DATA COMPRESSION TECHNIQUES IN WIRELESS SENSOR NETWORKS : A SURVEY

A. Jeevarathinam[1], Dr. S.Manju Priya[2]

ABSTRACT

Energy efficiency is a primary concern since sensor nodes have to report regularly the sensing data to the remote sink(s) for a very long time. So data transmission is one of the major factors for the energy efficiency in sensor networks. More research work is focused on minimizing the data using data compressing techniques. Data compression is used to decrease the number of bits required to transmit the data of particular information. It is useful because it helps people to minimize the use of resources such as data storage capacity. The goal of data compression is to eliminate redundancy in the data in order to reduce its size. This paper surveys different data compression techniques in WSN.

Keywords *:* Wireless sensor network; Data Compression;

A. INTRODUCTION

WSN is a key element of universal computing. It consists of tons of sensor nodes. Individual sensor node senses the environmental conditions like climate, atmospheric pressure and light and delivers the sensed information to a base station (BS). Since the sensor nodes are powered by fixed power batteries, low energy consumption is vital for sensor nodes, in order to improve the longevity of the network. WSNs are distributed computing systems that have various constraints like finite computing power and low memory. Typically, they can run on batteries, and have low frequency for transmission [8].

Usually a sensor node consists of three subunits: a sensing unit that is used to obtain the target events or exciting data' a processing unit prepared with partial memory that is used to handle the acquired data and a communication unit, usually a radio receiver that swaps the information among nodes [1]. Fig.1 represents the architecture of Wireless Sensor Networks. This architecture shows the traversing of sensor nodes from source to destination using gate ways.

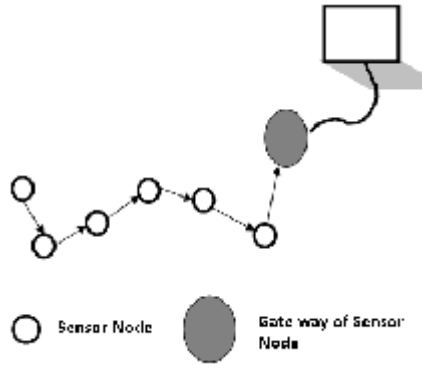

Figure 1: Sensor Node

At present, most of the research work shows that radio receiver is the major source of power utilization[11]. The network processing schemes solve the power utilization issues by minimizing the data which is to be transmitted by

[1]Research Scholar, Dept. of Computer Science, Karpagam Academy of Higher Education

[2]Associate Professor, Dept. of Computer Science, Karpagam Academy of Higher Education

means of aggregation techniques or Data Compression techniques [4].

Compression

Compression is the formation of coding. It will completely reduce the total number of bits, which will be essential to transmit data. It makes optimal use of finite storage volume to save time and helps to optimize the resources [10]. In general,

Data + Compression = information - redundancy

Compression ratio :

The term compression ratio means the ratio of the number of bits compressed before to number of bits compressed after. So after compression the data size of the sensed data will be small. This helps to transmit easily the information [9]. Figure 2 shows an architectural view of data compression method.

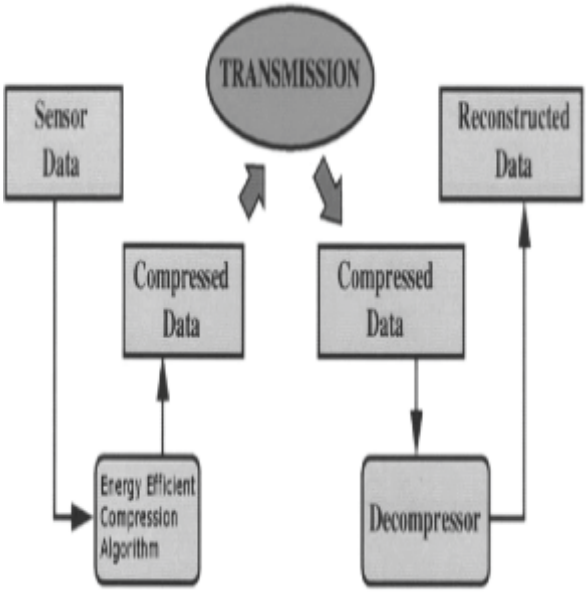

Figure 2: Data Compression Method

II. CLASSIFICATION OF DATA COMPRESSION

Data Compression is required for WSN applications, which have a huge amount of data to be sent across the network Depending on the importance and type of data, one may be better than the other. Data compression techniques include reducing the size of information before transmission. Decompression of data occurs at the base station. In compression it is important that no information is lost and independent readings are retained [10].

There are three categories of data compression namely lossless, lossy, and unrecoverable compression. A lossless compression means that after executing the decompression process, it can gain precisely the exact data as those which were previously executing the compression process. Huffman coding is one of the classical examples. A loss compression means that some detailed (and usually minor) features of the data may be lost due to the compression operation. Most of the image and video compression schemes such as JPEG2000 belong to this category. The unrecoverable compression means that the compression operation is irreversible and that there is no decompression operation. [2]. Figure 3 shows the process of lossless and lossy compression techniques.

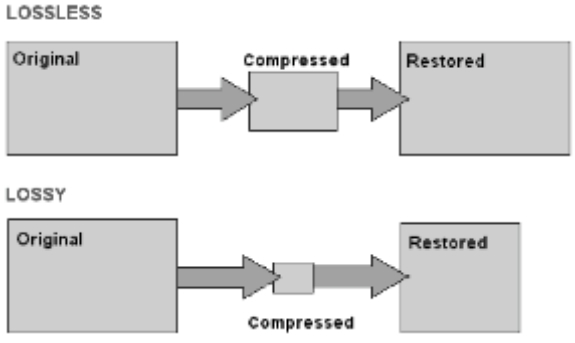

Figure 3: Lossless & Lossy Compression

III. COMPRESSION TECHNIQUES

This segment explains the major data compression techniques proposed in recent years for WSN's, and it focuses on the performance issues of each scheme [3].

The data compression techniques can be divided into five categories as follows:

- ***String-based compression techniques*** : This technique views the sensing data as a series of characters and then adopts the data compression strategy which will be used to handle text data to compress the sensing information. Inherited from these text data compression schemes, the string-based compression techniques can also contribute to lossless compression.
- ***Image-based compression techniques*** : This technique establishes a WSN into a hierarchical architecture and then adopts a few image compression schemes such as wavelet transformation to grant multiple resolutions of sensing data inside the network. Some minor features of sensing data may be lost due to the compression operations and thus the image-based compression technique supports the lossy compression.
- ***Distributed source coding techniques*** : This technique compresses sensing data within the network according to the Slepian-Wolf theorem, which proves that two or more related information streams can be encoded autonomously and then be decoded mutually at a receiver with a rate equal to their mutual entropy. So, the distributed source coding techniques can support lossless compression.
- ***Compressed sensing techniques*** : This technique suggests that compressible data can be exactly recovered from an adequate data of non-adaptive, randomized linear projection patterns. It can develop compressibility without relying on any past information or assumption on sensed data. This technique can provide lossless compression.
- Data aggregation techniques: This technique selects a subset of sensor nodes to collect and fuse the sensing data sent from their neighboring nodes and then transmits the small-sized aggregated data to the sink. Because the primary sensing data cannot be derived from these aggregated data, the compression of the data aggregation techniques is unrecoverable.

IV. COMPARISON OF DATA COMPRESSION TECHNIQUES IN WSN'S

To resolve the outcome of different kinds of data compression in wireless sensor networks, a survey was conducted and it is shown in table 1 [5].

Table 1. Comparison of data compression techniques in WSN

S. No	Compression Techniques	Compression Type	Method / Algorithm	Network structure or compression theory
1	String-based compression techniques	Lossless	Lempel-Ziv-Welch (LZW) algorithm LZW/ S-LZW	Sequence of Character
2	Image-based compression techniques	Lossy	Dimension / Multi-Resolution Compression and Query (MRCQ) framework	Hierarchic Architecture Wavelet Transformation

3	Distributed source coding techniques	Lossless	Slepian-Wolf theorem	Two or more correlated data stream.
4	Compressed sensing techniques	Lossless	Random projection	Non adaptive randomized linear Projection Samples
5	Data aggregation techniques	Lossy	Aggregation 1. Tree-Based 2. Cluster-Based 3. Chain-Based 4. Sector-Based	Aggregate data

V. EXISTING ALGORITHM

There are many existing algorithms for data compression in WSN. Some of them are as follows:

1. Basics of Information Theory
2. Run-Length Coding
3 Variable-Length Coding (VLC)
   - Shannon-Fano Algorithm
   - Huffman Coding
   - Adaptive Huffman Coding
4. Dictionary-based Coding
   - LZW Compression (Lempel Ziv Welch)
   - S-LZW Compression (Sensor-Lempel Ziv Welch)
5. Arithmetic Coding
6. Lossless Image Compression

The classification of the important algorithms for wireless sensor network is described in the following table 2 [6,7]

Table 2. Classification of algorithms

ALGORITHM	ADVANTAGES	LIMITATIONS
Shannon-Fano Algorithm	• Assigns code to each symbol based on its probability • Root-to-leaf method	• No unique code • Two different codes for the same symbol • Does not guarantee optimal solutions
Huffman Coding	• Easy to implement • Lossless technique • Produces optimal and compact code	• Relatively slow. • Depends upon statistical model of data. • Decoding is difficult due to different code lengths. • Overhead due to Huffman tree.

Adaptive Huffman Coding (OR) Dynamic Huffman	• Does not need the global statistics of the input data • Real time compression • 1 pass algorithm which makes the Huffman tree on the fly. code	• Both the coder and the decoder use exactly the same algorithm for updating code trees (otherwise decoding will not work !)
LZW Compression	• Does not need the global statistics of the input data • Real time compression • 1 pass algorithm which makes the Huffman tree on the fly. code	• Both the coder and the decoder use exactly the same algorithm for updating code trees (otherwise decoding will not work !)
LZW Compression	• Higher performance Data Compression • Simple and is dictionary based • Best technique for reducing the size of files containing more repetitive data	• Dictionary becomes too large • Dictionary reaches a certain size

VI. CONCLUSION

Since it is with the help of the battery that sensor nodes work, it is essential to preserve its energy for long. The use of data compression methods can be helpful to increase the life time of the sensor nodes. By compressing the data, the volume of the sensing data can be reduced. This paper has revealed a comprehensive survey on data compression algorithms and its techniques. Based on this survey, further study would investigate the proper compression algorithm to improve the energy efficiency of the sensor networks.

REFERENCES :

1. Shahina Sheikh, Ms. Hemlata Dakhore,"Data Compression Techniques for Wireless Sensor Network " Shahina Sheikh et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (1) , 2015, 818-821.
2. You-Chiun Wang," Data Compression Techniques in Wireless Sensor Networks" PERVASIVE COMPUTING
3. Dr. Shabana Mehfuz , Usha Tiwari "Recent Strategies of Data compression in Wireless Sensor Networks" Proc. of Int. Conf. on Advances in Electrical & Electronics, AETAEE Elsevier, 2013
4. A.Jeevarathinam, Dr.S.Manju Priya, "Comparison On Hierarchical Routing Protocols In Wsn "Karpagam JCS Vol.10 Issue 5 July - August 2016
5. You-Chiun Wang, "Data Compression Techniques in Wireless Sensor Networks", Pervasive Computing.
6. Emad M. Abdelmoghith, and Hussein T. Mouftah," A Data Mining Approach to Energy

Efficiency in Wireless Sensor Networks", IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications: Mobile and Wireless Networks, 2013.

7. M.A. Razzaque, Chris Bleakley, Simon Dobson, "Compression in Wireless Sensor Networks: A Survey and Comparative Evaluation", ACM Transactions on Sensor Networks, Vol. 10, No. 1, Article 5, Publication date: November 2013.

8. A Jeevarathinam, K Lakshmi, K Thilagam, K Rama," Overview Of Tenet: Architecture For Tiered Sensor Networks", International Journal of Engineering Science and Technology 1 (3), 379-387 ISSN : 0975-5462 Vol. 3 No. 1 Jan 2011

9. Ranganathan Vidhyapriya and Ponnusamy Vanathi "Energy Efficient Data Compression in Wireless Sensor Networks" The International Arab Journal of Information Technology, Vol. 6, No. 3, July 2009

10. Vandana Jindal, A. K. Verma & Seema Bawa "Impact of Compression Algorithms on Data Transmission "International Journal on Advanced Computer Theory and Engineering (IJACTE) ISSN (Print) : 2319 - 2526, Volume-2, Issue-2, 2013

11. Chetna Bharat Mudgule, Prof. Uma Nagaraj, Prof. Pramod D. Ganjewar "Data Compression in Wireless Sensor Network: A Survey"International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 2, Issue 11, November 2014