In this queuing model, the sensor nodes are sense environment and the sensed data is sending to cluster head with less energy consumption. This model is used to explained the arrival rate of packet and the scheduled period.
- In [9], R. K. Shakya and et al. explained the high degree of correlation between sensor nodes based on location. In wireless sensor network, many sensor nodes are deployed in an event area observe event information and send to sink node. Due to the physical properties related to sensed event and information is highly correlated according to the location of each sensor nodes.
- [10] This paper, M. A .S. Monfared and et al. described by an Artificial Immune system with sensor network are used to calculate weight of each node. Consider the remaining energy in sensor nodes has to be improved by energy consumption. The Artificial Immune Network used to reduce the energy consumption in WSN.
- [11] In this paper C. Carvalho and et al. proposed a multivariate correlation which is used to decrease the prediction errors by means of multiple linear regressions. The Multivariate correlation is applied to perform prediction by means of multiple linear regressions in each sensor node. The Pearson Correlation Coefficient method is used to measure the correlation between the sensor nodes.
- [12] In this paper, Y. Guo and et al. described a α-Local Spatial (LS) clustering algorithm, where the sensors are always distributed in high density. This algorithm is explained by weight computation, CHs selection and construction of clusters.

The main aim of this paper is to propose a Data Density Correlation Degree that measures the correlation between sensor nodes and its neighbouring sensor nodes. The main concept of DDCD clustering method is to save energy by selecting representative sensor nodes that are used to sense and send their sampled data to the sink node. The DDCD clustering method highlights the problem that the existing correlation models of nodes are not appropriate for measuring the correlation in a complex environment. The sensor nodes are highly correlated in this method.

II METHODOLOGY

This section describes the energy consumption of sensor nodes in sensor network using various clustering methods. These clusterings methods are used to reduce the data transmission and save energy.

A. IN NETWORK AND GRID-BASED DATA AGGREGATION

- In-Network scheme attempts to identify the sensor which has critical information and assigns that sensor as the data aggregator transferred packets to the sink node. In this scheme, the sensor network environment is divided into pre-defined set of grids or regions. Each region is responsible for observing and reporting events that occur inside the region to the sink nodes. Also each sensor device inside the region sends data to other sensor devices only inside the region. Only one sensor act as the data aggregator which sends the critical information received either from other sensor devices or by itself to the sink nodes.
- Grid based Data Aggregation is highly suitable for mobile environments where the time duration of an event at a particular place is small. This scheme is same as In Network data aggregation. Grid based data aggregation scheme has the notion of pre-defined data aggregators in fixed regions of the sensor networks region. In addition, one sensor device based on position with respect to either the sink or the centre of the grid is chosen as the data aggregator. The Grid-based scheme also increases congestion due to increased number of packets exchanged in the protocol compared to the in-network scheme. The mobility of events affects the performance of In Network and Grid based data aggregation. This scheme increased the end-end response time.

The above two schemes are used to reduce the energy consumption and also it increases throughput. The total numbers of sensor nodes are 100. The power of sensors radio transmitted is such that any node with a 100m radius within the communication ranges. The size of data packet was set to 200 bytes which is sent to sink and packet interval was set to 100 ms. The coverage area is high. So correlations between the sensor nodes are weak. So the chance of traffic will induce. The energy dissipated is reduced by a factor 2.2. These schemes are used to reduce the energy consumption of sensor nodes but limited amount of data is transferred to the sink. Both in-network and grid based schemes are used only for single hop transmission.

B. MAXIMUM LIFETIME ROUTING PROTOCOL

The objective of this paper is to maximize the lifetime of the network by jointly optimizing data aggregation and routing protocol in WSN. The Maximum lifetime problem cannot be solved directly using distributed methods. To avoid this the maximum lifetime routing protocol was proposed. The MLR algorithm can reduce the data traffic and improve the lifetime of the network. This algorithm can converge to the optimal value efficiently under all network configurations. Network lifetime depends upon sensors energy constraints. Network lifetime can be improved from two ways, one is to reduce the data traffic across the network by aggregation, that can reduce the power consumption of the nodes close to the sink node and the other is to balance the data traffic that avoids overwhelming bottleneck nodes. The MLR algorithm has better performance than the MEGA and MER algorithm. The MLR algorithm saves more energy compared with the Minimum Energy Gathering Algorithm [MEGA] and the Minimum Energy Routing [MER].

- MEGA-This algorithm tries to optimize the data aggregation cost for raw data and data transmission costs for compressed data.
- MER-This algorithm tries to minimize the overall energy consumption of delivery of a packet by using the shortest path from source node to the sink node.

The network lifetime obtained by MLR algorithm is almost twice of MEGA and MER algorithm. The source data rate is proportional to the network. The MLR algorithm consumes less energy compared to Minimum Energy Gathering and Minimum Energy Routing algorithms. The network lifetime of MLR algorithm increase gradually as the network size grows, while those or MER and MEGA algorithms decrease continuously. MLR algorithm con optimize both routing and data aggregation. It performs much better than MER and MEGA algorithms. In MLR algorithm the distance between neighbouring sensor nodes becomes smaller. So a node needs less power to send data to its neighbour node. The data correlation between two neighbouring nodes becomes highly. So the redundant data can be removed by this method.

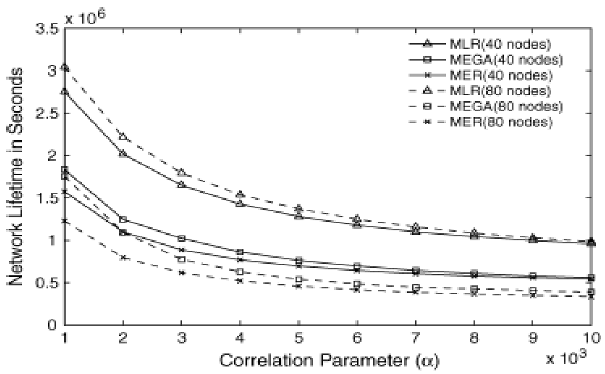

Fig 2. Correlation parameter vs. network lifetime

Fig 2. Shows that the average network lifetimes are given by MLR, MEGA, and MER algorithms as correlation parameter increases from 0.001 to 0.01. MEGA and MER algorithms achieve better network lifetime for the smaller network size (40 sensor nodes) than the larger network size (80 nodes) under correlation conditions. The performance of MEGA and MER algorithms degenerate as the data correlation becomes smaller. The normalized network lifetime obtained by the distributed MLR algorithm for various network sizes (20, 40, 60, and 80 nodes). The MLR algorithm has number of iterations required for the network lifetime to converge to over 90% of the optimal values is 5, 20, 35, and 40 iterations respectively for the network size ranging from 20 to 80 sensor nodes than MEGA and MER algorithm.

A let of sensor networks with the number of nodes ranging from 20 to 80. In each network, sensor nodes are randomly distributed on a 100mx100m square. The communication radius of all nodes is 20m. The communication radius is very small within the communication range less amount of packet will be transmitted to the end point.

C. ENERGY EFFICIENT CLUSTERING AND DATA AGGREGATION

The Sensor nodes use a large amount of energy during data transmission and data aggregation. Therefore, a new energy efficient routing protocols are required to minimize energy consumption. In this paper, an Energy Efficient Clustering and Data Aggregation (EECDA) protocol was proposed for heterogeneous WSN. Cluster head election and data communication mechanism is presented in this approach which is used to extend the network lifetime. In Cluster head election, a path with maximum sum of residual energy would be selected for data communication instead of the path with less energy consumption. The energy dissipated by the CH node during a round is given

$$E_{CH} = (n/k)\times L\times E_{elec}\times E_{EDA}+L\times \varepsilon_{fs}\times d^2_{BS}$$

n is the number of nodes, k is the number of clusters, E_{EDA} is the data aggregation and d_{BS} is the average distance between a CH and the BS.

The EECDA has the same initial energy as EDGA, LEACH and EEHCA gradually decreases in EDGA, EEHCA and LEACH over rounds. So, EDGA, EEHCA and LEACH have less residual energy left after certain rounds. The data aggregation technique is used to reduce the number of transmitted messages to the base station to save the energy and prevent the congestion. The goal of EECDA protocol is to maintain the energy consumption of sensor nodes in a single hop communication.

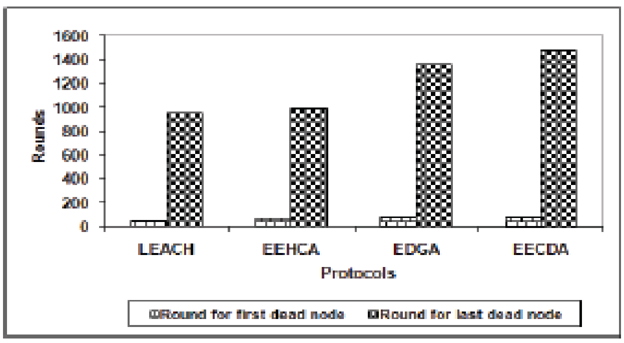

Fig 3. Network lifetime of sensor nodes over an area 100×100m^2

Figure 3 indicates that the number of alive nodes are more per round in case of EECDA as compared with EDGA, EEHCA and LEACH because a path with a maximum sum of energy residual would be selected for data transmission in spite of that path with minimum energy in case of EECDA.

In the performance of EECDA method 100 sensor nodes can be deployed over an area of size 100*100 square meters. The coverage area is large. So the nodes are not strongly correlated. The data size is 50 bytes so the limited number of packets is send to the sink node. Single-hop transmission is only suited for this method. It prolongs the network lifetime of 51% compared with the LEACH, EECHA AND EDCA protocols.

D. M/M/I QUEUING MODEL

The contribution of this paper is to evaluate the energy consumption in terms of data transmission, and aggregation using M/M/I queuing model. The main objective of data aggregation is decreasing energy consumption by decreasing need for redundant data transmission. All nodes in network are identical. The arrival of data packet to SNs transmitter queue with mean arrival rate is λ_{SN}. The service rates of exponential distribution are calculated with parameter μ_{SN} for transceiver queue and μ_p for cluster head processor head. The sensor nodes during its scheduled period of active time, remains in an idle state, switches to active state when n is equal to k and the sensor switches back from active state to sleep state. Such switching between idle state to active state and the active state is assumed to follow Poisson active state to sleep state and sleep state to idle state are referred as transitions. The aim of M/M/I queuing is used to minimize energy consumption of sensor nodes in wireless sensor network by reducing the number of transitions. In this model the SNs are periodically sense environment and send sensed data to cluster head node. The cluster head after reception of data will aggregate data and transmit the result to the sink node.

Utilization of sensor nodes is determined as

$$U_{SN}=\lambda_{SN}/\mu_{SN}$$

μ_{SN}=Bandwidth/packet size

λ_{SN}=mean arrival rate of sensor node's transmitter queue

μ_{SN}=mean service rate in sensor node transmitter queue

$P_{SN}(n)$-probability that there are n packets in queue sensor nodes in the steady state.

$$P_{SN}(n) =(1-U_{SN})U^{n}_{SN}$$

$P_{SN}(0)$-probability that there are 0 packets in queue sensor node's in the steady state:

$$P_{SN}(0)=(1-U_{SN})$$

So the total energy consumption in each sensor nodes for transmitting data to CH

$$E_{SN}=(P_{SN}(0)*E_{SLEEP})+U_{SN}(E_{IDLE}+E_{TRAN})$$

Finally the total energy consumption for N number of sensor nodes in per cluster is:

$$E_{SN}\text{-TOTAL}=(N-1)*E_{SN}$$

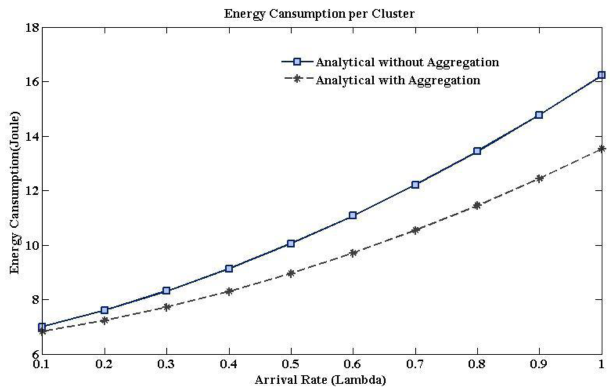

Fig 4. Energy consumption per cluster P=0.4

In Figure 4, X-axis is equal to the arrival rate (Lambda) and Y-axis is equal to the average energy consumption in per cluster. The total energy consumption of one cluster using data aggregation is higher than without aggregation. Consider elimination probability P=0.4. In M/M/I queuing model the number of sensor node per each cluster is 10 and the number of cluster is 10. The clusters have to select the cluster head. The packet size is 250 bytes. The elimination probability is decreased by 25% with data aggregation. The number of cluster head is large. So the distance between sensor nodes and cluster is more. If the transmission rate is reduced, the energy consumption is decreased. Then the traffic is increased. The M/M/I queuing model is used for multi-hop transmission.

E. ENERGY EFFICIENT CLUSTERING TECHNIQUE

In Wireless sensor network, the sensor field is divided into clusters and each cluster has representative sensor node. In circular cluster will cause overlapping to avoid this sensed field is divided into several virtual hexagons. The Energy Efficient Clustering Technique (EECT) was proposed in this paper to minimize the energy consumption and prolong network lifetime. This approach adopts the virtual hexagon as the structure of the sense field and forms the hexagon clusters in the sense field. This algorithm, concludes the optimal amounts of cluster heads and confirms the optimal cluster by energy model in the sense field. The advantage of adopting the virtual hexagon as the cluster structure is that the hexagon structure can avoid the load imbalance problem. For the cluster head, the energy consumption is composed of two parts: (1) The energy consumption of receiving and aggregating the data which comes from the normal nodes; (2) energy consumption of transmitting their own data.

Although LEACH and HEED are the attractive energy efficient schemes. These LEACH and HEED algorithms divide the sensed field into several random sub-fields. LEACH and HEED algorithm perform CH election frequently. Some of the algorithms use round shape as the structure to cover the sensed field, but round cluster will cause overlapping to the sensed zone, and this overlapping will cause the results that some sensors that join illogical clusters. In this algorithm, the distance between the centre location of the cluster head and the sensor nodes are the important reference for cluster head election. In cluster head election, firstly the hexagon cluster will be divided into circles and centre location of the cluster is their centre of the circles. LEACH and HEED are the cluster based routing protocols. LEACH resolves the energy balancing problem, but it ignores energy consumption issue of an intra clusters. As a result, the structure of clusters may not be optimal. The authors Younis and Fahmy presented an approach called HEED that periodically selects Cluster Heads based on sensor's residual energy and a secondary parameter, such as sensor proximity to its neighbors or sensor degree. The main concept of EECT on cluster head selection is that the cluster head is always closer to the center of a cluster than normal nodes. The sensor nodes are close to the center of the cluster and generally have much residual energy, since they transmit the data to the CHs via a short distance. In this way, the average distance between the cluster head and normal nodes are shortly. The EECT method achieves better performance than LEACH and HEED algorithms.

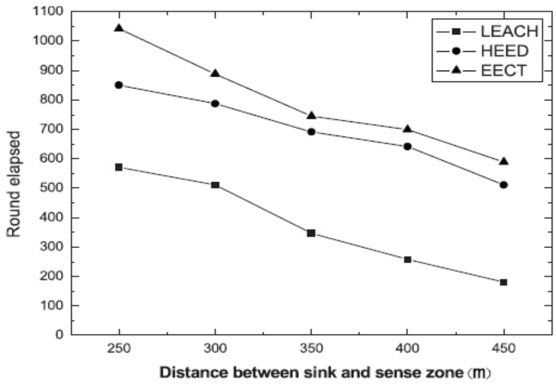

Fig 5. Distance between sink and sensed zone on the lifetime with 500 sensors

Figure 5. indicates that the comparison of EECT, LEACH and HEED with different locations of sink whose distances between sink and sensed field are 250, 300, 350, 400 and 450m. The EECT does not need to elect the cluster head frequently. Therefore, with the increasing of the distances between sink and sensed field, the energy consumption is much less than the LEACH and HEED. In this way, nodes transmit their data via a much longer distance and lead to longer latency of data packet. Meanwhile, LEACH and HEED form irregular cluster and further, increase latency of data.

In the performance of EECT method, the number of nodes is 500; the region radius of sensed field is 200 m; the initial energy of all nodes is 2.0 J; The average distance between the normal nodes and cluster head is short relatively.The EECT avoids the instance of excessive data latency since the distance between sensor nodes and cluster head is short.

III PERFORMANCE EVALUATION

- Consider the performance of Data aggregation technique in sensor networks. The total number of sensor nodes are 100 nodes in a network. This network, was deployed over a 100m radius within the communication range and is called a neighbour of a sensor. The size of each data packet was set to 200 bytes and the packet interval was set to 100 ms. The number of sensor nodes are high and the communication radius is also high within the communication radius. Limited amount of data is transmitted and the correlation between the nodes are less.
- In this method energy consumption is reduced by 2.3% and the throughput is increased by 2.2%. In the performance of Maximum Lifetime Routing algorithm a set of sensor networks with the number of nodes ranging from 20-80 are used. In each network, the sensor nodes are randomly distributed on 200 m.square. The coverage area of all sensor nodes are long. So the lifetime problem cannot be solved completely.
- The M/M/I queuing model is used to evaluate the energy consumption with mean arrival rate to SN's transmitter queue (E_{TRANS}=30mj). For decreasing the mean arrival rate the energy consumption is decreased. The number of clusters is 10. The total number of sensor nodes in each cluster is 10. The area of all sensor nodes are located at a distance of 250 m. This method does not measure the correlation between the two nodes. The energy consumption is reduced by 25%.
- In the performance of EECDA algorithm the number of nodes involved are 100 and the communication radius is 100×100 m square. The packet size of 50 bytes are transmitted to the sink node. The network lifetime is increased by 35%.
- In EECT, the sink node is located at the distance of 250 m with the sensed region; the number of sensors is 200; initial energy of all sensors is 2.0 J; the region radius of sensed zone is 200 m; the data transmision rate is 500b/s. The energy consumption is decreased by 2%. This algorithm is used for single-hop transmission.
- Comparing the overall performance, the Data Density Correlation Degree (DDCD) clustering method achieves better performance than the In-network data agggregation, MLR algorithm, M/M/I queuing EECDA and EECT algorithm. In DDCD method the communication radius is set as 6 m, each and every sensor nodes has 4-5 neighboring nodes. The communication radius is small and also correlaion between the nodes are high. So the total amount of (90%) energy will be saved. The DDCD method is used for both single and multi hop transmission .

IV CONCLUSION

In this paper, a DDCD clustering method is applied for data aggregation process. In DDCD method correlation between sensor nodes data and the neighbouring sensor nodes data are strong compared with other existing methods. The performance of DDCD method achieves lower distortion data because the communication radius is less and the energy consumption of sensor nodes are reduced. The DDCD method is adapted for all environmental conditions. In future, the work can further done by implementing the energy balancing network using all sensor nodes.

RERERENCES

[1]J.Yick, B.Mukherjee, D.Ghosal, ''Wireless Sensor Network survey,"Computer Net.,vol. 52, no.12, pp.2292–2330, 2008

[2]L. M. Oliveira and J. J. Rodrigues, "Wireless sensor networks: A survey on environmental monitoring," J. Commun., vol. 6, no. 2, pp. 1796–2021, 2011.

[3]N.R Karthikeyan Vaidyanathan, K.S Sathyantan Sur, J.Sundeep Narravula, M.S Prasun Sinha,''Data aggregation technique in sensor networks'', Int.Journal issue, vol. 8 no.5 march 2013.

[4]M. C. Vuran and O. B. Akan, ''Spatio-temoral characteristics of point and field sources in wireless sensor networks,'' in Proc. IEEE Int. Conf. Commu., June 2006.

[5] C. Hua and T.S Yum, "Optimal routing and data aggregation for maximizing lifetime of wireless sensor networks,"IEEE/ACM Trans. Netw., vol. 16, no. 4, pp. 892–903, Aug. 2008.

[6] X. Guan, L.Guan and X. G Wang ,''a novel energy efficient clustering technique based on virtual hexagon for wireless sensor networks'',International Journal of Innovative Computing, Information and Control, volume 7, number 4, April 2011.

[7]D. Kumar, T.C. Aseri, R.B. Patel, ''EECDA: Energy Efficient Clustering and Data Aggregation Protocol for Heterogeneous Wireless Sensor Networks'', Int. Journal of Computers, Communications and Control, ISSN 1841-9836, E-ISSN 1841-9844 Vol. VI (2011), No. 1 (March), pp. 113-124.

[8]Reza Rasouli, Mahmood Ahmadi, HadiTabata Malaz, Behnam Fallah, ''Analytical Model of Energy Consumption in Clustered WSNs''. International Journal of Advance Research in Computer Science & Software Engineering Volume 3, Issue 5, May 2013.

[9]R. K. Shakya, Y. N. Singh, and N. K. Verma, "A novel spatial correlation model for wireless sensor network applications," in Proc. 9th Int. Conf. WOCN, Dec. 2012.

[10]M. A .S. Monfared, Samira Nemati, and R.S Reza Azmi, ''An Energy aware Spatial Correlation Method in Event driven Wireless Sensor Networks ''IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 3, September 2012.

[11]C. Carvalho, D. G.Gomes, N. Agou-lmine, and J. N. de Souza, "Improving prediction accuracy for WSN data reduction by applying multivariate Spatio-temporal correlation," Sensors, vol. 11, no. 11, pp. 10010–10037, 2012.

[12]Y. Ma, Y. Guo, X. Tian, and M. Ghanem, "Distributed clustering based ag-gregation algorithm for spatial correlated sensor networks," IEEE Sensors J., vol. 11, no. 3, pp. 641–648, Mar. 2011.

[13]A. R. Rocha, L. Pirmez, F. C. Delicato, É. Lemos, I. Santos, D. G. Gomes, et al., "WSNs clustering based on semantic neighborhood relationships," Comput. Netw., vol. 56, no. 5, pp. 1627–1645, 2012.

[14]A. Rajeswari and P. Kalaivaani, "Energy efficient routing protocol for medium access control protocol compared with IEEE 802.11," in Proc. Int. wirelesssensor networks using spatial correlation based Conf. PACC, Jul. 2011, pp. 1–6.

[15]Matos, T.B.; Brayner, A.; Maia, J.E.B. Toward in-network data prediction in wireless sensor networks. In Proceedings of the ACM Symposium on Applied Computing, Sierre, Switzerland, 22–26 March 2010; pp. 592-596.

[16]Wang, H.; Agoulmine, N.; Ma, M.; Jin, Y. Network lifetime ptimization in wireless sensor networks. IEEE J. Sel. Areas Commun. 2010, 28, 1127-1137.

[17]J. Zheng, P. Wang, and C. Li, '"Distri- buted data aggregation using Slepian-Wolf coding in cluster-based Veh. Technol., vol. 59, no. 5, pp. 2564–2574, June 2010.

