An efficient communication protocol for wireless sensor network using differential encoding based compressed sensing technique

A. Parnasree Chakraborty, B. Tharini Prasad

***Abstract*—**Wireless sensor networks (WSNs) are typically resource constrained network due to restricted parameters like power supply, processing speed, memory requirement and bandwidth required for communication. Energy consumption is a key issue in the design of protocols and algorithms for WSNs due to their limited power supply. WSN operations involve sensing of data, computation, switching from node to node, transmission etc. In all these operations, energy efficiency is very essential. It is found in literature that, most of the energy is consumed in WSNs is due to the radio communications. In radio communication if the number of bits of data to be transmitted is reduced by some amount then it is possible to reduce the energy consumption. Hence it is essential to use data compression to reduce the number of bits to be transmitted. Researchers have investigated many energy efficient light weight compression algorithms suitable for WSN data. Still there is a requirement for efficient compression algorithms for WSN which minimizes the mean square error (MSE) of received data and hence in this paper differential encoding based compressed sensing (CS) algorithm is suggested. A CODEC design is suggested for improving the reconstruction quality. Simulation results show improvement in reconstruction quality and reduction in MSE value compared to standard compressed sensing technique.

***Keywords*—**Wireless Sensor Networks, Compressed Sensing, Differential Encoding.

I. INTRODUCTION

WIRELESS sensor network [1] , consists of large numbers of sensors which collect sensor data. Physical data sometimes are highly correlated in nature and thus can be utilized to improve compression ratio. Before transmission it is essential to reduce the dimension or size of the measurements in order to improve the transmission efficiency [2] as in literature it is found that most of the energy consumption is due to radio transmission [3]. Many algorithms for data compression [4] in WSN are considered in literature with the objective to improve the reconstruction quality. This paper is aimed to discuss compressed sensing (CS) algorithm with differential encoding as a better compression algorithm suitable for WSN. A CODEC design based on Lloyd's algorithms also suggested for further improvement of compression ratio with reduced Mean square Error (MSE) for data transmission in WSN. The encoded data is compressed by CS and quantized using a quantizer which works on Lloyd algorithm. This suggested technique is suitable for any scenario in WSN, where continuous data transmission is necessary.

Parnasree Chakraborty is with BSA Crescent Institute of Science & Technology, Chennai, India, (email:prernasree@crescent.education).
Tharini Prasad. is with BSA Crescent Institute of Science & Technology, Chennai, India,(email:tharini@crescent.education)

II. RELATED WORK

In [5], the authors have presented two methods for carrying out compressed sensing with quantized measurements: the regularized maximum likelihood, and another method based on regularized least squares. Results of numerical simulations show that both methods work relatively well, however the effectiveness of the methods are not discussed with respect to Mean Square Error (MSE).

In [6], the authors have proposed a simple and efficient CS encoder device to measure signals within sensor nodes of a WBAN. As the CS encoder and decoder are tightly coupled, a model of the overall acquisition chain is used in the first stages of development and validation. A SPICE model and a hardware prototype of the proposed CS encoder are also presented. This paper discusses the compressed sensing encoder useful for biomedical signal only however, it does not discusses regarding CS encoder useful for the other signals.

In [7], presented an effective compressed sensing based prediction measurement (CSPM) encoder compatible for wireless multimedia sensor networks. The compression performance of CSPM method is evaluated using metrics such as compression ratio and bit rate. The video is reconstructed by the orthogonal matching pursuit algorithm, however this paper discusses only about video signal encoding process.

In [8], compressed sensing (CS) algorithm is used for data compression in wireless sensors to address the energy and telemetry bandwidth constraints common to wireless sensor nodes. Circuit models of both analog and digital implementations of the CS system are presented that enable analysis of the power/performance costs associated with the design space for any potential CS application, including analog-to-information converters (AIC).

To upgrade the quality of CS reconstruction, in [9] a proposal is suggested on CS area, which includes gradient sparse operators and CS optimized reconstruction with combination blocks. The advantage of this algorithm is that no change is needed on the encoder side and the improvement is focused on the decoder side. This paper is fully concentrated on decoder design however no encoder design is explained here.

From the above literature survey it is obvious that even though CS based compression is used for data compression in WSN, there is no such CS algorithm to improve the sparsity of the data for better reconstruction quality. Hence the suggested differential encoding based CS is aimed to work in the direction to improve the sparsity of the sensor data which in turn improves the reconstruction quality by reducing the Mean Square Error (MSE).

III. SYSTEM DESCRIPTION

The system model for the proposed method for compression of sensor data in WSN is given in Figure 1.

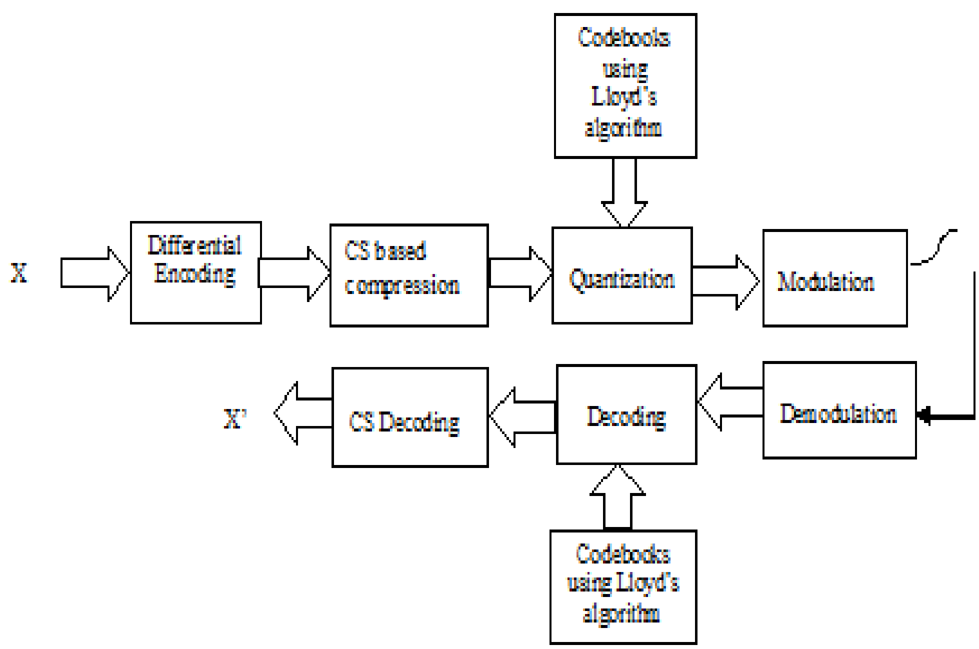

Fig.1 System Description

Explanation of the block diagram is given in following section. In Fig.1, x denotes input signal and x' denotes the reconstructed signal.

I. DIFFERENTIAL ENCODING TO INCREASE THE LEVEL OF SPARSITY

The input signal is a real time data (temperature data) obtained from the database and if we look into the data we can easily find that the data samples are highly correlated in nature. In order to increase the sparsity of the signal, difference sequence or delta sequence is generated by finding the difference between the consecutive samples of input data. After obtaining the difference signal the CS based compression is performed. CS based compression relies on two principles: Sparsity [9] and Incoherence.

Sparsity refers to that only very few entries in a vector or matrix is non zero. Incoherence says that unlike the signal of interest, the sampling/sensing waveforms have an extremely dense representation in transform domain and hence by exploiting these two principles of CS the reconstruction quality can be improved. In this paper the sparsity of the input measurements are improved by considering the difference of input samples rather than considering the raw input samples. In general L1 norm of a vector has the property of finding many coefficients with zero values or small values with very large coefficients and hence L1 norm can be used to find the level of sparsity of any measurement vector. In this paper the level of sparsity is measured with respect to raw input samples as well as difference signal and performance is compared.

II. CS ENCODING AND DECODING

The input data X is generated from pre-processing the real time sensor data. Here the temperature sensor values are collected from database. The data is differential encoded and converted in to a sparse signal using a basis function. Here discrete cosine transform is used as the basis function; the sparse signal is compressed using a random sensing matrix T. Equation (1) describes generation of the sparse signal (s).The compressed signal is then quantized using codebooks and appropriate indices are transmitted to the decoder.

$$s = Tx \tag{1}$$

In equation (1), (x) is the input signal of size (n x 1), (T) is the transform matrix or kernel of size (nxn) and (s) is the sparse representation of size (n x 1). After sparse generation, measurement matrix (M) of size (m x n) is used to compress the sparse representation's' of input image and compressed sparse output 'y' of size (m x 1) is generated. The compressed sparse output (y) is given in equation (2).

$$y = Ms \tag{2}$$

The design of the encoding algorithm is generic and any suitable reconstruction algorithm can be used. In this work, Orthogonal Matching Pursuit (OMP) is used as CS reconstruction algorithm.

III. QUANTIZATION USING NEAREST NEIGHBOUR CODING

Quantization involves mapping of samples from a continuous set to a discrete alphabet using fixed codebooks. The quantization is based on the concept of nearest neighbor coding. In this type of quantization, each entry of measurement vector is coded to its nearest code point. Therefore, given a fixed codebook associated with the measurement entry, the nearest-neighbour quantizer uses the encoding rule which minimizes the MSE. However, this approach does not necessarily guarantee that the end-to-end MSE is also minimized subject to fixed codebook sets. This is due to non-linear behaviour of the sparse reconstruction function and non-orthogonality of the CS sensing matrix. At the decoder, the fixed codebooks are used for decoding which is shown in Fig.1.

IV. CODEBOOK DESIGN BASED ON LLOYD'S ALGORITHM

In this work a data compression technique based on quantized CS is proposed, where signals are compressed by compressed sensing and then the compressed measurements are quantized and the quantization parameter are transmitted from transmitter to the receiver. In quantization process the input signal values from a large set is mapped into output values in a smaller set. Quantization parameters are mainly categorized as a partition and a codebook. Quantization of signal produces distortion and the distortion can be reduced by choosing appropriate partition and codebook parameters. In order to optimize the quantization parameters, the Lloyd's function is used in this work. The Lloyd's function optimizes the partition and codebook according to the Lloyd algorithm [10].

A quantization partition and codebook is explained here with the help of example. Partition in quantization is nothing but several neighboring, non-overlapping ranges of values within the set of real numbers. The concept of partition can be explained using an example. If the partition separates the input signal or real number line into the four sets as [$x(n)$: $x(n) \le 0$], [$x(n)$: $0< x(n) \le 1$], [$x(n)$: $1< x(n) \le 3$] and [$x(n)$: $3< x(n)$], then one can represent the partition as the three-element vector and hence partition = [0, 1, 3].

A codebook tells the quantizer which common value to assign to inputs that fall into each range of the partition. For the above example, the codebook can be given by [-1, 0.5, 2, 3].This codebook is one possible codebook for the partition [0, 1, 3]. There are many possible codebooks for the example explained here but the optimum codebook can be designed using the Lloyds algorithm. The Lloyd's algorithm steps are discussed below.

First, assuming that the codebook (yi) which is to be generated using Lloyd's algorithm are fixed. The input x is obviously nearest to one of the representative levels. The quantization result of the input x can be at most as small as that minimum distance.

Table.1 Comparison of l1 norm value and the level of sparsity

Data Size	**Type of Input Signal**	**L1 norm value**	**Level of Sparsity or Sparseness of (x)**
Data Set-1(1000 samples)	Input Data without differential encoding	$4.59 \text{ X } 10^{4}$	$2.1486 \text{ X } 10^{-6}$
	Input Data with differential encoding	$0.0023550 \text{ X } 10^{4}$	0.3209
Data Set-2(1200 samples)	Input Data without differential encoding	2249	$2.1425 \text{ X } 10^{-6}$
	Input Data with differential encoding	28.27	0.3199

Table.2 Comparison of the performance in terms of mean square error for normal data and difference input data with size of the codebook=8 with QAM Modulation

Data Size	**Types of Modulation**	**Compres-sion Ratio**	**MSE for normal input samples**	**MSE for difference input samples**
Dataset1 (1000 samples)	QAM	50%	2.0750	0.0161
Dataset2 (1200 samples)	QAM	50%	4.9079	0.0140

Table.3 Comparison of the performance in terms of mean square error for normal data and difference input data with size of the codebook=8 with QPSK Modulation.

Data Size	**Types of Modulation**	**Compression Ratio**	**MSE for normal input samples**	**MSE for difference input samples**
Dataset1 (1000 samples)	QPSK	50%	4.5443	0.0191
Dataset2 (1200 samples)	QPSK	50%	2.1303	0.0187

Table.4 Comparison of the performance in terms of mean square error for normal data and pre-processed data with size of the codebook=16 with QAM Modulation.

Data Size	**Types of Modulation**	**Compression Ratio**	**MSE for normal input samples**	**MSE for difference input samples**
Dataset1 (1000 samples)	QAM	50%	2.0383	0.0146
Dataset2 (1200 samples)	QAM	50%	4.8775	0.0139

Table.5 Comparison of the performance in terms of mean square error for normal data and pre-processed data with size of the codebook=16 with QPSK Modulation.

Data Size	Type of Modulation	Compression Ratio	MSE for normal input samples	MSE for difference input samples
Dataset1 (1000 samples)	QPSK	50%	4.5195	0.0136
Dataset2 (1200 samples)	QPSK	50%	4.587	0.0177

Table.6 Comparison of the PRD value for normal data and differentially encoded data with size of the codebook=16 with QAM Modulation.

Data Size	Types of Modula-tion	Compression Ratio	PRD in case of normal input samples	PRD in case of difference input samples
Dataset2 (1200 samples)	QAM	50%	12.4002	0.0339
Dataset3 (200 samples)	QAM	50%	14.01	0.2823

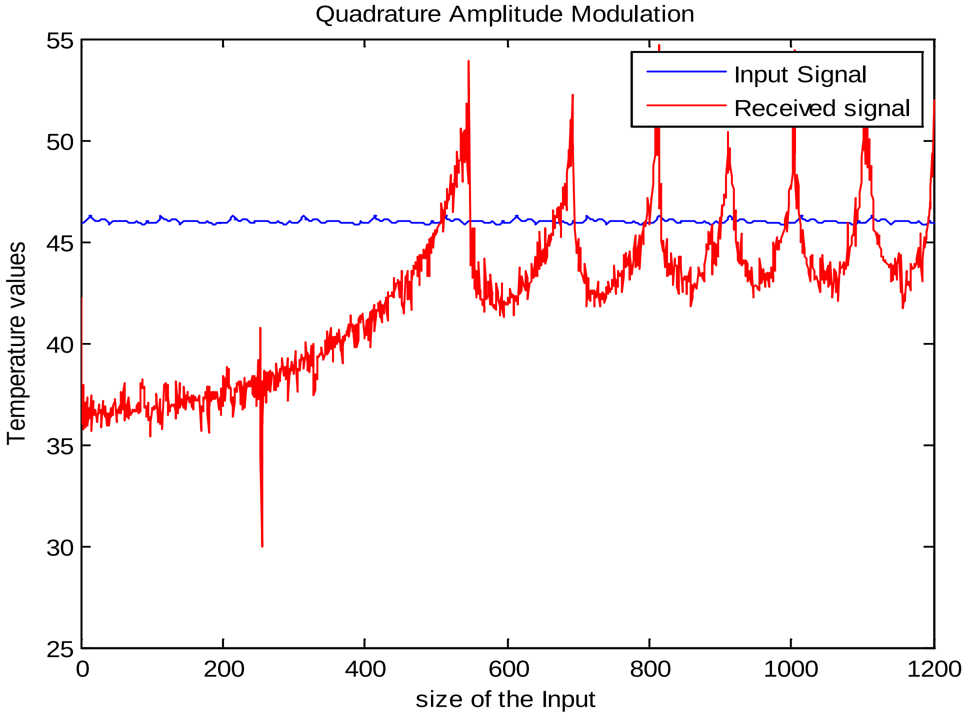

Fig.2 Input and Reconstructed (received) signal without differential encoding and QAM modulation

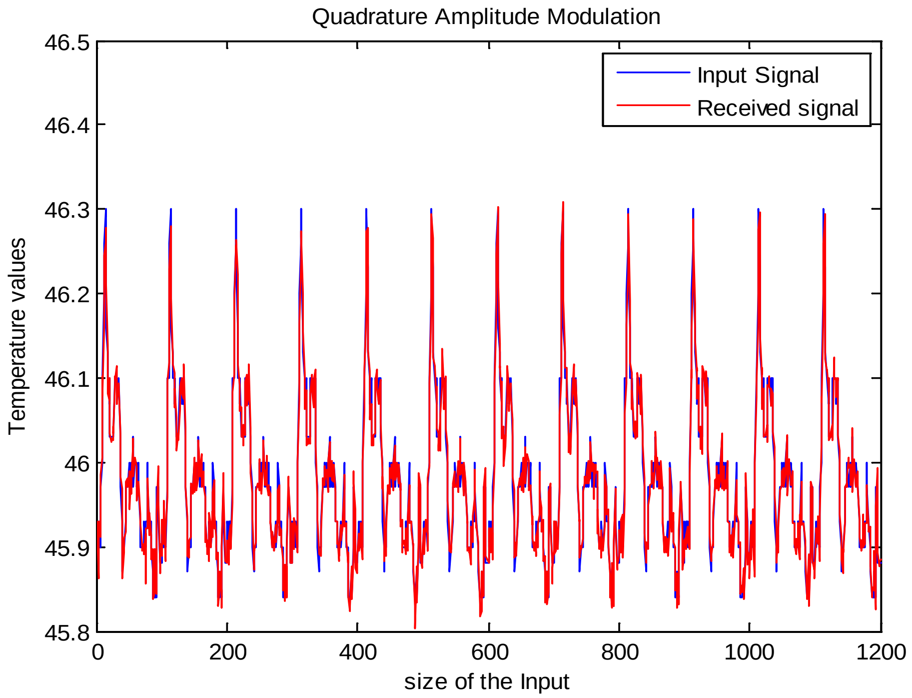

Fig.3 Input and Reconstructed (received) signal with differential encoding and QAM modulation

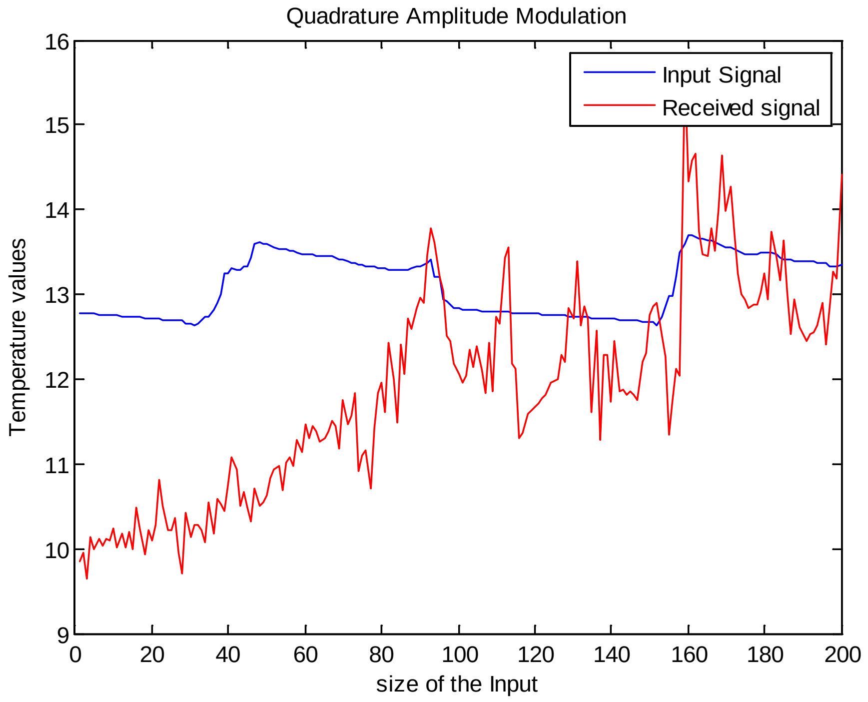

Fig.4 Input and Reconstructed (received) signal without differential encoding and QAM modulation

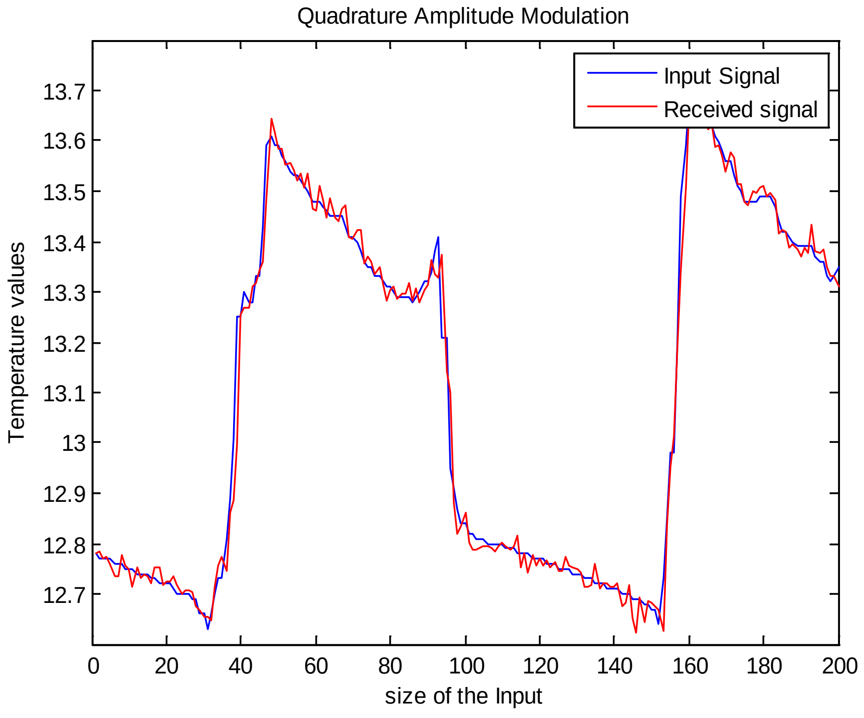

Fig.5 Input and Reconstructed (received) signal with differential encoding and QAM modulation

1. Start with initial representative levels y_i

2. Find the optimum interval boundaries (x_i) corresponding to the y_i's according to equation (3):

$$x_{i,\,opt} = 1/2[\, y_{i,\,opt} + y_{i-1\,,\,opt}\,]. \quad (3)$$

3. Recalculate new representative levels's according to the equation (4):

$$y_{i,\,opt} = \frac{\int_{xi,opt}^{x\ i+1,opt} x\, f(x)}{\int_{x\ i,opt}^{x\ i+1,opt} f(x)} \quad (4)$$

In above equation the probability density function (pdf) of input x is given by f(x).

4. If new y_i's are very different from old y_i's, then go to 2.

The above mentioned steps are repeated until the optimum quantization parameters are achieved.

V. MODULATION/ DEMODULATION

A proper modulation scheme can improve the bandwidth efficiency and energy efficiency of a WSN [15]. In a sensor network, the lifetime of the nodes depend on the energy consumption of transceivers. Therefore, an optimal selection of modulation techniques based on the characteristics of the communication channel improves overall system performance.

The compressed sparse signal is quantized using Lloyd's algorithm and required to be transmitted to the sink node .This is implemented by considering two different modulation schemes, i.e. QPSK and QAM. These two modulation schemes are widely used in a WSN. After quantization, the modulated data are transmitted considering the transmission channel as AWGN channel. QPSK modulation uses four different phases to distinguish binary 00, 01, 10 and 11.

QAM scheme is a combination of both analog and digital modulation. The two analog message signals, or two digital bit streams, are conveyed by changing (modulating) the amplitudes of two carrier waves, using the amplitude-shift keying (ASK) digital modulation scheme or amplitude modulation (AM) analog modulation scheme. The two sinusoids, are out of phase with each other by 90° and are thus called quadrature carriers or quadrature components. The final waveform is a combination of both phase-shift keying (PSK) and amplitude-shift keying (ASK), or (in the analog case) of phase modulation (PM) and amplitude modulation.

In this paper, the sensor readings are quantized using codebooks and the corresponding indices from the codebooks are transmitted to the decoder. The decoder uses the indices and decodes the values using the fixed codebooks available at the receiver side. The indices are modulated using any one of the modulation scheme mentioned above and transmitted to decoder. The decoder demodulates the received index values and decodes the transmitted data using codebooks. Fig 1 describes the system along with modulation.

The performance of the modulation schemes are analyzed considering ideal channel with addictive white Gaussian noise (AWGN). The quality of the reconstructed signal is measured in terms of mean square error.

VI. RESULTS AND DISCUSSION

The performance of the algorithm is evaluated in terms of mean square error for the different sizes of codebooks. The codebooks are designed and optimized using Lloyd's algorithm. The entries in the codebooks are optimized for the compressed signal. Two datasets are considered with size of 1000 (dataset1) and 1200(dataset2). The datasets are obtained from the database available in the literature. The real temperature sensor values are taken for analyzing the effectiveness of the algorithm. For measurement of level of sparsity there is no direct formula available but the level of sparsity can be measured by finding the L1-norm of the signal. If the value of L1-norm is less, then the level of sparsity is more [11] which is shown below in Table.1. One more parameter is used to define the sparsity [12] of the signal or vector given by equation (5).

$$\text{Sparseness of } (x) = \frac{\sqrt{k} - (\,\|x\|_1 / \|x\|_2\,)}{\sqrt{k} - 1} \quad (5)$$

Where 'x' is the vector whose sparseness to be measured and 'k' is the length of the vector. $\|x\|_1$, is the L1 norm and $\|x\|_2$ is the L2 norm. L1 norm or one norm of a vector **x** is given by $\|x\|_1 = \sum_{i=1}^{n} |x|$ (6)

and the L2 norm or two norm of a vector **x** is given by

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x|^2} \quad (7)$$

Whenever the vector is dense, the sparseness approaches to zero and if Sparseness of (x) approaches to 1, the vector is sparse [12]. Equation (5) can be explained better with the help of following example. Let the vector whose sparseness to be measured is given by x1= [1 2 3 4]; here the length of vector x is k= 4.Using equation (5) the Sparseness of (x1) is 0.1743. Let another vector x2= [1 2 3 0], in this case the Sparseness of (x2) is 0.3964.Sameway if another vector x3= [1 2 0 0] is considered, the Sparseness of (x3) is 0.6584. From this explanation it is clear that as the vector consisting more number of zeros the sparseness of the vector approaches towards '1'.

In this paper to increase the sparsity of the signal, differential sequence or delta sequence is generated by finding the difference between the consecutive samples of input data. Hence the number of zeros have been increased as the WSN dataset consist similar data value.

Table 2.and table 3. show that when the size of the input is 1000 samples, the MSE is 2.0750 for normal input data with QAM modulation, whereas with differential encoding the MSE is 0.0161.Thus the results shows that when 1000 measurements are compressed to 500 measurements, there is an reduction in MSE and quality of the reconstructed signal is improved with differential encoding based compressed sensing

technique. In case of another set of temperature reading of 1200 samples also shows the better result in case of differential encoding scheme with codebook size 8.

The size of the codebooks is based on the application and the memory availability. The performance results using 16 code points are tabulated in Table 4.and in Table 5. As the size of the codebook increases there is further reduction in the error. Fig.2 and Fig.3 show the reconstructed data without and with differential encoding with QAM modulation.

In Fig.3, the reconstructed data with differential encoding shows better reconstruction quality compared to without differential encoding shown in Fig.2. . The units of x-axis and y-axis values in fig.2, Fig.3, Fig.4 and Fig.5 are samples and degree Celsius respectively.

Another dataset consists 200 samples is also considered and reconstructed data with and without differential encoding is plotted below in Fig.4. and Fig.5.In this case also the result shows improved performance in proposed method.

The percentage of root-mean-square difference (PRD) between the reconstructed and the original signal is used as the evaluation index for reconstruction quality [13]. PRD is defined as:

$$PRD = \frac{[(\|(\hat{x} - x)\|_2).100]}{[\|x\|_2]} \quad (8)$$

Where $\hat{x}$ denotes the reconstructed signal and x denotes the original signal. $\|\cdot\|_2$, defines the L2 norm of the vector. The smaller the PRD, the better the reconstruction quality [14].

Table 6.shows the comparison of PRD values of data with and without differential encoding. It is very clear that the PRD value is nearer to zero which ensures the perfect reconstruction in case of differentially encoded data.

VII. CONCLUSION

The proposed communication protocol for wireless sensor network using differential encoding based compressed sensing technique is analyzed and the proposed algorithm improves the quality of the reconstructed data by reducing the mean square error value. Reconstructed signal quality is also measured by the percentage of root-mean-square difference (PRD) value. The PRD value between the reconstructed and the original signal is used as the evaluation index for reconstruction quality which is explained in this paper. The lesser the PRD value the better is reconstruction quality. The simulation result shown here confirms that with differential input signal the PRD is minimum whereas without differential input the PRD value is high. The numbers of measurements which are transmitted, is made sparse using compressive sensing technique. The compressed measurements are then quantized using Lloyd's algorithm and quantization parameters are transmitted using suitable modulation schemes. Not only the PRD value, the MSE is also found minimum in case of proposed technique. The comparison is done with normal data and differential data by keeping the compression ratios as constant. It is found from the result that, the reconstruction quality is improved using differential encoding based compressed sensing technique. Implementation of the proposed algorithm in hardware is suggested as future work.

REFERENCES

[1] Akyildiz I.F, Su W, Sankarasubramaniam Y, Cayirci E, "Wireless sensor networks: A survey," Elsevier, Computer Networks, Vol.38 2002, pp.393–422.

[2] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," Proceedings of the 33rd Annual Hawaii International Conference on System Sciences; Maui, HI, USA. 4–7 January 2000; pp. 3005–3014.

[3] Pottie G.J, Kaiser W.J,"Wireless integrated network sensors", Communications of the ACM, 43, 51-58.

[4] Razzaque M.A., Bleakley C., Dobson S, "Compression in wireless sensor networks: A survey and comparative evaluation," ACM Trans. Sens. Netw. 2013; 10:5:1–5:44.

[5] Argyrios Zymnis, Stephen Boyd, and Emmanuel Candès, "Compressed Sensing With Quantized Measurements," IEEE Signal Processing Letters, Vol. 17, No. 2, February 2010.

[6] Andrianiaina Ravelomanantsoa, Hassan Rabah, Amar Rouane, " Simple and Efficient Compressed Sensing Encoder for Wireless Body Area Network," IEEE Transactions on Instrumentation and Measurement ,Volume 63, Issue: 12, Dec. 2014.

[7] V. Angayarkanni Emailauthor,S. Radha, "Design of Bandwidth Efficient Compressed Sensing Based Prediction Measurement Encoder for Video Transmission in Wireless Sensor Networks," Wireless Personal Communications June 2016, Volume 88, Issue 3, pp 553–573.

[8] Fred Chen, Anantha P. Chandrakasan, and Vladimir M. Stojanović, "Design and Analysis of a Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors," IEEE Journal Of Solid-State Circuits, Vol. 47, No. 3, March 2012.

[9] Candès, E.J. and M.B. Wakin, "An introduction to compressive sampling," IEEE Signal Processing Magazine, 2008, 25: 21-30.DOI: 10.1109/MSP.2007.914731.

[10] S. P. Lloyd, "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory,* Vol IT-28, March 1982, 129-137.

[11] Available:https://math.stackexchange.com/questions/1904767/why-small-l1-norm-means-sparsity .

[12] Available:https://math.stackexchange.com/questions/117860/how-to-define-sparseness-of-a-vector .

[13] Zhang Y, Liu C, Wei S, Wei C, Liu F, "ECG quality assessment with a kernel support vector machine and genetic algorithm using a feature matrix," Journal of Zhejiang University Science C. 2014;15(7):654-573.

[14] Lalos AS, Alonso L, Verikoukis C, "Model based compressed sensing reconstruction algorithms for ECG telemonitoring in WBANs," Elsevier, Digital Signal Processing,2014.

[15] Parnasree Chakraborty and C. Tharini, "Analysis of suitable modulation scheme for compressive sensing algorithm in wireless sensor network," Sensor Review,Emerald Group Publishing Limited, 35(2),pp.168-173, DOI 10.1108/SR-06-2014- 666.