# Effectively Handle the Invalidate Data in Wireless Sensor Network Applications using W-Leach Algorithm

Aravind. P
Department Of Computer Science and Engineering
M.A.M College Of Engineering
Siruganur, Trichy, India.

Shanthi. P
Department Of Computer Science And Engineering
M.A.M College Of Engineering
Siruganur, Trichy, India.

***Abstract*—The communication between wireless sensor network applications is mostly energy consuming process, which is miserable in the WSN environment the number of sensor nodes have been placed to update the information about the capturing data with the help of sensory equipment for several computer applications, the data prediction approach has been applied previously for reducing the communication between the sensor nodes, which is achieved by eliminating the propagation of each and every sample data to the server, it is done with the use of a model to evaluate the value of captured data and by communicating with the server whenever the modification in the sampled data it will be able to appropriately described, the wireless sensor network applications burn the power not only for communicating the data even though used for various supervising operations urged by the network layer protocols, An approach to reduce the communication without concede the data characteristics is to foretell the way followed by the data being captured is the core of an any model, to overcome the deficiency in the wireless sensor network application environment to introduce the approach called derivative based prediction (DBP),it can improve the system performance by comparing the captured data each and every time with the suppression value calculated by the server with the help of initially captured data, if the value matches drop the sample data otherwise it gives the information about the changes in the environment, additionally applying the W-LEACH algorithm to randomly select the sensor node based on the highest battery power, then the highest battery power node act as a cluster head to collect the information from the other sensor nodes if any sensor node having the lowest battery power, the contents of the node have been transferred to the cluster head based on the highest battery power, which is help to prevent the data loss, improves the accuracy of data and also improves system lifetime.**

***Keywords— Wireless sensor network; Data prediction; Derivative based prediction (DBP); W-LEACH.***

## I. INTRODUCTION

Wireless sensor networks are spatially distributed autonomous sensors to measure the physical condition of the environment, sensor is a device which monitors or measures the physical characteristics of the environment sensors are repeatedly used to find and react to electrical or optical signals. A sensor transforms the physical properties into a signal which can be measured electrically. The main components of the sensors are, controller, transceiver, external memory, power source and sensors. With the numerous change in the WSN application environment it's requires high energy to capture the information so it is hard to ensure the long lifetime with the enough energy by means of using the rechargeable batteries. Generally sensors are classified into five major categories such as temperature sensor, infrared sensor, ultra violet sensor, touch sensor and proximity sensor. There are two types of sensors mostly used in WSN application environment such as active sensor and passive sensor, active sensors requires the battery power to capture the information whereas passive sensor s captures the information without requiring battery power. The major issues concerned with the wireless sensor network application environments are lifetime, energy and memory.

Data mining is the activity of finding and analyzing the huge volume of data in a data warehouse. A data warehouse is a huge backup of data pile up from extensive sources within an organization. The data prediction approach has been used to predict the method to follow up the data being captured. This can be applied whenever the data is informed repeatedly. The data prediction approach has three main drawbacks. First, it requires a communication between each sensor nodes to periodically give the information to the server. Second, requires more memory to store captured information. And the third one is it consumes more energy for capturing and transferring the information. The proposed data prediction techniques generate the model for capturing the validate information based on the sliding window calculation. It compares the latest captured information with initially captured information. If it satisfies the threshold value then update the information otherwise drop the information. Additionally using the W-LEACH algorithm to detect the highest battery power sensor node. it will act as a cluster head to store the updated validate information. Both of the can be used to improves lifetime of the sensors and minimizes the consumption of energy.

In this paper is categorized as following sections: section II represents the related works on a data prediction approaches in a both mobile sensor networks and wireless sensor network applications. Section III illustrates the derivative based prediction. Section IV illustrates the derivative based prediction with w-leach algorithm. Section V describes the system description. Section VI concludes the remarks of our proposed work.

## II. RELATED WORK

In the existing scenario there are several papers that observed on the improvement of collecting the validate information in a wireless sensor network applications that are monitoring the sensor networks, trace the active object, data processing in the scientific applications and finding the energy of the sensor nodes. Other papers are deal with concept of prediction, compression and estimation of the data.

Spiros Papadimotriou et al., [14] introducing the approach called streaming pattern discovery in multiple time series (SPIRIT).SPIRIT technique can proceeds based on the following criteria's, such as scalable, linear and adaptable. It monitors the problem of colleting the mutual relation and detecting the unmeasured values. SPIRIT can evaluate the number of data flow that can be affected within the time interval. It is used to periodically report the modification and directly measures the possible deviation in the flow of data. it can used in weather forecasting, estimating the missing values, summarizing the data flow and outlier detection. This technique has two drawbacks. First, huge volume of data travel at higher data rates. Second large amount time taken to load the process. Wei Wu et al.,[20] describe the approach Mobile data collector (MDC). That is used to analyze the data objects captured by mobile sensors depending on the user request. For capturing the data spatially distributed data forwarding protocol can be used and it uses the cache to increase the lifetime of the data. it takes the long time to answer the request and insecure to detect the noisy data. These two are the drawbacks of the above approach.

Daniela Tulone et al., [19] propose the Auto Regressive model (AR), which is used in all the sensors it foretell the all local sensor readings. It foretells the sensor readings without having the transmission to sensors. It sends the local sensor reading to the server when necessary it updates local readings. it is used to limit the transmission between the sensors. Sometimes it has flaw to predict the local readings. A.Silberstein et al., [17] addresses the data driven techniques, it can be used to capture the temperature readings, rate of the rainfall and soil conditions. For example this scenario has been applied in the forest areas to detect the growth and lifetime of the trees. This approach failed to handle the failures in the sensor readings. R.A.F Mini et al., [8] presents the method prediction based energy map, it is used to foretell the power used by the each sensor nodes. it consumes the power to construct energy map for the wireless sensor network applications. It foretells the energy of the sensor nodes based on power consumption of nodes in earlier observation. In this scenario the sensor node transmit information with the available power then the provisioning node can updates the energy of the node with the local information. It is difficult to predict the accuracy of the information based on the earlier observation and it can easily influence the jamming attack. N.Q.V Hung et al., [7] derive the model based compression techniques, which provides the benchmark to evaluate the performance of the compression based techniques. The outcome of the process is to offers the manageable and efficacious tool to maintain the changes in model based compression techniques. it is difficult to generate the predicted value of the captured data.

J.Rousse et al., [16] develops the technique called Synchronous random access protocols S-MAC and SCP-MAC.it is used to measures the traffic changes and network closeness. in the traffic changes it is used to evaluate the amount of power can be consumed based on improvement in the data rates and in the network closeness it is used to detect the movement of the sensor nodes. This technique failed to consider the storage area problems. Themis Palpanas [13] report the survey on real time processing in which real time data's are combined based on the temporal and spatial relationship then accept the deviation in data value and operate based on it. More expensive to combine the values and the performance can be reduced while in correct value occurs. Amol Deshpande et al., [2] discuss the probabilistic model which aims to provide the more reliable sensor readings in this scenario with the use of exponential time algorithm to evaluate the best solution for the optimization problem. Le Gruenwald et al., [6] focus on missing data in a mobile sensor network application. By using the data estimation approach to capture the missing data in mobile sensor networks. It mines the data based on the time and space with the help of virtual static sensors. It detects the missing data based on the conversion of mobile sensor reading to virtual static sensor readings. Sometimes the conversion value may not be efficient. Joseph Polastre et al., [15] demonstrate the approach called telos in which introduce the new sensor node that has the ability to provide less power, ease of use and reliability of the hardware and software. it can be used to improve the performance and throughput.

Omprakash Gnawali et al.,[5] uses the collection tree protocol for achieving the efficient, active and trust worthy routing. it ensures the accuracy using trickle algorithm. David Moss et al., [9] prescribed the method called BoX-MAC 1and BoX-MAC 2.These are the cross layer protocols. it describes the information about the physical layer and link layer. It achieves 40-50% of less energy consumption compare to other protocols. Christos Faloutsos et al., [18] explain the latent variable detection approach, in which the values are detected based on the calculation of sensor, processing and relationship between the sensors. This process having the connection with the sensors, then search the captured data from the sensor and finally find the measurements. Luca Mottola et al., [10] provide the study of three different tunnels in a wireless sensor network environment. An operational road tunnel provide the information about vehicles traffic, an non-operational road tunnel provide the information about the scenario without traffic. Ross Wilkins et al., [4] proposed the edge mining approach to evaluate the battery power of the sensor nodes and sensing the captured values. it is used reduced the use of energy. Suman Nath et al., [3] develop the general model called GAMPS. It combines the signals from the various sensors and then compressed. The

compression can be performed based on the polynomial time algorithm. Jing Jiang et al., [12] derive the adaptive filter approach for reducing amount of transmission from the sensor node to the server. The filter can reduce the arrival of the data flow while ensures guarantee of the data. Nicolas Burri et al., [1] present the technique called DOZER. This protocol correlates the MAC- layer, routing for reducing power to communicate to the subsystem. Using tree based approach the information can be evaluated and send to the server. Xu Ning et al., [11] discuss the two approaches called dynamic sleep time policy and optimal policy. The first approach calculates the time taken to information travel from sender node to receiver node. Second approach is used to reduce the power consumption of nodes.

In all the existing approaches less than one or two characteristics are considered to calculate the accuracy of captured data by the sensor nodes but in order to improve the overall performance of the system the data prediction and data estimation approaches are taken into the account. Hence in our proposed system the Derivative based prediction (DBP) and W-LEACH algorithms are included for improving the efficiency. The proposed schemes overcomes the deficiency of all the existing approaches and can improves the system lifetime, saves the energy and improves the performance.

## III. DERIVATIVE BASED PREDICTION TECHNIQUES

### *A. Basic Terminologies*

*1) Derivative based prediction*

The data captured by the sensor node is send to DBP model that built inside the server that calculates sliding window value of the data. The value can be compared to the currently updated value in the server.

*2) Suppression Ratio calculation*

The value can be calculated based on the below equation.

_____ (1)

*3) Weighted low energy adaptive clustering hierarchical aggregation algorithm*

This algorithm can be used to randomly select the sensor node as cluster head based on the power of the sensor node. The cluster head monitors the all the low power sensor nodes if the sensor node loss the power supply or a signal then the cluster replace the failure node.

## IV. DERIVATIVE BASED PREDICTION WITH W-LEACH ALGORITHM

As indicated in the diagram the process starts with detecting the set of sensor nodes locations, captured the information from the sensor nodes and the captured information can be send to the server then the server performs the calculation about captured information validated or in validated.

If it is a validate information stored in server. in this scenario the sensor node captured the modification in the environment. This information has been send to the DBP model that built inside the server that calculates the sliding window value (suppression ratio value) then the calculated value can be stored based on the calculation of currently finding value and previously stored value. The calculated value is same then drop the information or it varies update the information. Additionally adds the cluster head set up based on the W-LEACH algorithm it randomly select the cluster head based on power of the sensor node.

The cluster head monitors low power sensor nodes. If the low power sensor node loss the power or signal the cluster head can replace the failed sensor node. This scenario can improves the system lifetime, prevent the data loss and improve the performance. The block diagram of the proposed work is represented in fig1.

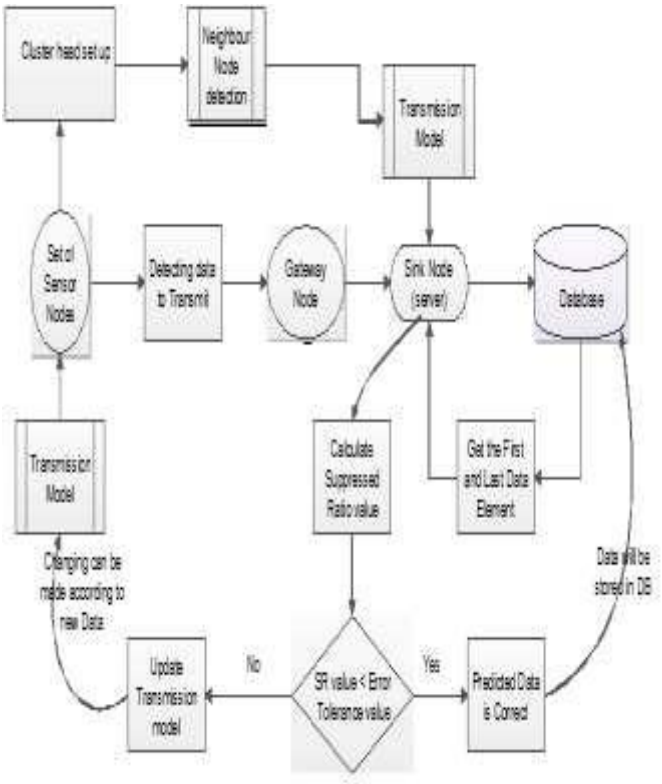


Fig. 1. System Architecture

## V. SYSTEM DESCRIPTION

This section represents the data prediction schemes and the procedures are described in detail to the following subsections.

### A. *Initial Information Processing*

The sensor nodes are deployed in the environment for capturing the information, the captured information can be send to the server with the help of the gateway node. The server before storing the captured information it detect weather it is valid information or in valid information. If the information is valid server store the information in database otherwise drop the information. Fig.2. represent the initial information processing.

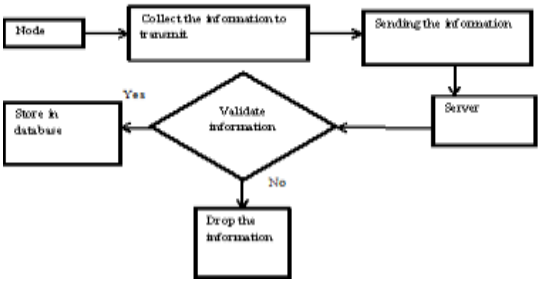


Fig. 2. Initial Information Processing

### B. *DBPModel Setup*

The sensor node captures the changes in information in a particular environment. The captured information can be send to the DBP setup that is built inside the server. This set up calculate the sliding widow value of currently captured information. Then the value can be compare with the previously stored information value. If the value matches drop the information otherwise updates the information. Fig. 3. represent the DBP model setup.

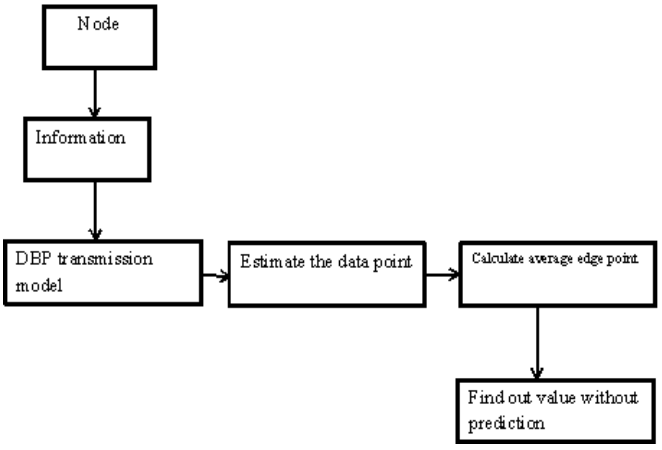


Fig. 3. DBP Model Setup

### C. *Suppression Ratio Value Claculation*

The value can be calculated based on the comparison of value with prediction and value without prediction. SR value calculation is represented in fig.4.

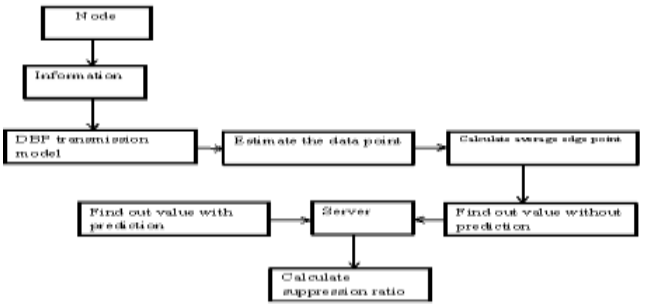


Fig. 4. SR Value Calculation

### D. *Updating The Transmission Model*

The calculated value (SR value) is compared with previously stored value. If the value matches then send the acknowledgement to the server. Otherwise update the transmission model. Updating transmission model is represented in fig.5.

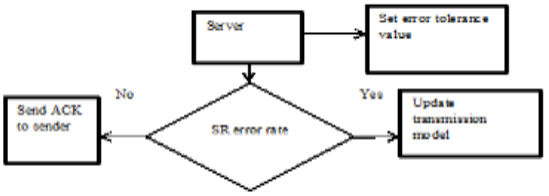


Fig. 5. Updating The Transmission Model

### E. *Cluster Head Setup*

Using the W-LEACH algorithm, it randomly chooses the cluster head based power of the node. If the node loss the power supply or signal the cluster can replace the node and improve the performance

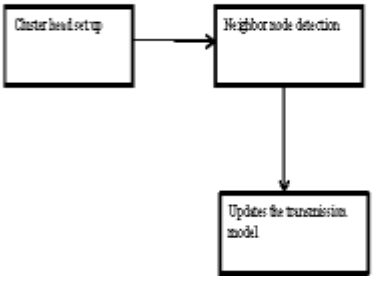


Fig. 6. Cluster Head Setup

## VI. CONCLUSION

In this paper derivative based prediction and W-LEACH algorithm are proposed. This data prediction technique construct based on initial information processing, DBP model set up, SR value calculation, updating the transmission model and cluster head set up. The both of the approach can be used to improve the system lifetime and prevent the loss of data, improving the accuracy of capturing data and saves the time, cost, and power.

## REFERENCES

[1] N. Burri, P. von Rickenbah, andR. Wattenhofer,―Dozer: Ultralowpower data gathering in sensor networks,‖ in Proc. Int. Conf.Inf. Process. Sen. Netw., 2007, pp. 450–459.

[2] A. Deshpande, C. Guestrin, S. Madden, J. M. Hellerstein, and W.Hong, ―Model-driven data acquisition in sensor networks,‖ inProc. 13th Int. Conf. Very Large Data Bases, 2004, pp. 588–599.

[3] S. Gandhi, S. Nath, S. Suri, and J. Liu, ―Gamps: Compressing multisensor data by grouping and amplitude scaling,‖ in Proc. ACM SIGMOD Int.Conf. Manag. Data, 2009, pp. 771–784.

[4] E. Gaura, J. Brusey, M. Allen, R. Wilkins, D. Goldsmith, and R.Rednic, ―Edge mining the internet of things,‖ IEEE Sensors J.,vol. 13, no. 10, pp. 3816–3825, Oct. 2013.

[5] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, ―Thecollection tree protocol,‖ in Proc. 7th Int. Conf. Embedded Netw. Sen.Syst. 2009, pp. 1–14.

[6] L. Gruenwald, M. S. Sadik, R. Shukla, and H. Yang, ―DEMS: Adata mining based technique to handle missing data in mobilesensor network applications,‖ in Proc. 7th Int. Conf. Data Manage.Sen. Netw. 2010, pp. 26–32.

[7] N. Q. V. Hung, H. Jeung, and K. Aberer, ―An evaluation of modelbasedapproaches to sensor data compression,‖ IEEE Trans.Knowl. Data Eng., vol. 25, no. 11, pp. 2434–2447, Nov. 2013.

[8] R. A. F. Mini, M. D. V. Machado, A. A. F. Loureiro, and B. Nath,―Prediction-based energy map for wireless sensor networks,‖ AdHoc Netw., vol. 3, pp. 235–253, 2005.

[9] D. Moss and P. Levis, ―BoX-MACs: Exploiting physical and link layer boundaries in low-power networking,‖ Stanford Inf. Netw. Group, Tech. Rep. SING-08-00, 2008.

[10] L. Mottola, G. Picco, M. Ceriotti, S. Guna, and A. Murphy, ―Not allwireless sensor networks are created equal: A comparative studyon tunnels,‖ ACM Trans. Sen. Netw., vol. 7, no. 2, p. 15, 2010.

[11] X. Ning and C. G. Cassandras, ―Dynamic sleep time control inwireless sensor networks,‖ ACM Trans. Sen. Netw., vol. 6,pp. 2722–2727, Jun. 2010.

[12] C. Olston, J. Jiang, and J. Widom, ―Adaptive filters for continuousqueries over distributed data streams,‖ in Proc. Int. Conf. Manag.Data, 2003, pp. 563–574.

[13] T. Palpanas, ―Real-time data analytics in sensor networks,‖ inManaging and Mining Sensor Data, C. Aggarwal, Ed. New York, NY, USA: Springer, 2012.

[14] S. Papadimitriou, J. Sun, and C. Faloutsos, ―Streaming pattern discoveryin multiple time-series,‖ in Proc. Int. Conf. Very Large Databases, 2005, pp. 697–708.

[15] J. Polastre, R. Szewczyk, and D. Culler, ―Telos: Enabling ultra-lowpower wireless research,‖ in Proc. 4th Int. Conf. Inf. Process. Sen.Netw. 2005, pp. 364–369.

[16] J. Rousselot, A. El-Hoiydi, and J.-D. Decotignie, ―Low powermedium access control protocols for wireless sensor networks,‖ inProc. Eur. Wireless Conf., Jun. 2008, pp. 1–5.

[17] A. Silberstein, G. Filpus, K. Munagala, and J. Yang, ―Data-drivenprocessing in sensor networks,‖ in Proc. Conf. Innovative Data Syst.Res., 2007, pp. 10–21.

[18] J. Sun, S. Papadimitriou, and C. Faloutsos, ―Online latent variabledetection in sensor networks,‖ in Proc. Int. Conf. Data Eng., 2005,pp. 1126–1127.

[19] D. Tulone and S. Madden, ―PAQ: Time series forecasting forapproximate query answering in sensor networks,‖ in Proc. Eur.Workshop Wireless Sen. Netw., 2006, pp. 21–37.

[20] W. Wu, H.-B. Lim and K.-L. Tan, ―Query-driven data collectionand data forwarding in intermittently connected mobile sensornetworks,‖ in Proc. 7th Int. Conf. Data Manage. Sen. Netw., 2010, pp. 20–25.