ISSN:0975-9646

Shahina Sheikh et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (1) , 2015, 818-821

Data Compression Techniques for Wireless Sensor Network

Shahina Sheikh#1, Ms. Hemlata Dakhore*2

#*Department of Computer science & Engineering*
G.H. Raisoni college of Engg & Tech. for Women
Nagpur, India

Abstract—**Data compression is an art used to reduce the number of bits required to transmit the data of particular information. The goal of data compression is to eliminate the redundancy in a data in order to reduce its size. Data compression can either be lossless or lossy. Lossless data compression recreates the exact original data from the compressed data while lossy data compression cannot regenerate the perfect original data from the compressed data. The relevance and importance of data compression has shown a growing trend from the past few years and is expected to continue remaining the same in the coming time. The proposed will provide an overview of popular adaptive data compression methods of general utility. It evaluated these existing algorithms with a newly proposed scheme with improvements to make the algorithm fully complete. The size data to be transmitted has an impact on the efficient working of WSN. This research work reports the simulation & implementation of efficient data compression algorithms for wireless sensor network.**

Keywords—Wireless sensor network; Data Compression; Huffman Algorithm

I. INTRODUCTION

Text databases are growing in the last years due to the widespread use of digital libraries, document databases and mainly because of the continuous growing of the Web. Compression comes up as an ideal solution that permits to reduce both storage requirements and input/output operations. Therefore, it is useful when transmitting data through a network. Even though compression appeared in the first half of the 20th century, in the last decade, new Huffman-based compression techniques appeared. Those techniques use words as the symbols to be compressed. They do not only improve the compression ratio obtained by other well-known methods (e.g. Ziv-Lempel), but also allow to efficiently perform searches inside the compressed text avoiding the need for decompression before the search. As a result, those searches are much faster than searches inside plain text.

Wireless Sensor Networks are popularly used to monitor a variety of physical environments. The network nodes gather data and process it to provide information in an appropriate manner as demanded by the primary application when required. Sensor Networks have some distinct features such as limited computation, distributed processing and degree of correlation. The gathered information is communicated to gateway nodes or sink nodes. Sensor nodes are designed with a limited battery power and they need to be unobtrusive. This mandates that they carry only a small battery, thus limiting the energy available for functioning and necessitating a low- power operation in order to conserve energy to prolong the useful life of the sensor node.

Radio communication is often the principal area of energy consumption. The extension of sensor node lifetime is generally achieved by reducing transmission/ receptions of data using data aggregation, compression techniques. The data compression algorithms are needed for these networks to reduce the amount of data communicated to the sink. The objective of the proposed work is to minimize the total number of bits required to be transmitted from the sensor node to reduce the energy consumed by the sensor node. The existence of temporal correlation in sensed data is considered as an advantage. A modified Huffman Data compression algorithm suitable for WSN is proposed in this synopsis. The proposed algorithm does not require the statistics of the sensed data though however encodes the difference of the current and the previous value of the sensed data. By this algorithm, a good compression ratio for both highly correlated and medially correlated sensor node data has to be achieved. The major problem in wireless sensor networks (WSN) is the limited energy supply. As the transmission of the data is the largest energy consumption, many studies have turned attention on reducing the amount of bits transmitted. The solution consists in data compression.

II LITERATURE REVIEW

Mo yuanbin, Qiu yubing, Liu jizhong, Ling Yanxia[1],"*A data compression algorithm based on adaptive Huffman code for Wireless Sensor Networks*", based on the principle of Adaptive Huffman Code, a novel compression algorithm for WSNs is proposed by making the full use of the phenomenon that the collected temperature data by the sensor nodes are continuous and slow changing in real environment. The Principle of the Adaptive Huffman code and the detailed procedure of the new algorithm are introduced in the paper. Unlike the predecessors to encode the data as a whole, the algorithm encodes the elementary characters in the difference value respectively. Based on the algorithm, compression experiments are carried out with Ambient Temperature Data, and the compression properties under the precision of 0.1, 0.01 and the mixed precision are analysed.

S. Renugadevi and P. S. Nithya Darisini [2], "*Performance comparison of Huffman & Lempel-Ziv Welch Data Compression for Wireless sensor node application*".The compression performance of the Huffman Algorithm & the LZW algorithm using various input data commonly measure by a wireless sensor node has been analysed. For the given tested data the Huffman Algorithm shows better performance when compared to the LZW in terms of compression ratio & computation time. From the experiment the Huffman algorithm is able to achieve an average of 43% data reduction. For double compression the LZH could provide up to 9% improvement in terms of data reduction but at the cost of an increase in computation time.

Mohamed Abdelaal and Oliver Theel [3], "*An Efficient and Adaptive Data Compression Technique for Energy Conservation in Wireless Sensor Networks"*. The Fuzzy transform has been proposed for a novel local adaptive data compression to minimize the bandwidth, the memory space, and the energy consumed in radio communication. An evaluation of the compression technique is provided. During this evaluation, the proposed technique is examined using real temperature data. The results have shown that the proposed technique can highly reduce the overall power consumption by up to 90%. Moreover, a modification of the proposed technique is presented which improves the accuracy of the recovered. A disadvantage of proposed technique, though, is the time delay between successive transmissions due to storing and processing the readings. Applications that require knowledge of recent or present measurements must wait for these readings to arrive.

Xi Deng and Yuanyuan Yang[4],*"Online Adaptive Compression in Delay Sensitive Wireless Sensor Networks"*. A new adaptive compression technique has been considered for lossless data compression which reduces the amount of data that must be passed through the network and to the sink and thus have energy benefits that are multiplicative with the number of hops the data travel through the network. Possible extensions in this paper, uses RUN LENGTH CODING to compress decompress the data in WSN. Static Huffman coding suffers from the fact that the uncompressed need have some knowledge of the probabilities of the symbol in the compressed files.

S. Renugadevi and P. S. Nithya *Darisini [5], "Huffman and Lempel-Ziv based Data Compression Algorithms for Wireless Sensor Networks"*.The effect of the compression on end-to-end packet delay in data gathering in WSNs has been studied in this paper. To accurately examine the effect of compression, author incorporated the hardware processing time of compression in the experiments by utilizing a software estimation approach to measuring the execution time of a lossless compressionalgorithm LZW on microcontroller TI MSP430F5418. While compression increases the maximumachievable throughput, it tends to increase the packet delay under light traffic loads and reduce the packet delay under heavy traffic loads.

A.V. R. Maheswari [6], "*Efficient Energy based Compression Techniques for Wireless Networks"*.The compression algorithm implemented in this paper the minimum variance Huffman compression algorithm, adaptive Huffman coding algorithm and variations of Huffman coding algorithm has been compared. From the analysis it is found that Huffman coding techniques are better than LZW techniques for achieving higher compression ratio in wireless sensor network. Among Huffman coding techniques and its variations, minimum variance Huffman coding technique performs better than others.

Francesco Marcelloni and Massimo Vecchio [7], *"A Simple Algorithm for Data Compression in Wireless Sensor Networks"*. In this paper the authors have discussed the use of static Huffman encoding to compress the data from a wireless sensor network. The proposed algorithm exploits the characteristics of high correlation of sensor data to compress the data. The data from the wireless sensor node is given to an ADC. The output of the ADC is digital or binary in nature. This binary output is sent to the entropy encoder which computes the difference between the successive values of the binary input. The difference values are encoded by using Huffman code from the table. The authors have considered the use of data collectors to increase the efficiency of the compression algorithm so that there is no requirement of knowing the statistics of the symbols prior to application of compression algorithms.

C.Tharini and P.VanajaRanjan [8], *"Design of Modified Adaptive Huffman Data Compression Algorithm for WSN"*.Authors have discussed the different variants of Huffman coding and advantage of one variant over the other in wireless sensor networks.

HuanZhang,Xiao-Ping Fan,Shao-Qiang Liu, ZhiZhong [9],*"Design and Realization of Improved LZW Algorithm for Wireless Sensor Networks.*"LZ 77, LZ78 and LZW algorithms are discussed for lossless data compression. LZ77 is the first of the LZ family and is about data compression when there is no prior Knowledge of source characteristics. The main drawback of this algorithm is its susceptibility to error propagation in the case of channel error. An improvement of LZ78 has the advantage of not transmitting the next non-matching symbol which is not the same with LZ78. LZW algorithm is applied to numeric data from wireless sensor network.

III. MEASURE OF PERFORMANCE

A compression algorithm can be evaluated in a number of different ways. We could measure the relative complexity of the algorithm, the memory required to implement the algorithm, how fast the algorithm performs on a given machine, the amount of compression, and how closely the reconstruction resembles the original. In this paper we will mainly be concerned with the last two criteria. A very logical way of measuring how well a compression algorithm compresses a given set of data is to look at the ratio of the number of bits required to represent the data before compression to the number of bits required to represent the data after compression. This ratio is called the compression ratio.

IV. ADAPTIVE HUFFMAN CODING

Huffman coding requires knowledge of the probabilities of the source sequence. If this knowledge is not available, Huffman coding becomes a two-pass procedure: the statistics are collected in the first pass, and the source is encoded in the second pass. In order to convert this algorithm into a one-pass procedure independently developed adaptive algorithms to construct the Huffman code based on the statistics of the symbols already encountered. These improved Theoretically, if we wanted to encode the (k+1)-th symbol using the statistics of the first k symbols, we could recompute the code using the Huffman coding procedure each time a symbol is transmitted. However, this would not be a very practical approach due to the large amount of computation involved—hence, the adaptive Huffman coding procedures.

The Huffman code can be described in terms of a binary tree similar to the ones shown in Figure 1. The squares denote the external nodes or leaves and correspond to the symbols in the source alphabet. The codeword for a symbol can be obtained by traversing the tree from the root to the leaf corresponding to the symbol, where 0 corresponds to a left branch and 1 corresponds to a right branch. In order to describe how the adaptive Huffman code works, we add two other parameters to the binary tree: the weight of each leaf, which is written as a number inside the node, and a node number. The weight of each external node is simply the number of times the symbol corresponding to the leaf has been encountered. The weight of each internal node is the sum of the weights of its offspring. The node number y_i is a unique number assigned to each internal and external node. If we have an alphabet of size n, then the 2n−1 internal and external nodes can be numbered as y1………..,y2n−1 such that if x_j is the weight of node y_j , we have x1≤ x2≤ · · · ≤ x2n−1. Furthermore, the nodes y2j−1 and y_{2j} are offspring of the same parent node, or siblings, for 1 ≤ j < n, and the node number for the parent node is greater than y2j−1 and y2j . These last two characteristics are called the sibling property, and any tree that possesses this property is a Huffman tree.

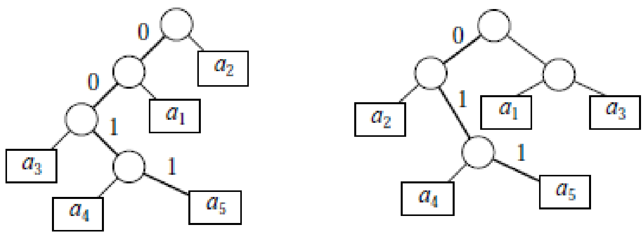

Figure1. Two Huffman trees corresponding to the same probabilities

In the adaptive Huffman coding procedure, neither transmitter nor receiver knows anything about the statistics of the source sequence at the start of transmission. The tree at both the transmitter and the receiver consists of a single node that corresponds to all symbols not yet transmitted (NYT) and has a weight of zero. As transmission progresses, nodes corresponding to symbols transmitted will be added to the tree, and the tree is reconfigured using an update procedure. Before the beginning of transmission, a fixed code for each symbol is agreed upon between transmitter and receiver.

When a symbol is encountered for the first time, the code for the NYT node is transmitted, followed by the fixed code for the symbol. A node for the symbol is then created, and the symbol is taken out of the NYT list.

Both transmitter and receiver start with the same tree structure. The updating procedure used by both transmitter and receiver is identical. Therefore, the encoding and decoding processes remain synchronized.

IV.1 Update Procedure

The update procedure requires that the nodes be in a fixed order. This ordering is preserved by numbering the nodes. The largest node number is given to the root of the tree, and the smallest number is assigned to the NYT node. The numbers from the NYT node to the root of the tree are assigned in increasing order from left to right, and from lower level to upper level. The set of nodes with the same weight makes up a block. Figure 2 is a flowchart of the updating procedure.

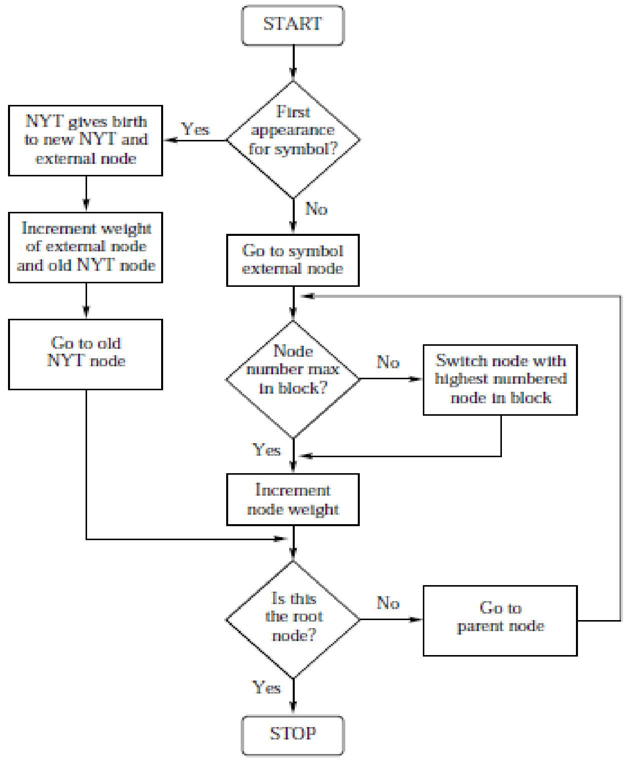

Figure 2. Update Procedure

TABLE I. COMPRESSION RATIO OF PROPOSED

Huffman Based Algorithms	Original bits	Compressed bits	Compression ratio
Static Huffman coding From [1]	32016	11099	65.3330 %
Minimum Variance Huffman coding	32016	7976	75.0875 %
Adaptive Huffman coding	32016	8149	74.5471 %
Variation 1	32016	8580	73.2009 %
Variation 2	32016	856	97.3263 %
Variation 3	32016	16348	48.9380 %

V. CONCLUSION

In this paper, based on the principle of Adaptive Huffman Code, a novel compression algorithm for WSNs is proposed. The principle of the Adaptive Huffman code and the detailed procedure of the new algorithm are introduced in the paper.

Unlike the predecessors to encode the data as a whole, the Algorithm encodes the elementary characters in the difference value respectively. Based on the algorithm, compression experiments are carried out.

REFERENCES

[1] Mo yuanbin, Qiuyubing, Liu jizhong, Ling Yanxia "A data compression algorithm based on adaptive Huffman code for Wireless Sensor Networks" 2011 Fourth International Conference on Intelligent Computation Technology and Automation

[2] S. Renugadevi and P. S. NithyaDarisini "Performance comparison of Huffman & Lempel-Ziv Welch Data Compression for Wireless sensor node application"American Journal of Applied Sciences 11 (1): 119-126 -2014

[3] Mohamed Abdelaal and Oliver Theel "An Efficient and Adaptive Data Compression Technique for Energy Conservation in Wireless Sensor Networks" IEEE –December 2-4 ,2013

[4] Xi Deng andYuanyuanYang"Online Adaptive Compression in Delay Sensitive Wireless Sensor Networks" IEEE Computer SocietyAug 2011

[5] S. Renugadevi and P. S. NithyaDarisini "Huffman and Lempel-Ziv based Data Compression Algorithms for Wireless Sensor Networks" International Conference on Pattern Recognition, Informatics and Mobile Engineering , 2013

[6] A.V. R. Maheswari "Efficient Energy based Compression Techniques for Wireless Networks" International Journal of Advanced Information Science and Technology, Vol.8, Iss.8, Dec2012

[7] FrancescoMarcelloni and Massimo Vecchio, "A Simple Algorithm for Data Compression in Wireless Sensor Networks", IEEE Communications Letters, vol. 12, no. 6, June 2008.

[8] C.Tharini and P.VanajaRanjan, "Design of Modified Adaptive Huffman Data Compression Algorithm for WSN ", Journal of Computer Science, June 2009.

[9] HuanZhang,Xiao-Ping Fan,Shao-Qiang Liu, ZhiZhong, "Design and Realization of Improved LZW Algorithm for Wireless Sensor Networks" ,International Conference on Information science and Technology, March 26-28, 2011,Nanjing, China.