International Conference on Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH14) 28 & 29 November 2014

A Huffman Based Lossless Compression Algorithm for Wireless Sensor Networks

Dr. Shabana Mehfuz[1], Usha Tiwari[2], Akanksha Rathore[3], Ankit Arora[4], Diksha Singh[5]

[1]Associate Professor, Electrical Engineering Department, Jamia Millia Islamia, New Delhi

[2] Research Scholar, Electrical Engineering Department, Jamia Millia Islamia, New Delhi,

[3,4,5] Student, Electronics and Instrumentation Department, Galgotias College Of Engineering and Technology

[1]mehfuz_shabana@yahoo.com, [2]ushapant@rediffmail.com,[3] akansharathor@gmail.com,[4] ankitarora03011992@gmail.com, [5]singhdiksha115@gmail.com

Abstract—**Utilization of energy efficiently has been a main area of research in Wireless Sensor Networks (WSN). As sensors are battery operated, battery cannot be recharged frequently in the remote location where sensors are deployed. So energy optimization is of paramount importance to increase the battery life and hence the sensor network lifetime. Data compression can be an effective tool to reduce power consumption. In this paper, we have proposed a modified version of Huffman coding (MHC) algorithm for WSNs. We have taken data set of temperature variation from a sensing node. The compression ratio obtained is up to 64.325%, 60.51% and 63.08% for Le-Genepi deployment, Patrouille des Glaciers and St-Bernard deployment respectively.**

Keywords **— Energy Consumption, WSN, Compression Ratio.**

I. INTRODUCTION

WSN is the most emergent field of science. Due to its wide applications in home automation system [1], health care monitoring system [2], agriculture or precise farming system [3], etc many knowledge seekers are showing their wide interest in this field.

A WSN is a combination of sensing nodes placed in a geographical location to measure geographical parameters such as temperature, humidity, solar radiations etc as shown in Fig.1. A sensor node has three main units namely sensing unit, processing unit and transmission unit as in Fig.1. Sensing unit senses the real world data. Processing unit processes the data accordingly and transmitting unit transmits the data from local station to the base station through radio communication or internet. Maximum power is required to exchange the information between the nodes and from nodes to base station. So a radio transmitter on board is the main factor of power consumption. Thus, till date several researches are aimed at reducing the radio communication to achieve sufficient energy saving. To achieve this objective data compression in wireless sensor network is required.

Along with the advantage of using WSN there are certain challenges that are being faced during the design of WSN. These are broadly classified into three major issues [4]. First, an information gathering architecture has to be designed to avoid the conflict of data. Also information loss should be minimised [5]. Second WSNs are placed in remote location where they are prone to various kinds of rancorous attacks. Hence security of data is of major concern. Therefore data should be encrypted and connection should be authenticated. Finally, as the sensor nodes are battery operated and deployed in a network consisting of hundreds to thousands of sensor nodes, recharging of battery is not feasible. So it raises the importance of power management in wireless sensor networks.

To overcome these problems an efficient data compression technique is highly required. We have proposed an algorithm that supports lossless data compression technique and pursue the characteristics that solemnly suit the demand of WSN such as less memory requirement and less complexity.

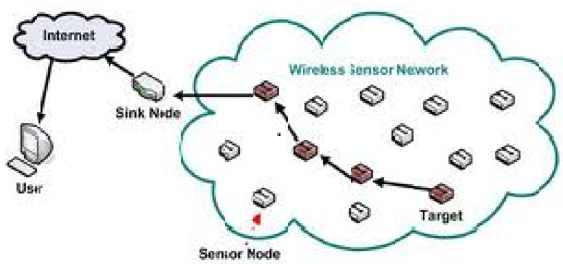

Fig. 1 Block Diagram of Wireless Sensor Network

II. RELATED WORK

Due to its wide are of application in real time, WSN is a very important topic of interest among the researchers. Its ability of being used in almost every functional area makes it very popular.

Based on the various researches and analysis, data compression techniques are broadly classified into two categories [5]: a Distributed Data Compression approach and a Local Data Compression approach. Distributed approach is mainly applied on dense sensor networks (containing large number of sensor node) and these schemes exploit the spatial correlation among the sensor nodes. While Local approaches perform data compression locally on each sensor node and these schemes usually exploits the temporal correlation among the sensor nodes and do not depend on the WSN's specific topology.

A distributed compression algorithm for multi hop, distributed sensor network based on the lifting factorization as in Fig 2. illustrates the Lifting algorithm for 5/3 wavelet ,that exploits the natural data flow in the network to aggregate by computing partial wavelet coefficient that are refined as the data flows towards the central node [6]. In addition to this, this algorithm also introduces an extra transmission in the network so as to give the node access to the neighbouring data making it possible to compute the transform coefficients. The another algorithm, exploits the natural data flow in the network to aggregate data by computing partial wavelet coefficient that are refined as the data flows towards the central node [7].

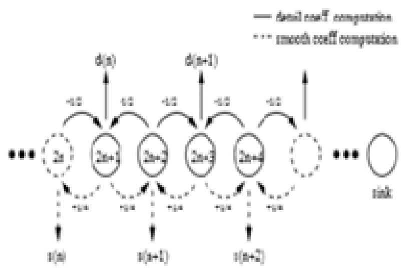

Fig 2. Lifting Implementation of Wavelet transform

Another approach is developed for collaboratively training networks of kernel-linear least square regression estimates as in Fig 3. The algorithm is shown to distributively solve a relation of the classical centralised least square aggression problem [10].

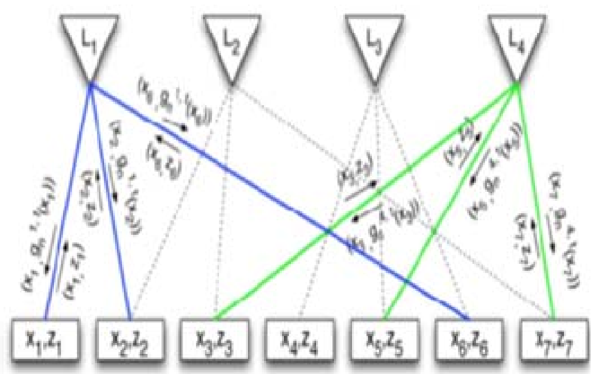

Fig 3. A Colloborative Training algorithm

As the distributed approaches have some limitations and issues, the Local approaches do not suffer from these issues. These kind of approaches provides a more robust data compression schemes as compared to the distributed one. We can do amalgamation of local approaches with distributed approaches to exploit both spatial and temporal correlation among the data in WSN. In Local approaches [8], author introduces a sensor version of standard LZW(Lempel-Ziv-Welch) algorithm. They worked on garnering additional significant energy improvement by devising computationally efficient lossless compression algorithm on the source node as in Fig. 5. Another Local approach as in [11], have proposed a simple lossless entropy compression (LEC) algorithm (particularly suited for WSN) which can be implemented in a few lines of codes and hence have obtained a compression ratio up to 70.81% and 62.08% for temperature and humidity data sets, respectively, and of the order of 70% for the other data sets. A sophisticated framework to systematically explore the temporal correlation in environmental monitoring WSN has been presented in [12].

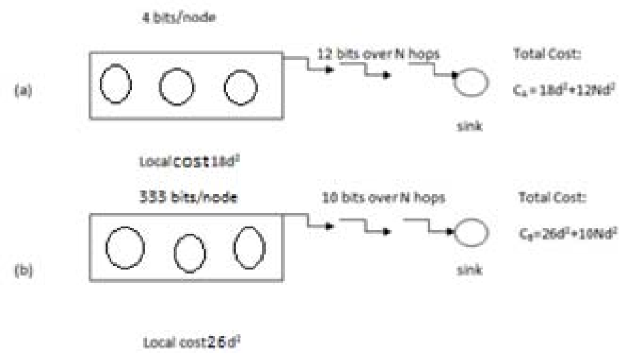

Fig4. (a) Method A: A simple encoding scheme is used; 12 Bits are sent to the sink node.
(b) A locally more expensive method is used, but Better compression is acheived;10 bits are sent to the sink.

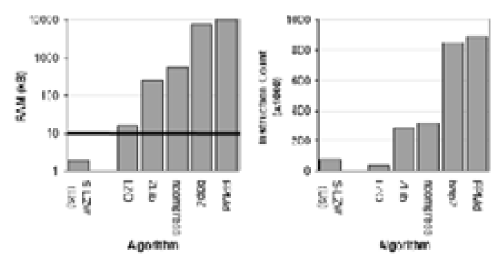

Fig5. Brief comparison of the compression algorithm in And S-LZW algorithm.

III. WORK PROPOSED

The MHC Algorithm uses a modified version of standard Huffman Algorithm [13]and where the length of code is based on the Exponential-Golomb code (Exp-Golomb) of order 0 which increases exponentially as the probability of occurrence decreases [14]. A little modification is done to work with negative integers.

We have taken real-world sensor data l_i and delayed data l_{i-1} and passed this data to an ADC to get the digital data b_i and b_{i-1}. We have calculated the difference d_i which is the difference between b_i and b_{i-1}. We have calculated the length of d_i given by n_i which is equal to $\log_2|d_i|$ which is at most equal to R. R is the resolution of ADC and for Sensirion SHT175(Sensor Module), its resolution is 14 bit for Temperature and 12 bit for Humidity. The encoder and decoder is shown in Fig 6 and Fig 7.Block diagram is based on the CCDS recommendation for Lossless compression algorithm as in [14].First block represents the pre-processing part and the second block shows the binary coding part.

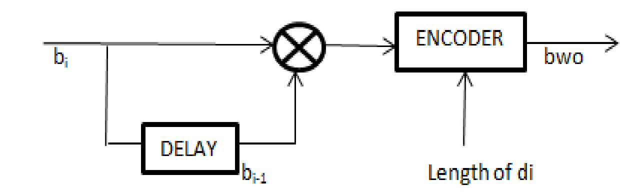

Fig 6. Block Diagram of MHC Encoder

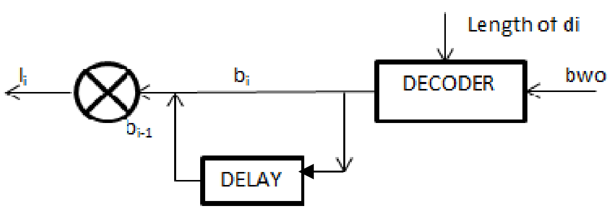

Fig 7. Block Diagram of MHC Decoder

In order to manage negative d_i will have taken index as in [14]

$$\text{Index} = \begin{cases} d_i & d_i >= 0 \\ 2^{n_i} - 1 - |d_i| & d_i < 0 \end{cases}$$

We have calculated n_i as per the following Table:

n_i	s_i	d_i
0	00	0
1	010	−1,+1
2	011	−3,−2,+2,+3
3	100	−7,....,−4,+4,....,+7
4	101	−15,....,−8,+8,....,+15
5	110	−31,....,−16,+16,....,+31
6	1110	−63,....,−32,+32,....,+63
7	11110	−127,....,−64,+64,....,+127
8	111110	−255,....,−128,+128,....,+255
9	1111110	−511,....,−256,+256,....,+511
10	11111110	−1023,....,−512,+512,....,+1023
11	111111110	−2047,....,−1024,+1024,....,+2047
12	1111111110	−4095,....,−2048,+2048,....,+4095
13	11111111110	−8191,....,−4096,+4096,....,+8191
14	111111111110	−16 383,....,−8192,+8192,....,+16 383

Table 1: Pattern used in MHC

This d_i is represented as bitwise output bwo which consist of two parts $s_i|q_i$. Where s_i is obtained from the table as per the corresponding n_i. And q_i is the representation of d_i.

This data bwo is sent to the base station and is then decoded using the MHC decoder. However it is to be noted that during decoding, the first sample should be known else decoding would not be possible.

The flowchart of the MHC algorithm is shown in fig 8.

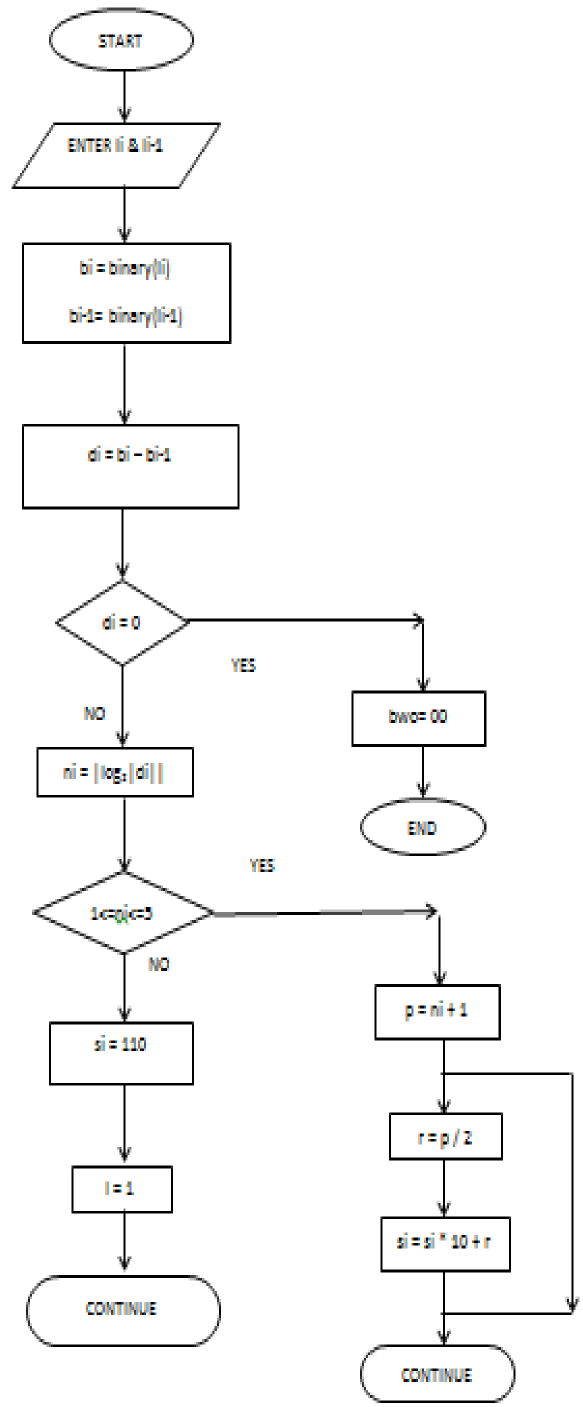

Fig 8. MHC Flowchart

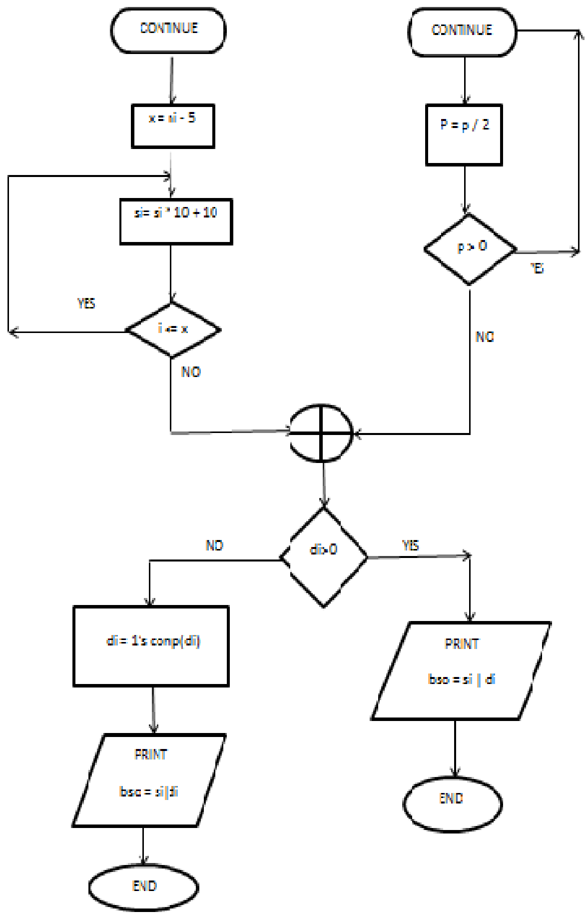

Fig 9. Flowchart for MHC algorithm

\

In order to check the correctness and effectiveness of our algorithm we checked it with the real world data. There are many places where sensing nodes are in use and have shown remarkable success. Hence we took data from such deployments and applied MHC algorithm to look at its utility. Table 2 shows the data samples details taken under consideration. We have taken three deployments obtained from sensorscope.com [15]. We have found the compression ratio.

COMPRESSION RATIO:

The compression performance of any data compression algorithm is defined by a parameter called *compression ratio.* It signifies how much the data is compressed. higher the compression ratio more effective the algorithm is and hence lesser the power is consumed.

Compression ratio is given by:

CR= **100*(1 – (C.S/O.S))**

Where C.S is the compressed size obtained after applying data compression algorithm and O.S signifies the original size of the sample data. Table 3 shows the compression ratio obtained from the given deployments using MHC algorithm.

DATA SET	ORIGINAL SIZE	COMPRESSED SIZE	CR (%)
GSB_ID03	8000	2953	63.08
LG_ID03	8000	2854	64.325
PDG_ID01	8000	3159	60.51

Table 3. Compression Ratio of Various Deployments using MHC Algorithm

DEPLOYMENT NAME*	NODE ID	SYMBOLIC NAME	NUMBER OF SAMPLES	TIME INTERVAL	
				FROM DAY	TO DAY
Le Génépi	03	LG_ID03	500	28 August 2007	29 August 2007
Grand-St-Bernard	03	GSB_ID03	500	13 September 2007	14 September 2007
Patrouille des Glaciers - Apr. 2008	01	PDG_ID01	500	15 April 2008	16 April 2008

*Data Obtained from Sensorscope [15].

Table 2. Characteristics of Various Deployments

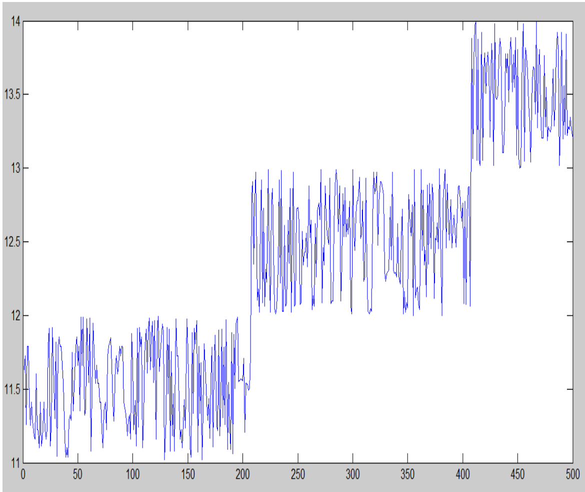

Fig 10. Variation in temperature of LG_ID03 versus no. of samples

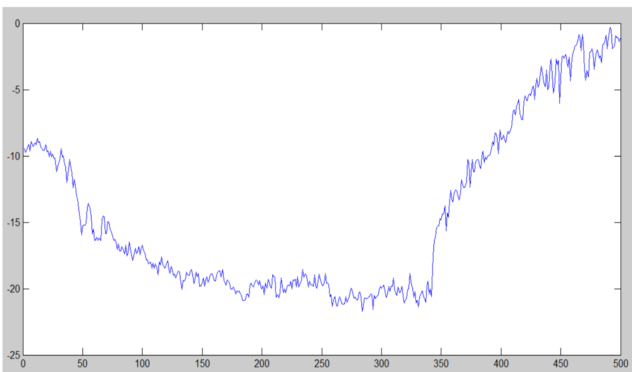

Fig 11. Variation in temperature of PDG_ID01 versus no. of samples

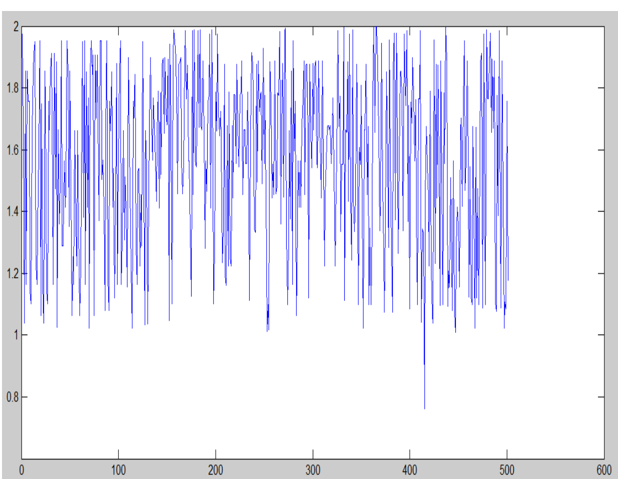

Fig 12. Variation in temperature of GSB_ID03 versus no. of samples

Fig 10, 11, 12 shows the variation of temperature as the number of samples increases.

IV. RESULT AND FUTURE WORK

In this paper, we have proposed MHC algorithm which is simple, consumes less memory and fits the requirement of a WSN. It not only possesses the worthy characteristics but also has the ability to obtain considerable compression ratio. We have obtained compression ratio upto 64.325%, 60.51% and 63.08% for Le-Genepi deployment, Patrouille des Glaciers and St-Bernard deployment respectively for temperature. After developing a framework for optimization, we would like to apply the following algorithm into hardware system to look into a practical aspect of WSN. For this purpose we are trying to develop an interface between MHC algorithm and hardware so that we can analyze the effectiveness of the MHC algorithm.

REFERENCES

[1] Song G, Zhou Y, Zhang W, Song A. Multi-interface gateway architecture for home automation networks. IEEE Trans Consum Electron 2008;54(3):1110-3.

[2] Gao T, Massey T, Selavo L, Crawford D, Chen B, Lorincz K, et al. the advanced health and disaster aid network: a light weight wireless medical system for triage. IEEE Trans Biomed Circuits Syst 2007;1(3):203-16.

[3] Riquelme JL, Soto F, Suardiaz J, Sanchez P, Iborra A, Vera J. Wireless Sensor Networks for precision horticulture in Southern Spain. Comput Election Agric 2009;68(1):25-35.

[4] Baronti P, Pillai P, Chook VWC, Chessa S, Gotta A, Hu YF. Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards. Comput Commun 2007;30(7):1655-95.

[5] Tossaporn Srisookai, Kamol Keamarungsi, Poonlap Lamsrichan, Kiyomichi Araki: Practical data compression in wireless sensor network35(2012)37-39.

[6] Alexandre Ciancio and Antonio Ortega: A Distributed wavelet compression algorithm for wireless multihop sensor networks using lifting: IEEE 2005.

[7] Alexandre Gomes Ciancio: Distributed Wavelets Compression Algorithm For Wireless Sensor Networks, December 2006.

[8] Christopher M. Sadler and Margaret Martonosi: Data Compression Algorithms for Energy-Constrained Devices in Delay Tolerant Networks: SenSys'06, November 1–3, 2006, Boulder, Colorado, USA.

[9] A.Ciancio, S.Pattem, A.Ortega and B. Krishnamachari: Energy-Efficient Data Representation and Routing for Wireless Sensor Networks Based on a Distributed Wavelet

Compression Algorithm: IPSN'06, April 19-21, 2006, Nashville, Tennessee, USA.

[10] Joel B. Predd, Sanjeev R. Kulkarni, H. Vincent Poor: A Collaborative Training Algorithm for Distributed Learning: IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 55, NO. 4, APRIL 2009.

[11] F.Marcelloni and M.Vecchio: An Efficient Lossless Compression Algorithm for Tiny Nodes of Monitoring Wireless Sensor Networks: The Computer Journal Advance Access, April 30, 2009.

[12]Yao Liang, Wei Peng: Minimizing Energy Consumptions in Wireless Sensor Networks via Two-Modal Transmission: ACM SIGCOMM Computer Communication Review Volume 40, Number 1, January 2010.

[13] Teuhola, J. (1978) A compression method for clustered bit- vectors. Inf. Process. Lett., 7, 308–311.

[14] Sayood K. Introduction to data compression 3rd ed. US: Morgan Kaufmann; 2006.

[15] Sensor Scope Deployments homepage(2014) http://sensorscope.epfl.ch/index.php/Main_Page. (Accessed April 25. 2014).