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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

Data Reduction Techniques in Wireless Sensor Network: A Survey

Thaker Maulik B.[1], Prof. Uma Nagaraj[2], Prof. Pramod D. Ganjewar[3]

PG Scholar, Dept. of CE, MIT Academy of Engineering, Pune, India[1]

Professor & Head, Dept. of CE, MIT Academy of Engineering, Pune, India[2]

Assistant Professor, Dept. of CE, MIT Academy of Engineering, Pune, India[3]

ABSTRACT: Wireless sensor network is a network which includes sensing and routing via nodes mainly called sensors and sink to monitor physical and environment condition and having many protocols according of usage .Recently, Wireless Sensor Network(WSN) used in many areas like military, hospitals, biological equipments, environment monitoring etc. The limitation of WSN includes lifetime of network, Battery, Bandwidth, Energy, data redundancy and routing etc. In this paper, we mainly focused on data redundancy and energy of sensor nodes.Data reduction is one of the data pre-processing techniques of data mining that can increase storage efficiency and reduce costs. Data reduction (DR) aims to remove unnecessary data while transmission. For this purpose many data reduction strategy introduce according to scenario of WSN. Recent algorithms and techniques based on data reduction are introduced in this survey which helps to increase energy as well as lifespan of network.

KEYWORDS: WSN; Data reduction; Energy efficiency;network lifetime

I. INTRODUCTION

WSN is an emerging technology that has wide range of potential applications including environment monitoring, surveillance, medical systems, robotic exploration, military etc.The individual nodes in a wireless sensor network (WSN) are inherently resource constrained: they have limited processing speed, storage capacity, and communication bandwidth. After the sensor nodes are deployed, they are responsible for self-organizing an appropriate network infrastructure often with multi-hop communication with them. Then the on-board sensors start collecting information of interest.

WSN consist large number of distributed nodes that organized themselves into many multi-hop wireless network. Each mode equipped with one or more sensors, embedded processor and low power radiosand is normally battery operated. A sensor node might vary in sizes and its cost. Generally, sensor node is a typical device that includes sensing (for data acquisition from environment), processing (for local data processing and storage) and communication (for data transmission).

In preceding existence, Wireless Sensor Networks (WSNs) [1],[2],[3] have gained an amplified attention from the research community and extended its boundaries in commercial, industrial and medical domains. Sensor nodes equipped with one or more sensors, memory, processor, power supply, radio and an actuator. Verities of sensors can be attached further in network to measure properties of environment. Sensor has limited power supply or low power device which can't easily replace.

Data mining is a broad area contains data pre-processing techniques like data cleaning, data smoothing, data reduction, data compression etc. Other methods like classification, clustering and association mining mechanism also give attention in data pre-processing. Data reduction schemes used when samples are unneeded or null values are sensed by sensors and unnecessary for transmission. Data acquisition is require when reducing the energy spent by sensing subsystems. In-network processing is necessary when data aggregation occurs in between sensor and sink. In this way the amount of data is reduced while traversing network towards the base station. Moreover data compression

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

can be applied to reduce amount of data sent by source nodes by involvingencoding at nodes and decoding at sink. Likewise many data pre-processing are used but here we focus on data reduction approach in wireless sensor network.

Data reduction is the process of minimizing amount of data that needs to be stored in a data storage environment. It increase storage efficiency and reduce cost. It is useful method in WSN to reduce some unnecessary data by unique features and saves energy of WSN. Here we focus on data driven approaches that are responsible for energy saving. Data reduction techniques force the sensor nodes to stop transmitting the data when it is confident about regenerating the future data at the sensor-sink based on existing, past and proximity observations thereby conserving the energy resourcesused for transmission of data. Data reduction method consists In-network processing, data compression and data prediction methods.

In this context, we present data driven approaches and various data reduction techniques classified as filter based, Tree based, Data stream based, Cluster Head based, Hybrid and other techniques.

II. DATA DRIVEN APPROACHES

Data driven approaches can be classified according to problem they addressed.(i) Data acquisition mainly target to reduce energy spent by sensing subsystem.(ii) Data reduction used when unnecessary samples are used.(iii) In-network processing do data aggregation at intermediate node between sensor and sink.(iv) Data compression reduce amount of information sent by source nodes with encoding at sensor nodes and decoding at sink.(v) Data prediction predicts the value sensed by sensor nodes within specific error bound.

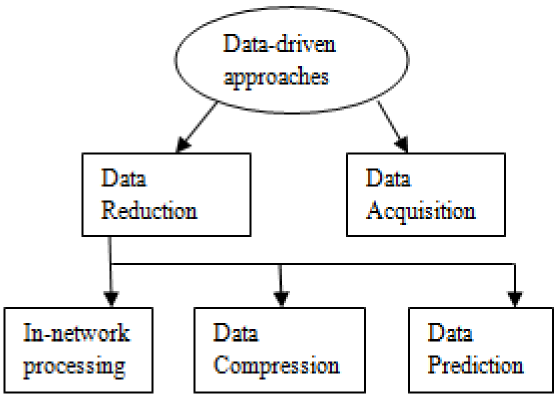

Fig-1: Data Driven Approach

III. DATA REDUCTION TECHNIQUES

A. *Adaptive Filter [10]:*

Filter is a device that reduces unnecessary features from a signal. Filter has two types stationary and non-stationary. Stationary filter doesn't allow variation of its component and non-stationary filters are required to track time variation and allow changing its coefficient. Adaptive filter works as a same manner. It works with algorithms like LMS (Least Mean Square) [5], Kalman Filter [6], LMS-VSS(LMS Variable Step-Size) [7], LMS-SSA, Extended Kalman Filter [8] etc. for giving fast prediction and better result by changing its coefficients. LMS-SSA (LMS Step Size Adaption) gives more appropriate result in case of filtering in WSN and doesn't require any prior knowledge about correlation functions and matrix inversions and requires few computations. It is suitable for tiny sensor nodes which mostly used in WSN environment. Prediction [11] in WSN provide less transmission for sensor nodes by adjusting its filter parameters. Data will be transmitted with some prediction values and compare with another values of data counter and if there is mismatching, then alert message issued by sink and both ends switched to initialization mode otherwise standalone mode [4]. LMS changes mode from normal to standalone after certain prediction error but it must be less than E_{max} ($E_{max}$is maximum prediction error handle by filter). Adaptive filter also works with Dual Prediction framework where

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

sensor and sink both has prediction model. It predicts local reading of sensor node and communication between nodes and sink will be reduced.

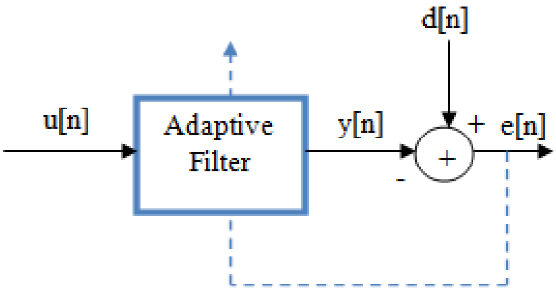

Fig-2 General Adaptive filter [10]

Prediction can be done in both time and space through a pre-defined model whose parameters depends on historical data or a prior knowledge. LMS was implemented on FPGA (Field Programmable Gate Array) [9] to reduce the communication between the sensor nodes and base station.

Weight calculation: LMS is stochastic gradient algorithm uses the estimates of gradient vector from the available data. It calculates weight of each signal in adaptive filter. It uses µ (step size filter parameter) to predict new weight and uses small values to stable filter prediction (small value prediction gives better result) The overall process is given below:

Filter Output	$y[k] = w^t[k]x[k]$
Error Signal	$e[k] = d[k] - y[k]$
Weights Adaption	w[k+1]= w[k]+µx[k]e[k]

Fig-3: LMS Weight Calculation

Where, y[k]= predicted output, e[k]= error between predicted and desired value d[k], x[k]= signal at instance k and µ= step size. Step size can calculate on every prediction with inputs. When normal mode is initializing the highest value of µ is calculated and must satisfy following condition.

$$0 \leq \mu \leq \frac{1}{E_x} \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Where, E_x is mean input power computed as:

$$E_x = \frac{1}{M}\sum_{x=1}^{M} x[k^2] \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

Where, M is number of iteration used to train filter. Advantage of LMS based filter technique is it is a lightweight approach and doesn't require much knowledge about statistical properties. Each sensor node has prediction model so no data prediction techniques required. Energy achieved up to 93% reduction in data transmission process. However this technique is not perfect for audio-video data.

B. *Tree Based Methods:*

It sometimes also based on in-network data reduction [12]. There are many tree based methods of data reduction in WSN. For mobility, we can use distance proportional energy consumption model like Khepera [36], Robomote [37],

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

LocalPos and FIRA [38]. The energy E_x(d) consumed by moving a distance d is modeled as $E_M(d) = kd$.The value of the parameter k depends on the speed of the node. Spanning tree is a graph that spans all nodes as vertices and does not contains cycles. SBTT and E-span are also energy aware spanning tree algorithms. But distance based algorithms only saves energy but not efficient enough for data reduction.

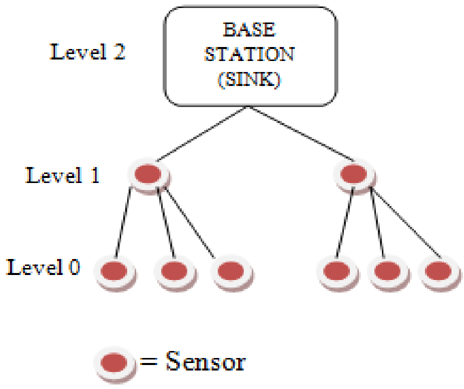

Fig- 4: Tree Based Aggregation

Hence one approach called association rule mining based tree method called MNDGT (Minimum Node Data Gathering Tree) [13] gives better energy efficiency in in-network data reduction in WSN. Association rule mining improves the QoS(Quality of Service) in wireless sensor network by taking part in resource management process. Sensor association rules (SAR) [20] capture temporal relation between sensors. In-network DR (data reduction) implemented on top of data gathering tree that will be used to relay data to the sink. SAR mines frequent patterns efficiently. It also used some data gathering algorithms like LEACH [14], PEGASIS [15], Chain based tree level Schema [16] etc. but they do not consider redundancy existing between sensor's activity.

MNDGT : Each sensor node has buffer to store data and epoch (time interval). MNDG-Tree is a tree which reduce energy and data redundancy by building a tree such as participating node construct new message format call Diff-AN and then it retransmit to the sink. Each node construct an activity set and sink begins broadcasting construction message which has Sensor-ID, Hop count(HC) and Frequency. Tree can be bottom-up or top-down that depends upon WSN environment.

Type	Sensor ID	Reference node	DIFF_SET

Fig-5: Data Message Header (DMH)

MNDGT-Construction algorithm build a tree by Sensor-ID, HC, and frequency and once tree is built then algorithm generates build message which contains only the difference set (DIFF) of active nodes (AN). After that INDR (In-network data reduction) process decides the mechanism for active and leaf nodes. It completely data reduction mechanism which is able to reduce number of message size. Greater advantage is sometimes inactive node also become mandatory nodes when neighbor nodes send build message. It deals with direct storage technique. So no other storage operation required. But it is necessary to build reference set to avoid data redundancy. In addition, TAG[51], Direct diffusion[52], DB-MAC[53], TiNA [54], DADMA[55] are routing approaches based on aggregation trees. For more aggregation techniques refer [58].

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

C. *Cluster-Head (CH) Based Reduction:*

CHRplays an important role to reduce energy in WSN thus life of network as well as scalability can be improved. Clustering is necessary when high-density networks because it can easily manage a set of cluster head (CH) from each cluster than whole nodes.

There are many algorithms like CACC [21] in which monitoring region is divided into cells (probably hexagonal cells) by taking location information of nodes. Each cluster consist approximately 6-7 cells. VAP-E[22] is used when more transmission power of every node is different and it's depend on Virtual Area Partition(VAP) in heterogeneous environment of sensor network. Moreover, PEZCA [23], VoGC [24], BARC [25], KOCA [26], CFL[27], EECS [28], EEUC [29], FoVs [47], PDCH[50] are energy saving algorithm but mostly used in wireless sensor network is LEACH(Low Energy Adaptive Clustering Hierarchy) distributed algorithm with its various descendants and used efficient techniques like LEACH-F, MS-LEACH [33], LEACH-C etc. for power consumption. In GROUP [48] only some sensor nodes responsible for making CHs, HAS Harmony Search Algorithms [49] is a music improvisation technique where musician constantly polish the pitches and minimize intra-cluster distance of network. These all are energy based cluster algorithms.

NACHO: It is divided into three phases: Initialization, in which base station calculates main parameters and transmits message to all nodes so that every node transmits the list of neighbour nodes. Cluster setup phase, based on four parameters: NRE (Node Residual Energy), Centrality, CH-frequency, Concentration. Data transmission Phase, started when CH receives data and forward resulting package to sink. This allowsupdating the energy level and whenever CH is dead, it will broadcast re-cluster message to all nodes and forcing them to disassociate from their CH. So NACHO is capable of creating better clusters, increasing energy and better packet reduction compared to all other algorithms.

SNCR(Single Node Cluster Reduction): single node cluster is a cluster which has only one member itself. MaxMin [30] algorithm is used the node address to select CH(s). Cluster head will select based on criteria like remain energy level, sensor proximity, degree of connectivity (DoC) and LQI(Link Quality Indicator). This selection procedure called floodmin phase of MaxMin algorithm. After that each CH has timer W_t inversely propostional to DoC. After certain time highest degree of connectivity is selected as CH of all other ones and sends CH-Info-Message. Overall, energy will be reduced by MaxMin algorithm is almost 99%. All cluster techniques saves energy as well as communication cost but choosing a CH is sometimes difficult when network is large.

D. *Data Stream Based Reduction:*
When data arrives in online as unordered, called data streams. Data stream reduction mostly requires when real time application used in WSN and not meet its deadline due to delay. Therefore reduction is necessary while transmission and to obtain less data delay in wireless sensor network. Sampling data stream techniques should be used to reduce data traffic.

Data Stream Reduction (DSR) used WSN as a distributed database where functions are computed and use resources available in sensor buffer or sensor database. Primary requirement of all these techniques are it should be applied on real-time application and afterwards we can generate delay metrics, impact of sample size etc. factors from appropriate techniques. Figure 5 indicates the example of real time application design based on data reduction. Sensor node receives a data stream from wireless environment and stream classified by stream organizer and stream processing chooses sample size or algorithm for generating sample.

Sampling: This algorithm provides balance between best data quality and network requirements like energy, routing, bandwidth etc. This algorithm first build the histogram of sensor data then it will create sample by randomly choose method and based on histogram. Afterwards it will sort the data samples according to its rank order in actual data stream. Analysis shows that overall complexity is O(n) (n is a number of samples) after reduction applied and communication complexity is O(md) where d is largest hop in WSN. But algorithm based on following assumption: random topology, EF-Tree [31], Stream generation, sample size, simulation in NS-2. All real-time applications are satisfy with this algorithm and it maintains the deadline in RTS, reduce energy consumption, maintain data quality and support data reduction decision.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

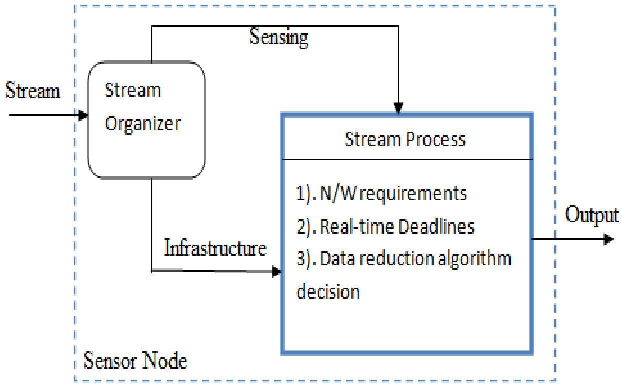

Fig-6: Real Time Application

E. *Hybrid Data Reduction (HDR):*

Data reduction techniques mainly depends on three levels: data driven reduction (where certain data management is required), event driven reduction (where some drastic changes are there in sensed parameters) and time driven (where sensor periodically sense and transmit the data) but in real-time application mostly event driven and data driven reduction is used but application like fire system, alert system also used time driven reduction approaches. So reduction techniques which work with all three aspects by data gathering protocol called HDR (Hybrid Data Reduction).

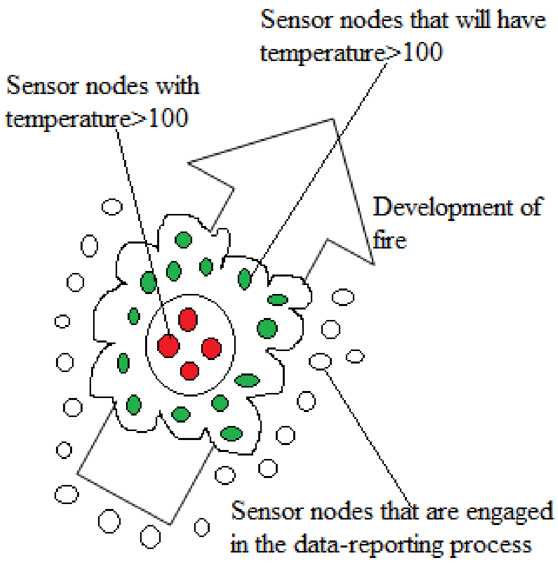

Fig-7: Hybrid WSN

Hybrid Data Gathering(HDG) protocol allow to switched between event driven data reporting(EDDR) and data driven data reporting(DDDR) such that energy will be saved and amount of data transmission will be reduced. Data reporting scheme based on WSN applications like fire system and alarm system where we will obtain the future event. For this purpose, related technology is present like cluster-based routing [32] [14]. It divide the whole network into clusters, compress the data and send messages to the sink but main challenge is to select the suitable cluster or CH(Cluster Head) and it will consume more energy and decrease the network life. HEED (cluster based data protocol)

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

can achieve the energy but it works only when network is static. Most known is Z-MAC [59] which switch between TDMA and CSMA. TDMA/CA [60] minimize data delivery time at sink in case of data gathering mechanism.

APTEEN [34]: It will allow nodes to transmit data but condition is it should maintain threshold. It combines reactive and proactive schemes.

SINA [35]: It selects the data and data collection approach based on nature of queries and active network status. But APTEEN and SINA doesn't maintain efficient data quality.

HDG protocol: Hybrid data gathering protocol works with algorithms like PED and PAD. Node will detect the future event based on current event and engaged with proactive data reporting process. HDG has following actions: (a) it switches dynamically among EDDR and DDDR and (b) nodes will detect the future event. Figure 8 illustrates the dynamic switching in HDG protocol.

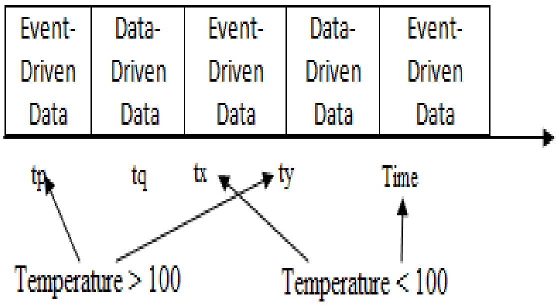

Fig-8: Dynamic Switching

To be aware when to switch, we have two algorithms called PED and PAD.

PED (Parameter based Event Detection) [39]: It has threshold value, threshold parameters and two counter variables. Counter variables decides the events where sensed value is increase or decrease compare to given threshold and whether value is above or below time interval of threshold. IT doesn't affect the normal behaviour of network and data capturing an events unless it transmit to the sink. Hence overall data will be reducing by capturing multiple events for transmission purpose.

PAD (Parameter based Area Detection) [39]: It is based on two parameters named TTL(time to live) and VT(valid time). This algorithm broadcast the change to engaged neighbour nodes in continuous data diffusion. TTL denotes number of hops within which sensor must switch to TDDR (time driven data reporting) approach and VT specifies upto when sensor nodes should use TDDR, thus not any additional reporting scheme is necessary.

ATA (Adaptive Transmission Algorithm): It is used for transmission purpose where transmission of data requires lowest energy. It stores the initial samples from datasets and sets its minimum and maximum frequency and reconstructs the minimum frequency samples and thus reduces the amount of data that will be transmitted. Moreover this technique also integrated with different abstraction levels.

F. *Data Prediction Based Reduction:*

Data prediction means to predict some value with the use of various algorithms and then choose the data accordingly. There are various types of algorithms which help for data reduction by prediction. For Example, LMS (Least Mean Square) algorithm mostly used because of its simplicity and less complexity [5]. As we shown above it also used with linear filters like adaptive filter.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

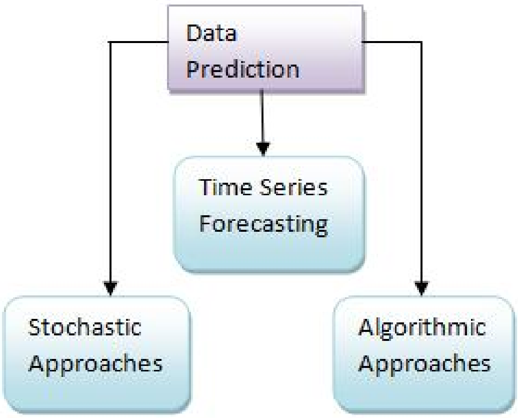

Fig-9: Data Prediction For Energy Management

1). Stochastic approach [17]: It determines depiction of random process such that probabilistic model can predict the sensed values. Ken [43] uses dynamic probabilistic model and reduce communication overhead at sensor nodes.

2). Time-Series forecasting: These models are simple but they mostly use in practical cases because of good accuracy. PAQ [18] uses auto regressive models and predict the result of every single sensor. SAF (Similarity based Adaptive Framework) [19] works with general linear time series model which includes time varying mechanism. It use to identify outliers and inconsistent data. Moving Average (MA), Auto Regressive Moving Average (ARMA) etc. are simple models and can used in practical cases. These methods give good accuracy and their implementation is simple and it's a lightweight approach.

3). Algorithmic approach: It is used to get predictions, starting from heuristic or behavioural character of sensed values. For example PREMON [44] take snapshots of network by monitoring operation continuously. It mostly suitable for cluster based WSN. Energy Efficient Data Collection (EEDC) [45] is good in inquiry-based applications, in which each sensor node relate with upper and lower bound and difference between bounds denotes the accuracy of sensed values. In EEDC, bounds are transfer to sink and later updated according to request. But algorithmic approaches are too complex for computation and sometime generate communication overheads. TGA [56] are used to improve coverage of WSN. Other prediction techniques are also useful in real time wireless sensor networks like LMS [5], WMA, ARIMA , Naïve Prediction model, PAQ[57] etc. with accuracy measures like Mean Absolute Error, Root Mean Square Error and Mean Absolute Percentage Error [46].

IV. OTHER TECHNIQUES

Further reduction techniques also used by different application based method like dimension reduction to convert high-dimensional data into low-dimensional data, Classification based approach to arrange data samples in proper way, linear aggregation based reduction etc.

IV.A. *Classification model based reduction*

These methods are heavily depend on aggregation. It uses some model like Naïve prediction model because it's a light-weight, simple and requires low memory and heavily depends on statistical parameters. WMA (weighted moving average) and LMS are good for time-series data. ARIMA [40] model provide good accuracy for time-series data. But classification model generally applies when application requires data sampling, data aggregation and these models sometimes require more statistical approach to fix the problem.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

IV.B *DR at sensor node*

Sometimes its mandatory to reduce data at sensor node before transmitting to sink to save energy of sensors. One approach called DR at sensor [41] works as same manner. It works as follows:

(a) It assume some conditions like create SB(sensor buffer), packet of 64 samples etc.
(b) If any of these conditions satisfied then transmit data and saves copy into SB.
(c) Now 2nd packet send by node and it checks the value with T_BUF.
(d) If values are same then prepare compression code and set DCB (Data Compression Bit) to '1', otherwise '0'.
(e) After that send packet to sink along with the non-repeated bytes and repeat procedure from (c). So, this way data is compress and it saves node energy before transmitting.

IV.*C Data reconstruction at base station (sink)*

It is necessary also that data should be reconstructed at sink and for that purpose data reconstruction [41] works as follows:

(a) Sink node maintains R_BUFF (received buffers)
(b) after that sink receive 1st packet from node and it saves the copy of the packet in R_BUFF.
(c) Now when another packet arrives, it will set DCB to '1' and collect repeated byte from R_BUFF
(d) Then it will collect non-repeated byte and reconstruct the original packet.This technique is helpful to reduce the energy of sink.

(IV) Data Reduction based on fuzzy logic: Fuzzy logic also called as multi-valued logic. It deals with faint parameters to obtain degree of truthfulness and falsehoods.

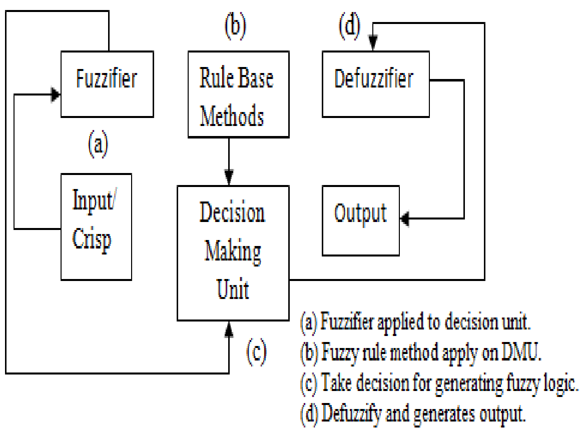

Fig-10: Fuzzy Logic System [42]

Figure 10 represents crisp set of input data are converted into fuzzy set using fuzzy linguistic variables (representation of system's input or output variables), terms and membership functions (used to map non fuzzy input values) this process called fuzzification. Now some rule based approaches applies and defuzzification step process denotes the output by its membership functions. Fuzzy logic is helpful to send only the decision of sensor nodes towards the sink node generally used for humidity, temperature, light intensity etc. It is used as a solution of power consumption in WSN via reducing messages which are sent to the sink.

V. DISCUSSION AND COMPARISON

In this section, we discuss merits and demerits of above all mention approaches in section III. However different data reduction techniques have various characteristics in wireless sensor networks and it depends on routing protocols and topology of 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)

Vol. 2, Issue 11, November 2014

No.	Reduction Technique	Advantages	Disadvantages	Suitable Algorithm
1.	Adaptive Filter	Not needed any prior knowledge of statistical properties in most techniques.	Some algorithms are complex and sometimes sink needed some transmissions to detect failures.	LMS is most suitable with adaptive filter.
2.	Tree Based Method	Fault node recovery, Able to reduce number of messages by 10-70% depends on techniques.	Need more experiments in some techniques.	MNDGT (Minimum Node Data Gathering Tree) algorithm is good for in-network data aggregation.
3.	CH/Cluster Method	Improve scalability of network, Spatial reuse of resources to gain system capacity.	Choosing a cluster head is not an easy task in case of centralized algorithms.	LEACH, NACHO are good for cluster head selection.
4.	Data Stream Method	Some streams provide security, Good for hard-WSN and reduce data traffic.	Mostly used for real time application.	It is based on data stream.
5.	Hybrid Data Reduction	Depends on combination of techniques but some approaches are good to detect future events	More complex and highly depended on protocols.	It is based on Protocol of WSN.
6.	Data prediction Method	Simple and less complex, Predict sensed values, Efficient for data reduction.	Some are complex and sometimes generates communication overheads.	Naive Bayes is best for data prediction method.

VI. CONCLUSION

The main goal of this survey paper is to evaluate the data reduction method to save energy in WSN. This survey paper delivered an overview of recent updates in reduction techniques. Many of the articles cited in this paper give their contribution to the real-world application as well as helpful for future work. Research organizations have putting their efforts to find best data reduction based approach in WSN but data reduction approach is based on WSN topology, application requirement and environment of network. Some algorithms provide good results but still having some bottleneck. More and more future work needed because each and every organization wants suitable technique for every kind of wireless sensor network. Many different types of techniques are combined to overcome individual limitations and benefit from each other's merit and measure the performance of data reduction technique in wireless sensor networks.

REFERENCES

1. F. L. LEWIS, The University of Texas at Arlington, "Wireless Sensor Networks", John Wiley , ed. D.J.Cook and S.K.Das,,New York,2004
2. Jennifer yick, Biswanathmukheerjee, Dipakghosal, "WSN survey", Elsevier, 2008.
3. KazemShoraby, Daniel minoly, Taiebznati ,"WSN,technology, protocols and applications", by john willey& sons, Inc. All rights reserved, copyright 2007.
4. R. Kwong and E. W. Johnson, "A variable step size LMS algorithm", IEEE Trans. On signal process., vol. 40, pp. 1633-1642, 1992
5. S. Haykin, "Least-Mean-Square Adaptive Filters", Inderscience,New York Willey, 2003
6. A. Jain, E.Y.Chang and Y.F.Wang: "Adaptive stream resource management using Kalman Filters". ACM SIGMOD/PODS Conference (SIGMOD '04). Paris, France, June 2004.
7. BiljanaStojkoska, dimitarSolev, DancoDavcev "Data Prediction in WSN using variable step size LMS Algorithm". SENSORCOMM 2011: The Fifth International Conference on Sensor Technologies and Applications.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

8. L.Ramrath, A.Schlaefer, F.Ernst, S.Dieterich, A.Schweikard, "Prediction of respiratory motion with a multi frequency based Extended Kalman Filter", in: Proceedings of the 21st International Conference and Exhibition on Computer Assisted Radiology and Surgery (CARS'07), Vol. 21, CARS, Berlin, Germany, 2007.
9. Nicholas Paul Borg, Dr. Carls James Debono, "An FPGA Implementation of an Adaptive Data reduction Technique for Wireless Sensor Networks", Dept. of communication and computer engineering, University of Malta.
10. G. Moschytz and M. Hofbauer: "Adaptive Filter". Springer Verlag, Berlin, 2000, ISBN 3-540-67651-1.
11. M. J. Murphy, S. Dieterich, "Comparative performance of linear and non-linear neural networks to predict irregular breathing", Physics in medicine and biology, (2006) 5903-5914.
12. Y. Yao and J. Gehrke, "The Cougar approach to in-network query" – processing in sensor networks, SIGMOD Rec., vol 31, no. 3, pp. 9-18,2002.
13. AnjanDas , "An enhanced data reduction mechanism to gather data for mining sensor association rules", 978-1-4244-9581-8/11 © 2011 IEEE.
14. Heinzelman, W.R., Chandrakasan, A., and Balakrishnan, H., "Energy-Efficient Communication Protocol for Wireless Microsensor Networks". In proceedings of the 33rd Hawaii international Conference on System Science. Hawaii, January 2000.
15. Lindsey, S., Raghavendra, C.S. "PEGASIS: Power Efficient GAthering in Sensor Information Systems". In proceedings of the IEEE Aerospace Conference. March 2002.
16. Lindsey, S., Raghavendra, C., and Sivalingam, K.M. "Data Gathering Algorithms in Networks Using Energy Metrics". IEEE Transaction on Parallel and Distributed Systems vol. (13)9, September 2002.
17. D.Chu, A. Deshpande, J.M.Hellerstein, W.Hong, "Approximate data collection in sensor networks using probabilistic models", in: proc. 22nd International Conference on Data Engineering (ICDE06), Atlanta, GA, , p. 48, April 3-8, 2006.
18. D. Tulone, S. Madden, "PAQ: time series forecasting for approximate query answering in sensor networks", in: proc., 3rdEuopean Conference on Wireless Sensor Networks(EWSN06), February 21-37, 2006.
19. D. Tulone, S. Madden, "An energy- efficient querying framework in sensor networks for detecting node similarities", in: Proc 9th International ACM Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM06), October 2006, pp. 291-300.
20. Boukerche, A., and Samarah, S., "A new representation structure for Mining Association Rules from Wireless Sensor Networks". In proc. IEEE Wireless Communications and Networking, Hong Kong 2007.
21. L. Chang-RI, Z. Yun, Z. Xian-ha, and Z. Zibo " A clustering algorithm based on cell combination for wireless sensor networks" In Second International Workshop on Education Technology and Computer Science, 2,74-77.
22. R. Wang, L. Guozhi, and C. Zheng" A clustering algorithm based on virtual area partition for heterogeneous wireless sensor networks", In International Conference on Mechatronics and Automation, 372-376.
23. F. Bai, H.Mu, and J. Sun "Power-efficient zoning clustering algorithm for wireless sensor networks", In Proceedngs of the Information Science and computer science, 1-4.
24. W. Yang, and W. T. Zhu " Voting-on-grid clustering for secures localization in wireless sensor networks", In proceedings of the IEEE International Conference on communication, 1-5.
25. K. Watfa, O. Mirza, and J. Kawtharani "BARC: A battery aware reliable clustering algorithm for sensor networks", Journal of Network and Computer Applications, 32,6, 1183-1193.
26. M. Youssef, A. Youssef, and M. Younis," Overlapping multi-hop clustering for wireless sensor networks" IEEE transactions on parallel and distributed systems, 20, 12, 1844-1856.
27. S. Zainalie and M. Yaghmaee "CFL: A clustering algorithm for localization in wireless sensor networks", In International Symposium on Telecommunications 435-439.
28. Y. Mao, L. Chengfa, C. Guihai, and J. Wu "EFCS: An energy efficient clustering scheme in wireless sensor networks", In Proceedings IPCCC, IEEE 24th International, 535-540.
29. A. Li, M. Ye, G. Chen and J. Wu, "An energy-efficient unequal clustering mechanism for wireless sensor networks", In IEEE International Conference on Mobile Ad-hoc and Sensor Systems, 604-611.
30. A.D. Amis, R. Prakash, T.H.P Voung, and D.T. Huynh. "Max-min d-cluster formation in wireless sensor networks". In 9th annual joint conference of the IEEE Computer and Communication Societies, 2000.
31. E. F. Nakamura, F. G. Nakamura, C. M. S. Figueredo, and A. A. F. Loureiro. "Using information fusion to assist data dissemination in wireless sensor networks", Telecommunication Systems, 30(1/2/3):237-254, November 2005.
32. Q. Fang, F. Zhao, and L. Guibas, "Lightweight sensing and communication protocols for target enumeration and aggregation," In proc. 4th ACM Int. Symp. Mobile Ad Hoc Netw. Comput., 2003, pp, 165-176.
33. T.Qiang, W. Bingwen and W. COM Zhicheng, "MS-Leach: A Routing Protocol Combining Multi-hop Transmission and Single-hop Transmission", Pasific-Asia Conference on Circuits, Communications and System, pp 1-4, 2009.
34. A. Manjeshwar and D. P. Agrawal, "APTEEN: A hybrid protocol for efficient routing and comprehensive information retrieval in wireless sensor networks. Parallel and Distributed Processing Symposium, Proceedings International, IPDPS, pp. 195-202, 2002.
35. Srisathapornphat, C. Jaikaeo, and C. Shen, "Sensor Information Networking Architecture", in proc. Int. workshop parallel process, 2000, pp. 22-30.
36. http://www.kteam.com/robots/khepera/index.html.
37. K. Dantu, M. Rahimi, H. Shah, S. Babel, A. Dhariwal, and G. S. Sukhatme," Robomote: enabling mobility in sensor networks," in IPSN, 2005.
38. J.-H. Kim, D.-H. Kim, Y.-J. Kim and K.-T. Seow, Soccer Robotics. Springer, 2004.
39. M. Rajalakshmi M. E., R. Karthika M. E., "Hybrid Data Reduction Scheme for Energy Saving in Wireless Sensor Networks", International journal of Science, Engineering and Technology Research (IJSETR), volume 2, issue 4, ISSN: 2278-7798, April 2013.
40. GuoruiLi and Ying Wang, "Automatic ARIMA modeling based data aggregation scheme in wireless sensor networks", Springer 2013.
41. Syed Misbahuddin, Member IEEE, MahjabeenTahir and SamiaSiddiqui, "An Efficient Lossless Data Reduction Algorithm for Cluster based Wireless Sensor Network", 978-1-4799-5158-1/14 © 2014 IEEE.
42. T. Ross, "Fuzzy logic with engineering application," John Willey & Sons Ltd., 2004.
43. D.Chu, A.Deshpande, J.M. Hellerstein, W. Hong, "Approximate data collection in sensor networks using probabilistic models", in: Proceedings of 22nd International Conference on Data Engineering 2006, pp. 48-59,2006.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 11, November 2014

44. Y.Le-Borgne, S. Santini, G. Bontempi, "Adaptive model selection for time series prediction in wireless sensor networks", Signal Process 87 (12) (2007) 3010-3020.
45. Q. Han, S. Mehrotra, N. Venkatasubramanian,"Energy efficient data collection in distributed sensor environments", in: Proceedings of the 24th IEEE International Conference in Distributed Computing Systems, pp. 590-597, 2004.
46. Femi A. Aderohunmu, GiacomoPaci, DavideBrunelli, Jeremiah D. Deng, Luca Benini, Martin Purvis, "An application specific forecasting algorithm for extending WSN lifetime", IEEE International Conference on Distributed Computing in Sensor Systems.
47. M. Alaei, and J.M. Barcelo-ordinas, "Node clustering based on overlapping FOVs for wireless multimedia sensor networks". In Proceedings of the IEEE wireless communication and networking, 1-6.
48. Y. Liyang, M. W. Neng, Z. Wei, and Z.Chunlei, " GROUP: A grid clustering routing protocol for wireless sensor networks", IEEE international conference on wireless communications, networking and mobile computing, 1-5.
49. A. C. Hoang, P. Yadav, R. Kumar, and S. Panda, "Node clustering based on overlapping FOVs for wireless multimedia sensor networks". In Proceedings of the IEEE on communications workshops (ICC), 1-5.
50. W. Linping, B. Wu, C. Zhen, and W. Zufeng, "Improved algorithm of PEGASIS protocol introducing double cluster heads in wireless sensor network". IEEE International Conference on Computer, Mechatronics Control and Electronic Engineering, 148-151.
51. S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, " TAG: a Tiny Aggregation Service for Ad-hoc sensor Networks", in OSDI 2002, Boston, MA, US, Dec., 2002.
52. A. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Direct diffusion for wireless sensor networking," IEEE/ACM Trans. Networking, vol. 11, pp. 2-16, Feb. 2002.
53. G.DiBacco, T. Melodia, and F.Cuomo, " A MAC protocol for delay bounded applications in wireless sensor network", in Med-Hoc-Net 2004, Bodrum, Turkey, Jun 2004.
54. A. Sharaf, J. Beaver, A. Labrinidis, and K. Chrysanthis, "Balancing energy efficiency and quality of aggregate data in sensor networks", the VLDB Journal, vol. 13, no.4, pp. 384-403, Dec. 2004.
55. Cayirci, "Data Aggregation and Dilution by modulus addressing in wireless sensor networks", IEEE Communication Letters, vol. 7, no. 8, pp. 355-357, Aug. 2003.
56. Shibo He, Jiming Chen, Xu Li, Xuemin (Sherman) Shen, and Youxian Sun, "Leveraging Prediction to Improve the Coverage of Wireless Sensor Networks", IEEE Transactions on Parallel and Distributed Systems, Vol. 23, No. 4, April 2012.
57. D. Tulone, S. Madden, "PAQ: time series forecasting for approximate query answering in sensor networks", in: Proceedings of the Third European Conference on Wireless Sensor Networks, , pp. 21–37,2006.
58. AnkitTripathi, Sanjeev Gupta, BhartiChourasiya, "Survey on Data Aggregation Techniques for Wireless Sensor Networks", International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 7, July 2014.
59. Rhee, A. Warrier, M Aia, J. Min, "Z-MAC: a hybrid MAC for wireless sensor networks", in Proc. ACM SenSys 2005, S Diego, USA, November 2005.
60. S. Mahfoudh, G. Chalhoub, P. Minet, M. Misson, I. Amdouni, "Node Coloring and Color Conflict Detection in Wireless Sensor Networks". Future Internet, vol. 2(4), pp. 469-504, 2010.