ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

Efficient Energy Utilization and Data Prognosis in Wireless Sensor Networks

Srividhya Lakshmi R[1*], Lokeshanand V[2]

Assistant Professor, Department of Computer Science and Engineering, Sri Venkateswaraa College of Technology, Chennai, Tamil Nadu, India [1*]

Assistant Professor, Department of Computer Science and Engineering, Sri Venkateswaraa College of Technology, Chennai, Tamil Nadu, India[2]

ABSTRACT: Wireless Sensor Network (WSN) not only forms the basis of emerging smart technology but also has its own unique challenges. A Wireless Sensor Network usually consists of hundreds or thousands of small sensor nodes which operate individually in different critical and uncontrolled environments. Since these battery operated sensor nodes are difficult to recharge, achieving energy efficient data transmission in WSN still remains a challenging task. Continuous data transmission among nodes in a WSN is found to be the main reason for earlier discharge of the sensor nodes' battery power. Data collection and prediction is a crucial and an effective operation to be done to achieve longer network lifetime by reducing the frequent communication among nodes. The most recent research work on data prediction enables each node to predict its own sensor data and to compare the predicted values with the actually observed data thereby generating a confidence value. The data samples collected during the data transmission and the prediction scheme are analyzed and it is proved that data prediction can drastically reduce energy consumption of sensor nodes. But, this consumes more energy by frequent communication and also an increase in the prediction cost on each node. Our proposed system implements a new efficient hybrid data prediction approach with data aggregation that reduces communication and computational cost without compromising data quality. We further ensure security while transmitting the predicted data. The proposed system saves energy, improves network lifetime and hence improves the performance of the network with optimal cost.

KEYWORDS: Wireless Sensor Networks; Data Aggregation; Data Prediction; Energy Efficiency; Clustering

I. INTRODUCTION

A Wireless Sensor Network (WSN) is a self-configuring network consisting of small sensor nodes communicating among themselves using radio signals, and is deployed in quantity to sense, monitor and understand the physical world. Since sensor node is a small device that runs on batteries the communication is a primary energy drain [1].

The key responsibility of wireless sensor networks is to forward the sensing data gathered by sensor nodes in the sensing field to the base station. The design of protocols and applications for such networks has to be energy aware in order to prolong the lifetime of the network because it is quite difficult to recharge node batteries [2]. However, since a battery's lifetime is limited, the power resource is at a premium.

Several theoretical studies clearly demonstrate that the approaches to achieve energy efficiency suppress the vast majority of data reports at the source nodes. Nevertheless, the techniques employed are relatively complex, and their feasibility on resource-scarce WSN devices is often not ascertained. More generally, the literature lacks reports from real-world deployments, quantifying the overall system-wide lifetime. Energy efficiency can be improved by forming cluster based networks which reduces the communication between the source and the sink node [7], [10]. The lifetime of the network can be improved by grouping few sensor nodes into the sensor field and electing the cluster-head. Data aggregation can be done for specific time period in order to improve the energy efficiency [2]. Thus communication

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 5, Issue 4, April 2017

overhead is greatly reduced by minimizing the transmission of the data from each sensor node to the sink (Base Station).

The data prediction technique can greatly reduce energy consumption of nodes during communication. This data prediction technique relies on a heuristic or a state-transition model describing the sensed phenomenon [3]. Data prediction is done in order to reducing the communication that does not affect the data quality. An approach to reduce communication without compromising data quality is to predict the trend followed by the data being sensed an idea at the core of many techniques. This data prediction approach [12], [13], is applicable when data is reported periodically—the common case in many pervasive computing applications. In these cases, a model of the data trend can be computed locally to a node. This model constitutes the information being reported to the data collection sink, replacing several raw samples.

The remainder of the paper is organized as follows: Section II discusses the overview of data prediction and its various approaches, Section III describes the summary of Energy Conservation in Wireless Sensor Network, Section IV present and evaluates proposed system, and Section V provides conclusion and future work.

II. OVERVIEW OF DATA PREDICTION

Energy efficiency is an important factor in WSN. Conservation of energy is a critical issue in the design of sensor networks since the sensor nodes are battery-powered. The design of protocols and applications for such networks has to be energy aware in order to prolong the lifetime of the network because it is quite difficult to recharge node batteries. In order to improve the energy efficiency by reducing the communication and computational cost prediction techniques is used.

A. Data Prediction:

Data prediction consists of building an abstraction of a sensed phenomenon. The prediction can be done on a set of historical values (the time series) obtained by periodical samplings and are used to predict future values in the same series [6], [13]. The model can predict the values sensed by sensor nodes within certain error bounds, and reside both at the sensors and at the sink. If the needed accuracy is satisfied, queries issued by users can be evaluated at the sink through the model without the need to get the exact data from nodes. Data prediction reduces the number of information sent by source nodes to improve the lifetime of network [12].

B. Prediction Metrics:

There are three standard classifications of prediction models

1) Stochastic Approaches: It is possible to map data into a random process describe in terms of a probability density function (pdf). Data prediction is then obtained by combining the computed pdf with the observed samples. A state space representation of the phenomenon can be derived, so that forthcoming samples can be guessed by filtering out a non-predictable component modeled as noise.

2) Time Series Forecasting: Time series forecasting [5], [8], set of historical values (the time series) obtained by periodical samplings are used to predict future value in the same series. Main difference is that time series analysis explicitly considers the internal structure of data. It is represented as a combination of a pattern and a random error. Once the pattern is fully characterized, the resulting model can be used to predict future values in the time series. Fig. 2 shows Time series forecasting.

3) Algorithmic Approaches: This data prediction technique [11], [12], relies on a heuristic or a state-transition model describing the sensed phenomenon. Such algorithmic approaches derive methods or procedures to build and update the model on the basis of the chosen characterization.

$$x_t - \hat{x}_t > \in$$

Where, $\in$ is the error bound. xt is the sensor reading at time instance t. $\hat{x}_t$ is the predicted data values of xt

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

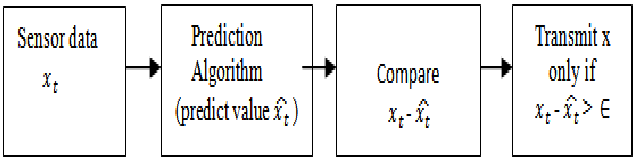

Fig. 3. Function of Source Node.

Fig. 3 shows the prediction algorithm predicts the value of the sensor node based on the previous values. If the predicted value is less than the predetermined threshold value then the sensor node will not communicate the sensed value. If the predicted value is more than the threshold then the sensor node communicates the current data to the sink node. This prediction framework greatly reduces the transmission cost and hence reduces the energy consumption of the WSN.

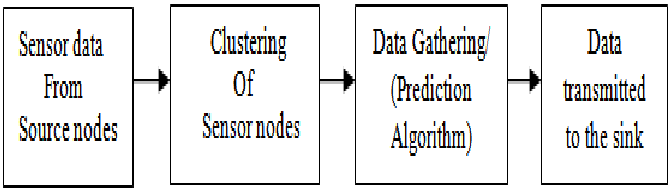

Fig. 4. Function of Sink Node.

Fig. 4 shows the sink node collects the data from all the sensor nodes. The sink node computes the magnitude similarity and trend similarity of the sensed data from the nodes. Then the nodes are partitioned into clusters using the clustering algorithm. After obtaining disjoint group of sensor nodes the sink nodes provides the schedule with which the sensor node communicates to the sink node.

III. Energy Conservation in Wireless Sensor Network

Wireless signal transmission, reception, retransmission, and beaconing operations all consumes the battery power. The design of protocols and applications for such networks has to be energy aware in order to prolong the lifetime of the network because it is quite difficult to recharge node batteries. Energy conservation is a critical issue in the design of these sensor networks since the nodes are battery-powered. We propose the clustering technology with data aggregation technique and data prediction approaches in order to improve the energy efficiency without compromising data quality.

A. Cluster-Based Wireless sensor Networks:

The approach to the fulfilment of this task is direct data transmission. In this case, each node in the network directly send the sensed data to the base station as observed in LEACH [9], the direct approach would work best if the base station is located close to the sensor nodes or the cost of receiving is very high as compared to the cost of transmitting data. However, if the base station is remote from the sensor node, the node will soon die for suffering excessive energy consumption for delivering data frequently to sink. To solve this problem by saving energy the cluster based sensor networks are used.

1) Cluster Head Election: The lifetime of the network can be improved by grouping few sensor nodes into the sensor field and electing the Cluster-Head. The Cluster Heads are elected based on the battery, memory capacity and processing ability for each cluster and only the cluster head aggregate the received data from its cluster sensors and sends to the Base Station. Clustering technology [7] improves the energy efficiency and maintains the network topology. The main objective of the cluster based networks is to reduce the communication between the source and the sink node. Fig. 5 shows clustered architecture.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

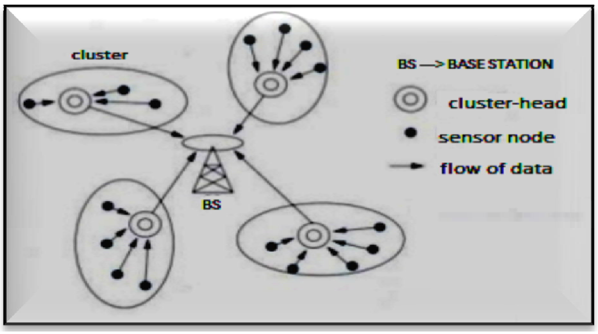

Fig. 5. Clustered Architecture.

B. Data Aggregation in Wireless Sensor Network:

Data aggregation is the process of collecting and aggregating the useful data. Data aggregation is considered as one of the fundamental processing procedures for saving the energy [2]. In WSN data aggregation is an effective way to save the limited resources. The main goal of data aggregation is to gather and aggregate data in an energy efficient manner so that network lifetime is enhanced [14]. Fig. 6 shows the sensor nodes are divided into clusters. These clusters can satisfy the intended parameter requirements and conditions that verify the amount of nodes that will be considered in a cluster. Thereafter a cluster head (CH) is selected among nodes lies within the each cluster. CH are going to be responsible for administration of all different nodes inside several cluster and collecting the data from the nodes within the cluster and transferring the information to the neighboring cluster head for more information exchange and updation. Thereafter in the data aggregation approach, all the data collected and aggregated is stored at a storage location in database server. Finally the data aggregated are going to be transfer to the base station for further use.

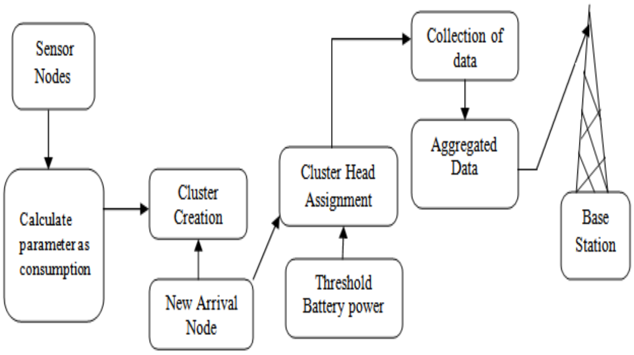

Fig. 6. Data Aggregation in Wireless Sensor Networks.

IV. PROPOSED SYSTEM

In the proposed methodology, we first segregate into three different types of clusters based on distance and range covered by the sensor nodes in order to reduce the communication overhead between source nodes to base station. We further do data prediction by generating the confidence value on each Cluster Head (CH) based on processing the historical data in the Base Station. The packets are dropped at the Cluster Head (CH) if it is below confidence value to avoid redundancy. We further aggregate the sensor data on Cluster Head (CH) for some time period. Therefore drastic reduce in communication cost and computational cost. We also provide security by encrypting and applying signature to the sensor values using efficient algorithm. Fig. 7 shows the architecture diagram to achieve network lifetime in WSN

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

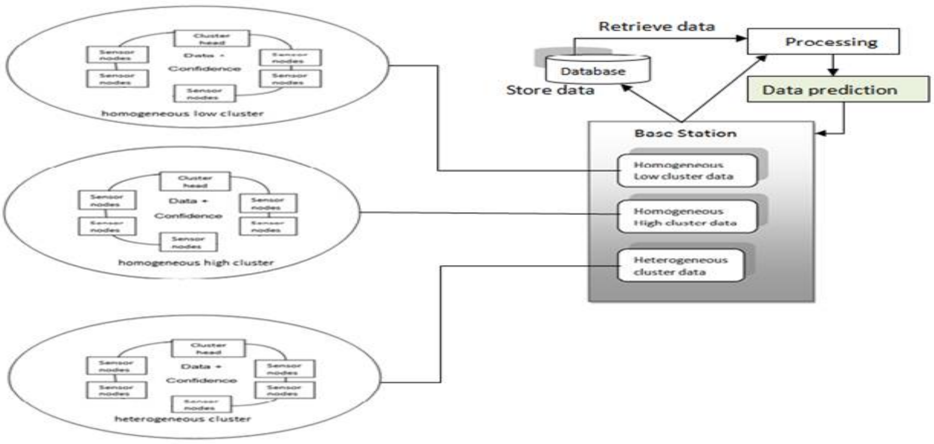

Fig. 7. Architecture Diagram.

A. Network Formation and Cluster Head Election:

We are segregating completely three different types of clusters in a Wireless Sensor Network namely,

1) Type 1 - Homogenous Low (HL) Sensors Cluster.
2) Type 2 - Homogenous High (HH) Sensors Cluster.
3) Type 3 – Heterogeneous (HE) Sensor Cluster.

The partition of three clusters is based on the distance D and coverage range R. The coverage range of current sensor nodes i is calculated as follows.

$$\mathbf{i_{min} = D + R} \qquad (1)$$

$$\mathbf{i_{max} = D - R} \qquad (2)$$

The Cluster Heads for each type is based on as follows.

$$\mathbf{S = B + M + P}$$

Calculating the cluster head based on the strength S of each senor nodes. Where B = Battery, M = Memory, P=Processing ability. The neighbor nodes are calculated based on the range R covered by the current sensor node. Other sensor nodes that are covered within the given range are considered as the neighbor node of that sensor. Calculating neighbor node from Eq. 1 and Eq. 2 is as follows. Where i = Current node and j = neighbor node.

$$((j_{min} \geq i_{min})\&\&(j_{min} \geq i_{max}))|| \\ ((j_{max} \geq i_{min})\&\&(j_{max} \geq i_{max})) \equiv \begin{cases} 0 & if\ i = j \\ Add\ j & if\ i \neq j \in R \\ \infty & if\ i \neq j \notin R \end{cases} \qquad (3)$$

B. Data Transmission and Aggregation:

Select the number of sensor nodes under each type. Sensors send its own sensing data to its cluster head using shortest path algorithm. If the source node is i and the destination node is j, k is an neighbor node in the random walk from i to j. the data get passed from source node to cluster head is as follows.

$$\mathbf{dist[i][j] = dist[i][k_1] + dist[k_1]\,[k_2] + \ldots + dist[k_n][j]} \qquad \mathbf{(4)}$$

Fig. 8 shows data transmission and aggregation is done for type 2 and type 3 sensors at the cluster head to reduce the communication overhead. No data aggregation is done at type 1 cluster due to memory constrained. Aggregation is calculated is as follows.

if(value == 4*aggregate)

aggregate (sensor node, path route, temperature, humidity, visibility, date, time)

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

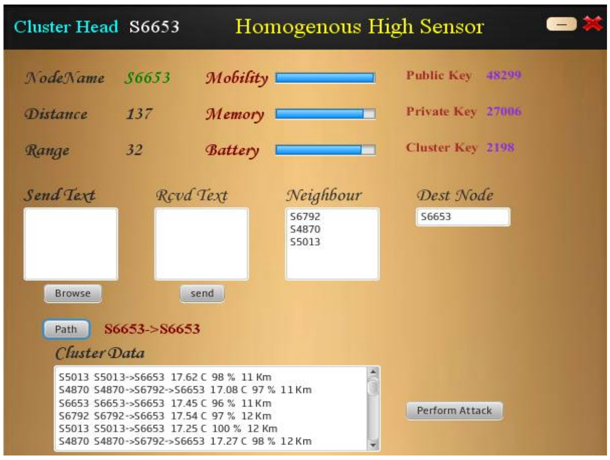

Fig. 8. Data Transmission and Aggregation.

C. Recovering Data using Signature:

Once the aggregation time is out, aggregated data's at the Cluster Head are encrypted using Elgamal Encryption algorithm and the Elgamal signature is generated for the same and data's are converted into a single packet at the cluster head and send it to the Base Station. Once the Base Station receives the sensor data it recovers each and every data by verifying signature thereby ensuring data Integrity and Authenticity. Fig. 9 shows the Encryption and Signature verification using Elgamal Algorithm.

$$\beta_A \equiv {\alpha_A}^{dA} \pmod{p_A} \qquad (5)$$

$$r \equiv \alpha_A^k \pmod{p_A} \qquad (6)$$

$$t \equiv \beta_A^k M \pmod{p_A} \qquad (7)$$

Where p_A = large prime, α_A modulo p_A=primitive element, d_A integer is chosen as private key, (P_A, α_A, β_A) is a Public Key. M is the message, k is the random integer, (r,s) is encrypted message. Signing process of the message M:

$$m = xr + ks \bmod p - 1 \qquad (8)$$

Where m is a hashed value of message M. (m,(r,s)) is the signed message. From Eq. 8 we get,

$$s = k^{-1}(m - xr) \bmod p - 1 \qquad (9)$$

Verifying the signature and decrypting the message M at the Base Station.

$$\alpha^m = y^r * r^s \qquad (10)$$

From Eq. (6) and Eq. (7)

$$tr^{-dA} \equiv \beta_A^k M (\alpha_A^k)^{-d_A} \pmod{p_A}$$

$$tr^{-dA} \equiv (\alpha_A^{d_A})^k M (\alpha_A^k)^{-d_A} \pmod{p_A}$$

$$tr^{-dA} \equiv M \pmod{p_A} \qquad (11)$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

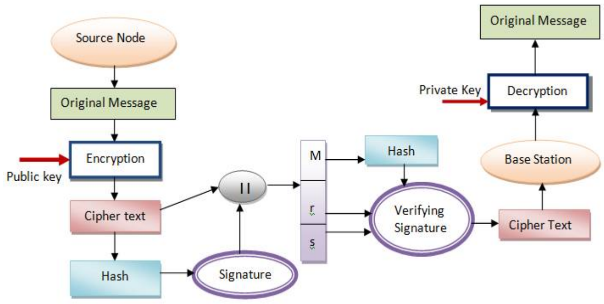

Fig. 9. Encryption and Signature verification.

D. Data Prediction and Query Processing:

The Broadcasted encrypted binary packets are being accumulated on Base Station and it is fed to a Database after Verification of Packets from various Clusters. The Historical data thus formed by time are subjected to four kinds of Query Processing. Table I shows the Query processing values based on its types.

1) Top- K Based Query Processing (Top Ranked Values on each Cluster).
2) Necessary Set Based Query Processing (Values Should be Present).
3) Sufficient Set Based Query Processing (Values that are more than enough).
4) Boundary based Query Processing (Ranked values in a range).

TABLE I
Query processing values based on its types

source node	Path	Temp	Humidity	Visibility	Dist	range
S9272	S8309->S0127	23.36 C	95 %	8km	84	41
S8540	S8540->S0127	25.97 C	78 %	5km	102	19
S5431	S9272->S0127	28.22 C	83 %	2 Km	96	25
S9272	S9272->S0127	26.26 C	96 %	6km	124	32

Data Prediction is done at the Cluster Head CH in order to reduce the redundancy of data. We are checking the observed values from the sensor nodes. If the values are not equal to the values cluster head, then sensed values are added to the Cluster Head (CH).

$$\begin{cases} \textbf{if CHECK(values)} != \textbf{CH(values)} & \textbf{CH} - \textbf{ADD(values)} \\ \qquad \textbf{else} & \textbf{Otherwise 0} \end{cases}$$

The Low Confidence data from the sensor nodes are dropped in cluster head by data prediction strategy on each cluster head. Hence the redundant data's are free from communication reducing overhead. Fig. 10 shows the packets are dropped at the Cluster Head (CH) are shown in a graph for redundancy evaluation by time vs. drop count.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

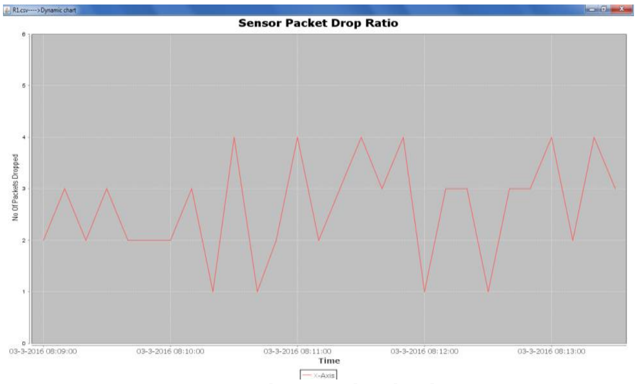

Fig. 10. Data prediction and packet drop ratio.

E. Advantages:

- The nodes are segregated to the different clusters in order to reduce the cost.
- Prediction is done at the cluster head to avoid redundancy.
- Cluster head aggregates the data for a period of time and compressed into one single packet to reduce the communication cost.

V. CONCLUSION AND FUTURE WORK

This paper has been presented for hybrid Wireless Sensor to design, develop and evaluate a energy efficient, secure and reliable system to transmit and process historical data on different types of cluster using elgamal signature scheme aggregation methods and data prediction strategies. Data prediction exploits the fact that many applications can operate with approximate data, as long as it is ensured to be within certain limits of the actual data. This allows huge reductions in communication. These prediction techniques applied, to over 13 million data points from 4 real-world applications, showing that it suppresses up to 99% of the application data, a performance often better than other approaches places minimal demands on resource-scarce WSN devices.

REFERENCES

1. R. Rathna, and A. Sivasubramanian, " improving energy efficiency in wireless sensor networks through scheduling and routing" in international journal of advanced smart sensor network systems (ijassn), vol 2, no.1, january 2012.
2. Sumit Chaudhary , Neha Singh , Avinav Pathak, and A.K Vatsa, "Energy Efficient Techniques for Data aggregation and collection in WSN"in International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.2, No.4, August 2012.
3. E. Gaura, J. Brusey, M. Allen, R. Wilkins, D. Goldsmith, and R. Rednic, "Edge mining the internet of things," Sensors Journal, IEEE, vol. 13, no. 10, pp. 3816–3825, Oct 2013.
4. D.Goldsmith, J.Brusey, "The Spanish Inquisition Protocol—Model based transmission reduction for wireless sensor networks" Sensors, 2010 IEEE conf,nov. 2010.
5. Themis Palpanas, Michail Vlachos, Eamonn Keogh, Dimitrios Gunopulos,"Streaming Time Series Summarization Using User-Defined Amnesic Functions", IEEE Transactions on Knowledge & Data Engineering, Published by the IEEE Computer Society, 2008 vol.20 ,Issue No.07 - July.
6. D. Laiymani, and A. Makhoul, "Adaptive data collection approach for periodic sensor networks",Wireless Communications and Mobile Computing Conference (IWCMC), IN IEEE International conf, July 2013.
7. Alia Sabri , Khalil Al-Shqeerat,"Hierarchical Cluster-Based Routing Protocols for Wireless Sensor Networks", IJCSI International Journal of Computer Science Issues, Vol. 11, Issue 1, No 2, January 2014 ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784. www.IJCSI.org
8. D. Tulone and S. Madden, "PAQ: Time series forecasting for approximate query answering in sensor networks," in Proceedings of the European Wkshp. on Wireless Sensor Networks (EWSN), 2006.
9. H. Jiang and S. Jin, "Leap: Localized Energy-Aware Prediction for Data Collection in Wireless Sensor Networks," Proc. IEEE Int'l Conf. Mobile Ad Hoc and Sensor Systems (MASS), 2009.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

10. Jia Yunjie, Liu Ming, Zhu Song, Dong Pengtao ,"A Clustering Routing Algorithm Based on Energy and Distance in WSN", Computer Distributed Control and Intelligent Environmental Monitoring (CDCIEM), in IEEE International Conference,March 2012.
11. Babar Nazir ,Halabi Hasbullah and Sajjad A Madani, "Sleep/wake scheduling scheme for minimizing end-to-end delay in multi-hop wireless sensor networks" 2011 Nazir et al; licensee Springer EURASIP Journal on Wireless Communications and Networking 2011, 2011:92 doi:10.1186/1687-1499-2011-92,vol.14. http://jwcn.eurasipjournals.com/content/2011/1/92 .
12. H. Jiang, S. Jin, and C. Wang, "Prediction or not? An energy efficient framework for clustering-based data collection in wireless sensor networks," IEEE Trans. on Parallel Distributed Systems, vol. 22, June 2011.
13. V.Mathibala , Mrs. Seiya Susan Thomas, Prediction Based Data Collection in Wireless Sensor Network in International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 3, March 2013.
14. C.Tharini and P. Vanaja Ranjan, "an energy efficient spatial correlation based data gathering algorithm for wireless sensor networks" in international journal of distributed and parallel systems (ijdps) vol.2, no.3, may 2011.
15. L. Gruenwald, M. S. Sadik, R. Shukla, and H. Yang, "DEMS: a data mining based technique to handle missing data in mobile sensor network applications," in Proc. of the International. Conf. on Data Mgmt. for Sensor Networks (DMSN), 2010.

BIOGRAPHY

Srividhya Lakshmi was born in Vellore, India, in 1993. She received the B.E. degree in Computer Science and Engineering of Anna University in 2014 and the M.E degree in Computer Science and Engineering of Anna University in 2016. She joined as an Assistant Professor in Department of Computer Science Engineering in Sri Venkateswaraa college of Technology. Her current research interests include Wireless Networks, Network Security, Cloud Computing, Big Data Analytics, Data Science, and Data Mining.

Lokeshanand was born in Vellore, India, in 1991. He received the B.E. degree in Computer Science and Engineering of Anna University in 2012 and the M.Tech degree in Software Technology of VIT University in 2015. He also has a professional working experience in Web Development for one year. Later, He joined as an Assistant Professor in Department of Computer Science Engineering in Sri Venkateswaraa college of Technology. His current research interests include Wireless Networks, Network Security, Cloud Computing, Big Data Analytics, Data Science, and Data Mining.