Modified Distributed Source Coding using Syndromes for WSNs

Adish Jain (1), Doshi Yashkirti(2), Patil Abhishek Sanjay(3), Patil Amit Mahaveer(4)

Department of Electronics & Communication Engineering
Bahubali College of Engineering, Shravanabelagola, Hassan Dist

Abstract

Wireless Sensor Networks (WSN) have become a very interesting area for researchers. WSN is a network of small devices called sensors which sense the physical entity from the surrounding and convey the same to a central entity called the Base Station (BS). Sensor nodes are powered by batteries. As WSNs are generally deployed in hostile environments, recharging or replacement of the batteries is infeasible. Thus the data communication in WSNs has to be energy efficient. Source coding is one such technique used to achieve the better energy efficiency by exploring the correlation of data in the network. As the BS is not energy constrained, it can be over burdened than the sensor nodes. This fact is used to develop a Distributed Source Coding (DSC). Many DSC algorithms were proposed in the literature. Distributed Source Coding Using Syndrome (DISCUS) is one of the most popular algorithms used for DSC. In this paper we have made an attempt in modifying the DISCUS algorithm so as to use it for larger networks.

Keywords: Wireless Sensor Networks, Distributed Source Coding, Distributed Source Coding Using Syndrome.

1. INTRODUCTION

Wireless Networks plays a very important role in daily life for example cell phone, Bluetooth, Hiper LAN, Computer connection, Wi-Fi, etc. Wireless Networks are highly distributed networks of small light weighted nodes. It is having a flexible connection. It is a network that uses radio waves as carrier for communication. It gives us a remote operation. It provides a powerful communication for remote areas.

Wireless Sensor Network (WSN) is a network of sensors. Sensor is the device which connects physical world to digital world by sensing the analog data from surrounding area, converts it in the form such a way that it can be processed and sends data to the BS [1, 2].

Figure 1 shows a typical WSN where a CH is associated with 4 nodes. As shown in the figure, the nodes are collecting the data from surrounding and send to the Cluster Head (CH). CH is one of the sensor nodes which is chosen by the network itself based on the energy level of the node. CH collects all the data from other nodes, compresses it and then sends it to the BS. CH ask the node for the data by using poll command and if the node is having any data to transmit then it send data to CH otherwise ignore the poll command.

In WSNs, numbers of nodes are deployed in a geographical area. Generally WSNs are deployed for specific applications like animal tracking, temperature monitoring, pressure or humidity measuring, etc. As the nodes are densely deployed in most of the cases, a sort of correlation exists amongst the data sensed by the nodes. Generally nodes are placed in hostile environments where the battery replacing or recharging is not possible. Under such condition energy efficient transmission is required. Distributed Source Coding (DSC) is one of the efficient ways of achieving an energy efficient transmission in WSNs [3, 4, 5].
Coding for physically separated sources in a joint manner with little or no mutual communication for energy efficient data transmission is called DSC. Thus DSC is the compression of multiple correlated information. In DSC information from all nodes is gathered at one of the nodes having more energy, compression of data is performed using correlation between neighboring nodes thus reducing the data size and then will be conveyed to the BS. In WSN the nodes do not communicate with each other, hence the name DSC [6, 7].

To motivate DSC, let us consider an example for the temperature monitoring of a particular area where four sensors are placed say P, Q, R and S. As the nodes are placed in same area, the data sensed by the nodes have correlation. Let the number of bits required to represent the temperature is 8 bit. If the direct transmission is used, we require 4*8= 32 bits. However as the data is correlated the difference is temperature amongst the neighboring nodes can be efficiently represented using 2 bits. Thus if the first data is sent as a whole and the for the remaining sensors if only the difference is sent, then we require just (8+2+2+2)=14 bits instead of 32 bits.

The figure 2 depicts an example of a DSC. Here four nodes P, Q, R and S are considered and they are expected to have a high degree of spatial correlation. Each node has data to be communicated. If each node establishes an independent connection to the BS, lot of energy is required for the communication. Instead if node P is considered as a CH, thus aggregating the data from the nodes Q, R and S and sends an aggregated data to the BS.

II. EXISTING IMPLEMENTATION OF DISCUS

Distributed Source Coding using Syndromes (DISCUS) algorithm is one of the most popular DSC algorithms used to compress the data. In DISCUS syndromes are used as a key feature for data fusion [8, 9].
Let us consider a simple example. Consider two distributed sources X and Y which generate equiprobable 4 bit binary data. The data of X and Y highly correlated and can differ in at most one bit position. In other words, the Hamming distance between X and Y is less than or equal to 1. The possible data sequences generated by X are {0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111}. Thus we need 4 bits to transfer X to the decoder, but if Y is available at the decoder side say {1100}, sending 4 bits for X is not required. Instead a 3 bit index can be sent using which a probable set of estimation for X can be located. A grouping can be done such that the distance between the entries in the set is of maximum distance {(0000, 1111), (0001, 1110), (0010, 1101), (0011, 1100), (0100, 1011), (0101, 1010), (0110, 1001), (0111, 1000)}. Each of the entry in the set can be indexed as {000}, {001}, {010}, {011}, {100}, {101}, {110}, {111} respectively. Let us consider the transmitted vector Y as {1100} and the index for X as {011}. Thus the probable set of X as {0110}. Since the Hamming distance between Y and X is at most 1, the proper value of X can be estimated depending upon its distance from received Y as {1101}.

The above idea can be easily modeled using a Linear Block Code (LBC) which is represented in the form of (n, k, d_{min}), where n is the length of the data, k is a design parameter and d_{min} is the minimum distance of the code. For any integer ‘m’ > 1, we can have n= 2^m-1 be the length of the data to be transmitted by the sender. Then we can always represent it as an (n, n-m) LBC whose generator matrix is given by [I_k : P] where k= n-m. A parity check matrix H can be computed from matrix G as [P^T : I_{n-k}]. A standard array for the proposed (n, k) LBC can be computed where each coset consists of 2^k , n-tuples and there will be 2^m such cosets. The first element of each coset is called as a coset leader. Then the syndrome is computed by multiplying the coset leader with H^T matrix. Each coset can be uniquely represented using an ‘m’ tuple syndrome.

The encoding of data is done as follows. The first data will remain as it is and other data will be encoded using syndrome [10, 11]. The syndrome bit will be calculated by multiplying the data with H^T matrix. This syndrome bit along with the data will be sent to the receiver.

At the decoding side the received block of data containing a concatenation of a data and a syndrome. Using the received syndrome coset can be identified. After finding the coset, proper decoding is done depending on the Hamming distance property i.e. Hamming distance between two data must be less than or equal to 1. Each element of the coset is compared with the data and which is having least Hamming distance that will be consider as the original data.

Hence the better compression is achieved by sending an (n-k) bits of syndrome instead of an ‘n’ bit data.
III. MODIFIED IMPLEMENTATION OF DISCUS

In DISCUS algorithm the generation of data from two sources has been considered, but the scenario with more number of sources was not considered. In this paper we have implemented for the same the scenario for networks with more number of nodes. Here we have considered more than two correlated sources. The Hamming distance between the data is considered to be less than or equal to 1. The data of one of the nodes is sent as it is however for other nodes the corresponding syndrome bits are sent.

Consider an example of a four distributed sources P, Q, R and S which generate equiprobable 4 bit binary data. These data are correlated in a manner such that the Hamming distance of S with P, Q and R is less than or equal to 1. Let the codes for P, Q and R are {0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111}. We need 4 bits for each to transfer P, Q and R to the decoder, but if S is available at the decoder side say {1100}, sending 4 bits for P, Q and R are not required. Instead a 3 bit index can be sent using which a probable set of estimation for P, Q and R can be located. A grouping can be done such that the distance between the entries in the set is of maximum distance {(0000, 1111), (0001, 1110), (0010, 1101), (0011, 1100), (0100, 1011), (0101, 1010), (0110, 1001), (0111, 1000)}. Each of the entry in the set can be indexed as {000}, {001}, {010}, {011}, {100}, {101}, {110}, {111} respectively. Let us consider the transmitted vector S as {1100} and the index for P, Q and R as {011, 100, 010} respectively. Thus the probable set of P, Q and R as {0011, 1100}, {0100, 1011} and {0010, 1101} respectively. Since the Hamming distance S with P, Q and R is at most 1, the proper value of P, Q and R can be estimated depending upon its distance from received S as {1100, 0100, 1101} respectively.

IV. RESULTS AND ANALYSIS

A comparison of the obtained compression ratio from our algorithm for various number of nodes is considered. Figure 5 shows a plot of compression rate achieved for more than two sources. From the figure it is clear that increasing number of nodes reduces the achievable compression ratio for a particular value of ‘m’. Hence for a constant value of m, the compression decreases with the increasing number of data. Moreover, for a particular number of nodes, the compression ratio still decreases for increasing the value of m. Thus the algorithm performs reasonably well for a small set of nodes transmitting the data simultaneously.

V. CONCLUSION

DISCUS is one of Distributed Source Techniques used to compress the correlated data generated from two correlated sources. In our paper we have made an attempt to implement the DISCUS algorithm for WSNs. Also an improvement over the same is made by considering more than two correlated sources generating correlated data having a Hamming distance of less than or equal to 1. The results are validated with simulation. From the simulations, it is shown that a better compression is achieved for smaller networks.

![Number of Nodes Vs Compression Ratio](image-url)
ACKNOWLEDGEMENT

We would like to take this opportunity to thank one and all who have provided their valuable advice, without their guidance this work would not have been a success, we have to thank who have helped us directly or indirectly since they have given us more than just guidance.

We want to thank to Prof. K S Shivaprakasha, Associate professor, Department of Electronics and Communication Engineering, Bahubali College of Engineering, Shravanabelagola, who have guided us in our work and enrich it with his faithful experience.

Our profound thanks to Mrs. Ramamani K, Head of the Department, Department of Electronics and Communication Engineering, Bahubali College of Engineering, Shravanabelagola, for her invaluable advice and constant encouragement to complete this work in a successful manner.

REFERENCES


AUTHOR’S PROFILE

Adish Jain
Final Year, Electronics and Communication Engineering, Bahubali College of Engineering, Shravanabelagola.

Doshi Yashkirti
Final Year, Electronics and Communication Engineering, Bahubali College of Engineering, Shravanabelagola.

Patil Abhishek Sanjay
Final Year, Electronics and Communication Engineering, Bahubali College of Engineering, Shravanabelagola.

Patil Amit Mahaveer
Final Year, Electronics and Communication Engineering, Bahubali College of Engineering, Shravanabelagola.