CHAPTER - 2

THRESHOLD BASED COMPRESSED SENSING ALGORITHM FOR
WIRELESS SENSOR NETWORKS

2.1 INTRODUCTION

Wireless Sensor Networks (WSNs) are comprised of sensor nodes which are spatially distributed for monitoring and recording of environmental conditions. Each node in WSN contains sensors, processors, power supply units and transceivers for communication. Irrespective of the field of WSN applications, the data gathered in nodes must be delivered to the sink node for further processing. However the limitations due to constrained bandwidth, frequent node failure, unreliable and insecure communications in WSN poses great challenges and enforces to develop energy efficient mechanisms to resolve these issues. One such solution is to develop data compression scheme for enhancement of the transmission efficiency in WSN. One such energy efficient data compression is Compressed Sensing (CS). This chapter focuses on threshold based CS on the one dimensional input data. Threshold based CS is used to increase the sparsity of the signal. The advantages of CS based compression is that, the signal is sparsified first and then the sparsified signal is converted to lesser dimension using CS encoding. The encoded signal is transmitted from node to node in WSN. Since the encoded signal is of much lesser dimension than the original signal, hence the number of bits to be transmitted in WSN is reduced which reduces the transmission cost [27].

2.2 RELATED WORK

The field of CS is existing since last four decades. The concept of CS is first introduces in the field seismology in 1970.In [28], Clarebout and Muir suggested attractive alternatives of Least square solutions.
In [29] and [30] the norm of random matrices are described which is useful for CS based reconstruction. The idea of l1 norm was introduced by Santosa and Symes in [31] during mid-eighties.

CS based reconstruction using total variation minimization in image processing was discussed by Rudin et.al in [32]. The concept of total variation minimization is similar to the concept of $l_1$ optimization.

Some more contributions of this era towards CS are given in [33], [34], [35], [36], [37] and [38]. The concept of Compressed Sensing got a new life when Donoho et.al listed the important results and mathematical formulation of CS in [18] in the year 2009. A series of contributions in the field of CS in last few years are witnessing achievements almost on a daily basis.

In [39], a matrix completion and CS based data accumulation in WSN by the sensor nodes are well discussed. Exploitation of CS based techniques in the form of low rank and sparse nature of readings is carried out in this work. A high compression ratio of readings is achieved by making use of CS based compression. The experimental results of collected practical dataset compression show the efficiency of the method suggested. However, this technique did not accommodate the data readings based on their correlation. To achieve this, the threshold based CS algorithm is proposed in this chapter.

2.3 INTRODUCTION TO COMPRESSIVE SENSING (CS)

A signal processing technique known as Compressive Sensing or Compressed Sensing (CS) is used in almost all the field of communication engineering in recent years. For efficient acquisition and reconstruction of a signal, CS can be used as a data reduction tool. In CS based reconstruction, underdetermined linear systems [40] are solved using linear programming and by finding solutions to the underdetermined system, perfect recovery of the signal is done. In [19], it is explained very well how the CS theory can be used to recover signals from very few number of samples (measurements) than traditional and conventional methods. The working principle of CS is
completely different than the normal sampling theorem. The normal sampling theorem says that for successful reconstruction of any signal, the sampling rate must be at least twice the number of highest frequency component present in the signal [41]. According to CS theory, the conventional sampling theorem need not be followed.

To make this possible, CS relies on two principles: one is sparsity and the other principle is incoherence. Sparsity [40], depends on the density of the signal and incoherence, which depends on the sensing modality. Both the principles are discussed in detail as follows.

Sparsity: A sparse matrix is a matrix or a sparse array is an array in which few elements are non-zero and most of the elements are zero. On the other hand, most of the elements are nonzero in a dense matrix. The parameter sparsity can be found by finding the number of zero-value elements divided by the total number of elements in a matrix or array. The sparsity can also be found by subtracting the density of the matrix from unity. In general many signals are sparse or compressible by nature. This behaviour of signal is exploited by CS theory. The meaning is, the signals have sparser representations when they are expressed in the proper basis. Mathematically if we have a vector ‘x’ of length ‘N’ (i.e. \( x \in \mathbb{R}^N \)), then the vector ‘x’ can be expanded using an orthonormal basis \( \Psi = [\Psi_1, \Psi_2, \ldots, \Psi_N] \) using the equation (2.1).

\[
s = \sum_{i=1}^{N} x_i \Psi_i
\]  

(2.1)

Where sparse representation of input signal ‘x’ is denoted by ‘s’. The number of non-zero components in the vector ‘s’ are much lesser than the non-zero components in vector ‘x’ if the spreading basis is sparse. From equation (2.1) it is clear that when a signal has a sparse representation, small coefficients can be discarded without much perceptual loss.

Incoherence: Incoherence principle is extended by the duality between time and frequency. It expresses the idea that objects having a sparse representation in a basis ‘\( \Psi \)’ must be spreaded out in the domain in
which they are acquired. The concept of incoherence can be made much clear with the help of a spike signal or impulse signal. In the time domain a Dirac or a spike signal is concentrated to one particular time instant where as when the same signal is represented in frequency domain, it spreads out in the frequency domain. Similar concept can be used in CS based sparse representation of the signal. Suppose given a pair of \((\phi, \psi)\) orthobases of dimension \(R^N\) each. For explanation of incoherence let ‘\(\phi\)’ basis is used for sensing the signal ‘\(x\)’ and the other basis is used to represent ‘\(x\)’ in sparse domain. The relation between these two bases decides the reconstruction quality in CS recovery and hence it is essential to determine the degree of correlation between the pair of orthobases. The mutual coherence of a pair of orthobases \((\phi, \psi)\) is the measure of highest similarity factor between any two columns of ‘\(\psi\)’ and ‘\(\phi\)’. If ‘\(\phi\)’ and ‘\(\psi\)’ contain correlated elements, the coherence is large, otherwise the mutual coherence value is small.

The mutual coherence between the sensing basis ‘\(\phi\)’ and the representation basis ‘\(\psi\)’ is given by the equation (2.2).

\[
\mu(\phi, \psi) = \sqrt{n} \cdot \max (\phi_k, \psi_j) \quad \text{Where,} \quad 1 \leq k, j \leq n; \quad (2.2)
\]

In CS based reconstruction, small mutual coherence value is desirable as it provides better recovery of the signal.

2.3.1 CS encoding

The CS algorithm can be better explained with the help of mathematical expressions given in this section. Conversion of the acquired signals or images into a sparse signal is done by the use of sparsifying kernel in CS. Sparse signals are the signals which are very few in number and has complete structure and meaningful information which are essential to reconstruct the original signal or image. The generation of the sparse signal \(s\) in matrix form is given in equation (2.3).

\[
s = \psi x \quad (2.3)
\]
In equation 2.3, the input signal is denoted by ‘x’ and ‘s’ is the sparse representation of the input data and ‘Ψ’ is the transform matrix or kernel. After sparse generation, measurement matrix ‘A’ is used for compression of the sparse signal or image. The output ‘y’ which is the compressed version of input ‘x’ is given in equation (2.4),

\[ y = As \] (2.4)

Where ‘A’ is called the measurement or sensing matrix, ‘s’ is the sparse representation and ‘y’ is the compressed output.

### 2.3.2 CS reconstruction

CS involves a variety of methods for representation of a signal on the basis of a limited number of samples. CS based recovery of signals involve reconstruction of signal from these reduced number of samples. The major issue in CS recovery is how to effectively reconstruct the original signal from the compressed data without the loss of generality. Currently, there are varieties of reconstruction algorithms available in literatures which are defined either based on the framework of convex optimization, or greedy approaches. These are namely:

a) Convex Relaxation  
b) Non Convex Minimization Algorithms  
c) Greedy Iterative Algorithm  
d) Combinatorial / Sublinear Algorithms  
e) Iterative Thresholding Algorithms  
f) Bregman Iterative Algorithms

The complexity of the CS recovery algorithm and also the computational speed of the algorithms are very important as these factors influence the recovery of the signal in practical situations. In general, the convex relaxation method is used for CS reconstruction for its simplicity and computational speed. Convex relaxation based recovery method has high computational speed compared to other reconstruction algorithms mentioned.
above. The convex relaxation reconstruction method is used in this research work which is explained below.

2.3.3 Convex relaxation method for reconstruction

Convex relaxation methods of CS based reconstruction was developed in the year 1980s. Donoho and his collaborators suggested the idea of convex relaxation in late eighties [42, 43]. It is purely depends on the concept called linear programming [44]. In CS, the reconstruction is done from far fewer samples of original signal and hence the exact reconstruction is a challenging task as the computational complexity is much higher. Some of the reconstruction algorithms namely Basis Pursuit [45], Basis Pursuit De-Noising (BPDN) [45], Least Absolute Shrinkage and Selection Operator (LASSO) [46], Least Angle Regression (LARS) [47] etc. are example of such convex relaxation algorithms.

Basis Pursuit (BP) algorithm involves decomposition of a signal into optimal dictionary elements. The optimal elements have smallest $l_1$ norm and hence based on the $l_1$ norm coefficient values one can find the optimal decomposed components of the signal. Basis pursuit method finds the best representation of the signal by optimizing the $l_1$ norm. Ideally the components of signal nearer to zero value are preferable. BP method not only recovers the sparse data but also can recover compressible data with the help of minor modifications in the algorithm. The application of basis pursuit method ranges from small to larger scale.

Basis pursuit method is suitable in large scale optimization problems due to recent development in linear and quadratic programming. Linear and quadratic programming uses the interior point methods. In [46], wide variety of models uses $l_1$ penalties for optimization is shown. Application of these models are allowed by the advanced computational algorithms to huge number of datasets which enables the statistical and computation gain by exploiting the sparsity behaviour of data. In many fields of engineering, statistics, mathematics and computer science, Lasso ($l_1$) penalties based optimization techniques are used for solving the problems. Matrix versions of
signal recovery method called as $||M||_1$ Nuclear Norm optimization [48] is used in latest work and shows the improved capabilities of the algorithm compared to conventional methods.

### 2.4 THRESHOLD BASED CS ALGORITHM

Threshold based CS is proposed in order to sparsify the data to improve the reconstruction quality of the 1 D signal when the correlation of the samples in a signal is very low. The threshold values (upper threshold and lower threshold) are selected for further improvement of sparsity of the signal so that CS can be efficiently used for dimensionality reduction in the signal of interest. The threshold is chosen based on three sigma rule of normal distribution. Three sigma rule expresses the values in a dataset, to lie within their standard deviation of mean and empirically three sigma rule is useful to treat 99.7% probability as near certainty.

Reconstruction of the signal is done using basis pursuit method. The steps involved in threshold based CS is shown in Figure 2.1. In order to perform a comparative study of standard CS and threshold based CS, input signal sets with various mean autocorrelation coefficients are considered and both the algorithms for compression i.e., standard CS and threshold based CS are tested.

Figure 2.1 shows the basic steps involved in threshold based CS. A Gaussian distribution based upper threshold and a lower threshold is calculated in order to enforce the sparsity in transformed data. The sparsity of the transformed data is increased with the help of thresholding technique which in turn improves the compression ratio. The sparsity enhancement is essential in compressed sensing for achievement of high quality reconstruction and hence in order to increase the sparsity, the thresholding is performed.
Threshold based CS and standard CS algorithms are simulated on the datasets having different correlation coefficients. The correlation coefficient parameter is often used in data analysis which refers how close data samples are or having linear relationship with each other.

Simulation result shows that for highly correlated data, threshold based compressive sensing algorithm does not provide significant improvement but standard CS works well whereas for low correlated data, threshold based CS is much suitable compression technique to be used in practice for desired signal to noise ratio and less mean squared error.
Algorithm: Threshold Based Compressed Sensing (CS) algorithm

Step 1: Start with input data and compute the mean.

Step 2: Perform Subtraction of the mean value from each data point and find square of the results.

Step 3: Calculate mean of results obtained in Step 2.

Step 4: Find the standard deviation by taking square root of result obtained in Step 3.

Step 5: Multiply the standard deviation by 1 or 2 or 3 etc. for 1 Sigma, 2 Sigma and 3 Sigma rule respectively

Step 6: perform the addition of result obtained in Step 5 with the statistical average value of the original data set and calculate the upper threshold point.

Step 7: Result of Step 5 to be subtracted from the statistical average value of the original data set to find the lower threshold point.

Figure 2.2: Threshold based Compressed Sensing (CS) algorithm

In Figure 2.2, the upper and lower threshold values are calculated in step 6 and step 7 respectively. After the calculation of threshold limits, the data falling beyond the upper and lower threshold limits are made as zeros which in turn sparsify the data making it much suitable for CS algorithm. The zero enforcement in the dataset based on threshold values increases the sparsity of dataset which is the prime necessity for CS based compression.
2.5 RADIO ENERGY MODEL OF WSN

In order to compute energy required for transmitting compressed data, radio model from [49] is used. It is assumed that ‘k’ bits of data are transmitted over a distance of ‘d’ meters. The transmitting energy $E_{Tx}(d)$ and the receiving energy $E_{Rx}(k)$ are shown in the first order radio model in Figure 2.3. The formulae for calculation of transmission energy and receiving energy for sending and receiving ‘k’ bits data are given in equations (2.5) and (2.6).

![Figure 2.3: First order radio model for WSN](image)

\[
E_{TX}(k,d) = E_{TX\text{-}elec}(k) + E_{TX\text{-}amp}(k, d)
\]

\[
E_{TX}(k,d) = E_{elec} \cdot k + \epsilon_{amp} \cdot k \cdot d^2
\]  

(2.5)

\[
E_{Rx}(k) = E_{Rx\text{-}elec}(k)
\]

\[
E_{Rx}(k) = E_{elec} \cdot k
\]  

(2.6)
Where, ‘$E_{\text{elec}}$’ is the energy dissipation by the radio to run the transmitter or receiver circuitry. ‘$E_{\text{amp}}$’ is the energy consumption of the amplifier for maintenance of the radio of reliable transmission.

The transmitter in the radio model of Figure 2.3 shows the energy expenditure for operating its radio electronics and the power amplifier. Even the receiver of radio model needs energy for operating its radio electronics [49]. To compute the energy usage based on the distance between transmitter and receiver, the radio model is used.

2.6 RESULTS AND DISCUSSIONS

Temperature sensor data taken from online dataset [55] of size 200 samples, 400 samples, 1000 and 2000 samples are used as input signals and compressed using proposed CS algorithm. Initially a dataset consists samples of 200 numbers are used whose mean correlation coefficient is 0.65. The parameters namely Signal to Noise Ratio (SNR) and Mean Squared Error (MSE) are calculated for analyzing the performances of the standard CS algorithm and proposed threshold based CS algorithm. For both the algorithms the sparsifying basis used as discrete cosine transform (DCT) as DCT kernel is much suitable sparse basis for one dimensional signal. CS reconstruction of the signal from its compressed version is performed using basis pursuit method which is a convex relaxation method discussed in section 2.3.3. Basis pursuit method is a mathematical optimization problem of the form given in equation (2.7).

$$\min ||s||_1 \quad \text{subject to} \quad y = As$$

In the equation 2.7, dimension of the vector's' is $(N \times 1)$. The dimension of compressed vector 'y' is $(M \times 1)$. In equation 2.7, the matrix 'A' is $(M \times N)$ dimension transform matrix also termed as measurement matrix where $M < N$.

Basis pursuit method uses $l_1$- norm optimization for recovery of the signal. $l_1$- norm optimization is a convex type problem and can be considered
as a linear programming problem. In order to reconstruct the signals, $l_1$-norm optimization is used in this chapter.

For the purpose of CS recovery, a small mutual coherence value of matrix (A) is desirable. For a given matrix ‘A’, the mutual coherence can be defined by the largest absolute normalized inner product between different columns from ‘A’. Here ‘A’ is measurement matrix given by equation (2.4). The size of matrix ‘A’ is dependent on dimension of original signal and the number of compressed samples. Suppose $A = [a_1 \ a_2 \ \ldots \ \ldots \ \ a_n]$. Then the mutual coherence of matrix ‘A’ is given by equation (2.8):

$$\mu(A) = \max_{k,j,k\neq j} \frac{|(a_k)(a_j)|}{\|a_k\|_2 \cdot \|a_j\|_2}$$ (2.8)

In equation (2.8), $|[(a_k)(a_j)]|$ is the maximum absolute normalized inner product between columns $a_k$ and $a_j$ of matrix ‘A’. $(\|a_k\|_2)$ and $(\|a_j\|_2)$ are the $l_2$ norms values of the columns $a_k$ and $a_j$ of matrix ‘A’. Mutual coherence values are computed for varied compressed sized data. The input signal of size 200 samples of a temperature dataset [55] of mean correlation coefficient 0.65 is chosen and with respect to different compression ratio, the SNR and MSE value is tabulated and is given in Table 2.1. The mutual coherence value of measurement matrix ‘A’ is also tabulated in Table 2.1. The measurement matrix ‘A’ for the dataset of 200 samples is generated using the equations (2.3-2.4). The mutual coherence values in Table 2.1 are calculated using equation (2.8).
Table 2.1: Impact of mutual coherence on signal recovery

<table>
<thead>
<tr>
<th>Number of Input data (Temperature) samples</th>
<th>Number of Compressed data samples</th>
<th>SNR (dB)</th>
<th>MSE</th>
<th>Mutual Coherence</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>50</td>
<td>34.98</td>
<td>0.03</td>
<td>$1.21 \times 10^{-15}$</td>
</tr>
<tr>
<td>200</td>
<td>67</td>
<td>36.5154</td>
<td>0.0502</td>
<td>$3.52 \times 10^{-15}$</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>51.2187</td>
<td>0.0017</td>
<td>$9.28 \times 10^{-16}$</td>
</tr>
</tbody>
</table>

Table 2.1 shows mutual coherence value is inversely proportional to the SNR value. It is clear from Table 2.1 that when the mutual coherence value is low ($9.28 \times 10^{-16}$) then the SNR value is high (51.2187 dB). Hence for better recovery of CS based compression, low mutual coherence value is desirable. The SNR is found using the equations (2.9) and to find MSE equation (3.0) is used.

\[
\text{SNR (dB)} = 10 \log \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right) \tag{2.9}
\]

\[
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} [ \overline{X_i} - \overline{X} ]^2 \tag{3.0}
\]

In above equation (2.9), the signal power is denoted by $P_{\text{signal}}$ and the noise power is $P_{\text{noise}}$. The power of input data is $P_{\text{signal}}$ whereas difference power of input signal and reconstructed signal is termed as noise power $P_{\text{noise}}$. In equation (3.0), the ‘$\overline{X_i}$’ is input vector and ‘$\overline{X}$’ is reconstructed signal vector of length ‘$n$’.
2.6.1 Performance analysis of standard CS and proposed threshold based CS for highly correlated data

The temperature data samples with size 200 numbers of mean correlation coefficient 0.65 is considered as input signal for performance analysis of standard CS and threshold based CS. Table 2.2 shows the SNR and MSE for both proposed and standard CS with different compression ratio.

Table 2.2: SNR and MSE of standard CS and threshold based CS for highly correlated data

<table>
<thead>
<tr>
<th>Number of Input data (Temperature) samples</th>
<th>Number of Compressed data samples</th>
<th>Standard CS</th>
<th>Threshold based CS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>SNR (dB)</td>
</tr>
<tr>
<td>200</td>
<td>50</td>
<td>0.0179</td>
<td>40.988</td>
</tr>
<tr>
<td>200</td>
<td>67</td>
<td>0.0049</td>
<td>46.643</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>0.0013</td>
<td>52.261</td>
</tr>
</tbody>
</table>

Result of Table 2.2 shows that for varied compression ratio, standard CS has better performance than threshold based CS. The original and reconstructed temperature samples are plotted in Figure 2.4 for standard CS and for threshold based CS it is plotted in Figure 2.5.
In Figure 2.4, the reconstructed signal is generated from 50% samples of compressed data using standard CS. From the plot in Figure 2.4 it is clear that the input and reconstructed signal both are almost similar. The mathematical proof of this similarity index is measured by calculating SNR and MSE values. The MSE is a function corresponds to expected value of squared error loss and it is always strictly positive. The SNR value calculated here is also a function of MSE. High SNR and low MSE value ensures that the reconstructed signal is similar to original input signal.
Figure 2.5: Original and reconstructed signals with threshold based CS

In Figure 2.5, the reconstructed signal is generated from 50% samples of compressed data with threshold based CS. From the Figure 2.5, it is clear that both the input and reconstructed graph are almost similar which ensures the better reconstruction.

In order to compare the SNR values in standard CS and proposed threshold based CS algorithms for temperature data of 200 samples of high correlation coefficient is shown in Figure 2.6. The MSE values are compared between standard CS & threshold based CS for same data and is shown in Figure 2.7. From the Figure 2.7, it is observed that the MSE value is low in case of standard CS.
Figure 2.6: Comparison of SNR (in dB) for standard CS & threshold based CS for data with high correlation coefficient

Figure 2.6 shows the SNR values with respect to number of compressed data samples for standard CS and threshold based CS. It is clear from the graph that in case of standard CS, the SNR value is very high which is 52.2617 dB with 50% compression ratio. This result is obtained for the data with high correlation coefficient. Similar kind of analysis is also performed for the data with low correlation coefficient but in this case standard CS does not perform well. When the data correlation is low, then standard CS can not increase the sparsification of data and hence cannot provide better result.
Figure 2.7: Comparison of MSE for standard CS & threshold based CS for data with high correlation coefficient

Figure 2.7 shows that in case of standard CS, the MSE value is very low and it is 0.0017 with 50% compression ratio.

Further the performance analysis of standard CS and threshold based CS is done for another dataset of 400 samples of temperature data. Figure 2.8 shows the original and reconstructed data with 50% compression ratio. The correlation coefficient of this dataset is equal to 0.93. The calculated mean squared value is 0.0413 and the signal to noise ratio is 37.3666 dB when compression ratio is 50%.
Performance analysis of threshold based CS for the dataset of 400 samples temperature data with 50% compression ratio is shown below in Figure 2.9. The upper threshold and lower threshold values for the dataset of 400 samples of temperature data are chosen by considering three sigma rule as mentioned in algorithm of Figure 2.2. Application of three sigma rule for choosing the upper and lower threshold value is based on the statistical nature of the data. In this case when the sigma rule is applied, the data samples lying outside the three sigma region of distributed data are made zero which in turn increases the sparseness of the data. The correlation coefficient of this dataset is equal to 0.93 (very high). The calculated mean squared value is 0.1689 and the signal to noise ratio is 31.2443 dB when compression ratio is 50%.

Figure 2.8: Original and reconstructed data with standard CS for highly correlated data with 50% compression ratio
3.2

Figure 2.9: Original and reconstructed signal with threshold based CS for highly Correlated Data

Results of Figure 2.8. and Figure 2.9. show that for highly correlated data threshold based compressive sensing algorithm is not suitable but standard CS works well. In case of highly correlated data, the dataset do not possess the Gaussian distribution and hence when the sigma rule is applied for selection of threshold values, it does not have any impact on zero enforcement in data. Thus the sparsity of highly correlated data does not get increased by threshold based CS and hence threshold based CS does not provide better SNR for highly correlated data.

2.6.2 Performance analysis of standard CS and threshold based CS for data with low correlation coefficient

A set of pressure data with mean correlation coefficient 0.31 (very low) is considered and the comparison of proposed algorithm with standard CS is done and shown in Table 2.3.
Table 2.3  SNR, MSE comparison of standard CS and threshold based CS for data of mean correlation coefficient 0.31.

<table>
<thead>
<tr>
<th>Number of Pressure data samples</th>
<th>Number of Compressed data samples</th>
<th>Standard CS</th>
<th>Threshold based CS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean Squared Error (MSE)</td>
<td>SNR (dB)</td>
</tr>
<tr>
<td>200</td>
<td>50</td>
<td>1.0496</td>
<td>23.3115</td>
</tr>
<tr>
<td>200</td>
<td>67</td>
<td>0.7960</td>
<td>24.5130</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>0.6126</td>
<td>25.6501</td>
</tr>
</tbody>
</table>

Table 2.3 shows the performance comparison of proposed algorithm with standard CS. The performance comparison in terms of SNR between standard CS & threshold based CS is done for pressure data of low correlation coefficient and is shown in Figure 2.10. MSE comparison is shown in Figure 2.11. Result shows that threshold based CS performs much better than standard CS when the data is low correlated in nature.
Figure 2.10: Comparison of SNR (in dB) of standard CS & threshold based CS for data with low correlation coefficient.

Figure 2.11: Comparison of MSE value of standard CS & threshold based CS for data with low correlation coefficient.
Results of Figure 2.10 and Figure 2.11 show that threshold based CS produces improved results when the data is low correlated in nature. In order to find the suitability of proposed threshold based CS, another dataset of very low correlation coefficient is chosen and the performance of the suggested algorithm is discussed in following section.

2.6.3 Performance analysis of standard CS and threshold based CS for audio data

An audio data set of very low correlation coefficient is considered for analyzing the performance of threshold based CS on low correlated data compression. Figure 2.12. below shows the comparison of standard CS and threshold based CS for an audio dataset of length 2000 samples.

![audio Signal original and reconstructed](image)

Figure 2.12: Original and reconstructed audio signal with standard CS for very low correlated audio data

The reconstruction is performed from a compressed dataset of size 1000 samples. This input dataset has mean correlation coefficient of very
low value equals to 0.07. The audio signals possess very low correlation and in literature there are not much work carried out for low correlated signal compression and hence this proposed method is mainly focusses on the efficient compression algorithm for low correlated data and signal. The measured mean squared error and signal to noise ratio are \( \text{MSE} = 1.1513 \times 10^{-5} \) and \( \text{SNR} = 20.4872 \text{ dB} \) respectively when the standard CS is applied on the low correlated audio data.

![Graph showing original and reconstructed audio signal](image.png)

**Figure 2.13:** Original and reconstructed audio signal for very low correlated audio data with threshold based CS

The above Figure 2.13. shows the original and reconstructed audio signal with threshold based CS. In this case the measured MSE value is \( 3.9669 \times 10^{-6} \) and SNR value is 25.1144 dB. In wireless transmission of audio signal, for better quality of audio the minimum SNR required is 25 dB to 40 dB [67] and less than 25 dB of SNR leads to poor quality of signal. Hence from the result of Figure 2.13. and Figure 2.14., it is clear that threshold based CS can provide the better quality of audio signal than standard CS based compression.
2.6.4 Comparison of standard CS and threshold based CS for different size datasets with 50% compression ratio

Comparison of standard CS and threshold based CS for various dataset with 50% compression ratio is shown in Table 2.4. The analysis helps to identify easily how the proposed scheme improves the reconstruction quality by means of improved signal to noise ratio and minimum mean squared error for low correlated signals irrespective of data size.

Table 2.4: Comparison of standard CS and threshold based CS for different sizes dataset with 50% compression ratio

<table>
<thead>
<tr>
<th>No. of input data samples</th>
<th>No. of Compressed data samples</th>
<th>Mean Correlation Coefficient</th>
<th>Standard CS</th>
<th>Threshold based CS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean Squared Error (MSE)</td>
<td>SNR (dB)</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>0.65</td>
<td>0.0013</td>
<td>52.261</td>
</tr>
<tr>
<td>400</td>
<td>200</td>
<td>0.93</td>
<td>0.0413</td>
<td>37.366</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>0.31</td>
<td>0.6126</td>
<td>25.650</td>
</tr>
<tr>
<td>2000</td>
<td>1000</td>
<td>0.18</td>
<td>2.9998 x 10^{-5}</td>
<td>24.190</td>
</tr>
<tr>
<td>2000</td>
<td>1000</td>
<td>0.07</td>
<td>1.1513 x 10^{-5}</td>
<td>20.487</td>
</tr>
</tbody>
</table>
From the results of above Table 2.4, it is clear that threshold based CS gives better result in terms of higher SNR for low correlated input signal. Similar analysis on highly correlated data does not show significant improvement.

2.7 ENERGY COMPUTATION USING RADIO MODEL

In order to analyze the energy efficiency of proposed threshold based CS, percentage energy saving is computed for various sampling ratio for a temperature sensor dataset [55] of size 1000 samples and shown in Table 2.5. Equation 2.5 - 2.6 are used to compute the transmission energy, energy saving, percentage energy saving etc.

Table 2.5: Radio energy computation

<table>
<thead>
<tr>
<th>No. of input samples (N)</th>
<th>No. of compressed samples (M)</th>
<th>Sampling ratio (SR= N/M)</th>
<th>Transmission energy of input samples $E_{TX}(N)$</th>
<th>Transmission energy of compressed samples $E_{TX}(M)$</th>
<th>Energy savings $E_{saving} = ((N-M)/N)^*E_{TX}$</th>
<th>% Energy Saving</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>750</td>
<td>1.33</td>
<td>1050μJ</td>
<td>787.5μJ</td>
<td>262.5μJ</td>
<td>25%</td>
</tr>
<tr>
<td>1000</td>
<td>500</td>
<td>2</td>
<td>1050μJ</td>
<td>525 μJ</td>
<td>525μJ</td>
<td>50%</td>
</tr>
<tr>
<td>1000</td>
<td>250</td>
<td>4</td>
<td>1050μJ</td>
<td>262.5μJ</td>
<td>787.5 μJ</td>
<td>75%</td>
</tr>
<tr>
<td>1000</td>
<td>125</td>
<td>8</td>
<td>1050μJ</td>
<td>131.25 μJ</td>
<td>918.8 μJ</td>
<td>87.5%</td>
</tr>
</tbody>
</table>
Few assumptions are made for calculation of transmission energy in above Table

Assumptions:
Type of node : MICA2DOT (MPR 500) sensor node

\( E_{\text{elec}} : 50 \text{ nJ} / \text{bit} \)

\( \varepsilon_{\text{amp}} : 100 \text{ pJ} / \text{bit} / \text{m}^2 = 0.1 \text{nJ} / \text{bit} / \text{m}^2 \)

\( d : 100 \text{ m} \)

Results show that energy efficiency of data compression in WSN is improved by adopting threshold based CS as the proposed scheme provides more sparsity and hence degree of compression ratio can be chosen higher which provides better energy saving in WSN. Calculation of transmission energy, energy savings, percentage savings of energy etc. are done using the formulae given in equations (2.3) and (2.4).

**2.8 CONCLUSION**

The proposed threshold based CS was applied on highly correlated data, low correlated data and also on uncorrelated audio data. Results show that, threshold based CS performs better for low correlated signal compression. Selection of suitable threshold limits increases the level of sparsity which is an essential requirement in CS based compression. Hence threshold based CS can be proposed as a compression algorithm for audio signal (low mean correlation co-efficient) compression.

To improvise the performance of compression algorithm of highly correlated data, modification in the compression technique is suggested by performing differential encoding of the input data followed by an efficient CODEC which is discussed in the next chapter.