# THE AUDIO STEGANALYSIS SYSTEM BASED ON MUTUAL INFORMATION APPROACH

**By**

**SU SU HALING**

**D. C. M**

**A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Applied Science (M. A. Sc.)**

**University of Computer Studies, Yangon**

**June, 2022**

# ACKNOWLEDGEMENTS

First and Foremost, I would like to express my sincere gratitude to **Dr. Mie Mie Khin**, Rector of the University of Computer Studies, Yangon and **Dr. Soe Lin Aung**, Pro-rector of University of Computer Studies, Magway, for their kind permission to conduct this thesis.

I would also like to offer my deep and sincere gratitude to my supervisor, **Dr. Yawai Tint**, Associate Professor, Faculty of Computer Science and Technology, University of Computer Studies, Magway, for her effort, time in reading and patience to help me in accomplishing this paper. It was a great privilege and honor to work and study under her guidance.

My sincere thanks and regards go to **Dr. Htar Htar Lwin**, Pro-rector, Head of Faculty of Computer Systems and Technologies, University of Computer Studies, Yangon, for her kind management throughout the completion of this thesis.

I would like to express my deeply thanks to **Dr. Amy Tun**, Professor, Course coordinator of Master's (CT), University of Computer Studies, Yangon, for her painstaking suggestion and encouragement throughout the development of the thesis.

I would like to thank all my teacher for their motivated, encouragement and recommending the thesis.

Furthermore, I am extremely grateful to my companions who have given me their precious ideas and invaluable knowledge throughout this thesis. Finally, I want to extend my heartfelt thanks to my family for their encouragements and support to accomplish this work.

# ABSTRACT

Steganography is a technique whereby we put the existence of a message to question by simply covering it up within another file image or video. This system used the input audio signals as hidden message with audio signal which is embedded based on the stenographic tool and pure audio signal. The proposed system tested the steganalysis technique on audio signals embedded with Invisible Secret tools which are available from the Internet. The aim of the system is to develop the steganography detection system using the concept of data analysis approach. To evaluate the performance of proposed system, three types of audio signal are analyzed with different number of bit stream. Overall detection accuracy for various types of music is around 90% and experiments showed that the proposed passive steganalysis is to detect hidden message in audio files.

# TABLE OF CONTENTS

**Page**

# LIST OF FIGURES

**Page**

# LIST OF TABLES

**Page**

CHAPTER 1

# INTRODUCTION

Steganography is the method of hiding secret data within a conventional, non-secret, file or message in order to avoid detection; the secret data is then taken out at its target. The goal of steganography is to support convert communication by hiding data in digital covers such as images, audios and videos, etc. Numerous steganography approaches and software have been broadly applied. Correspondingly, steganalysis techniques are technologically advanced to detect the existence of hidden information. Steganalysis is the scientific technology to decide if a medium carrier some hidden messages or not and if possible, to determine what the hidden messages are. Figure 1.1 and Figure 1.2 show representations of Steganography and Steganalysis, respectively.

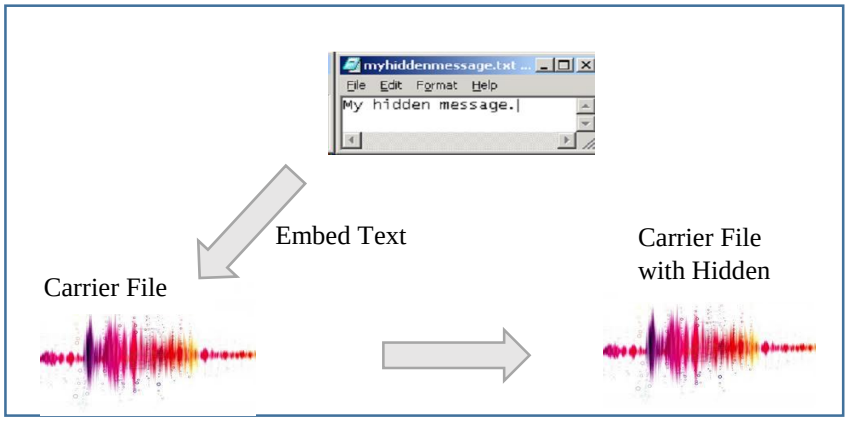



**Figure 1.1. Representation of Steganography**

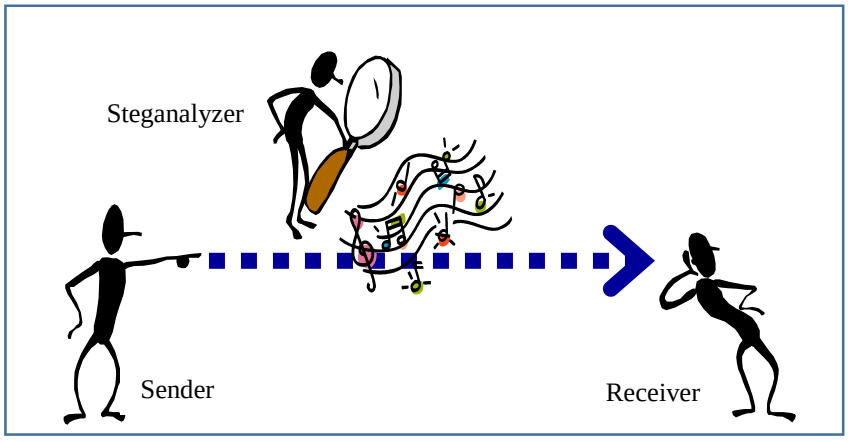



**Figure 1.2. Representation of Steganalysis**

Technique-specific steganalysis and universal steganalysis are the most popular research approach to analyze the steganoraphic algorithms. The former group of techniques achieves very accurately when used against the steganographic technique it is targeted for. The latter group of technique, on the other hand, is effective over a wide-ranging variety of techniques, while carrying out less accurately overall. However, ever since universal steganalysis is well suitable to the practical setting, it attracted more awareness and many effective steganalyzers are proposed.

Audio is a crucial communication way for society, and so is a convenient medium secure communication. Audio steganography is a methodology of hiding information within an audio signal. As data is embedded in the signal, it comes to be transformed and this change should be made indistinguishable to the human ear [1]. An effective audio steganography may yield an output which is similar with original audio and not capable to sense the alteration by human ear [2]. This system emphasizes on MP3 audio files. With the aim of discriminating stego audios from pure ordinary ones, that set in random data into a (possibly) stego file.

## 1.1. Audio Steganography

Audio steganography is using the advantage of the psychoacoustical masking phenomenon in human auditory system. Hidden message are embedded within an inaudible tones in original audio signal. According to the process of audio steganography, secret message are hidden into digitalized signal. The value of the audio sequence are slightly altered for being the process audio steganography. Figure 1.3 describes the concept of audio steganography system.

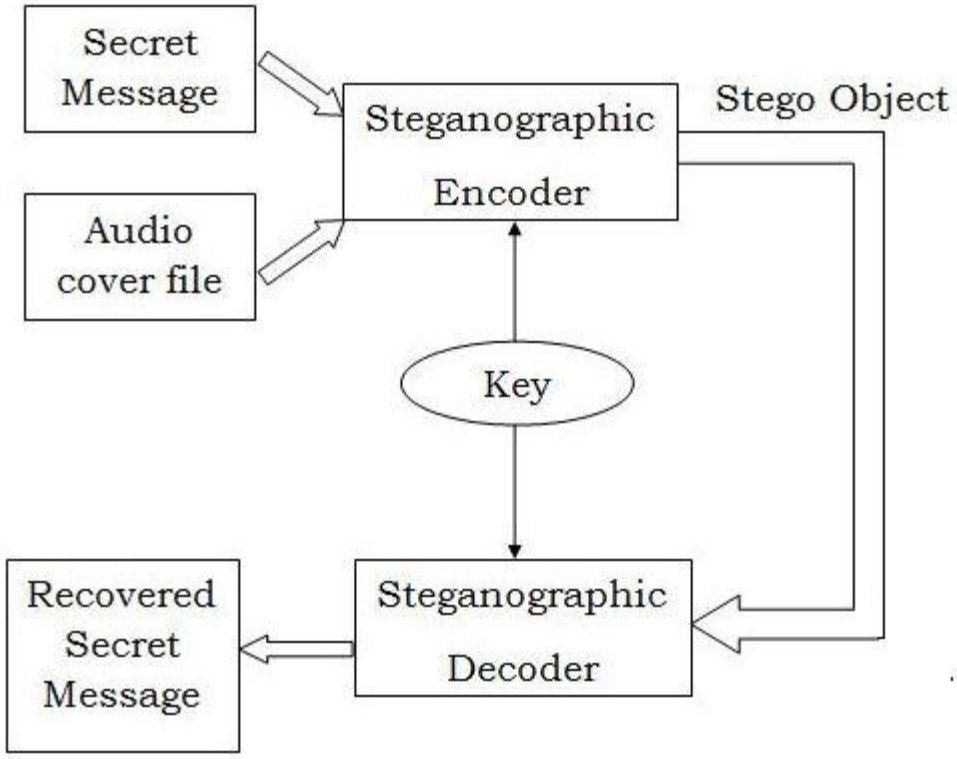



**Figure 1.3. Process Diagram of Audio Steganography**

## 1.2. Applications Area of Steganalysis

Convert channels, embedded data and digital water marking are the most popular research areas in a digital signal processing environment. For a converting channels, some information are needed to hide within the TCP/IP header for user identification, verification and authorization.

Nowadays, hidden message are embedded within a container for secure communication. Audio data hiding are applied to embed a secret chemical formula in this day. For copy right protection in digital media, mostly audio and video steganography techniques are used.

A universal steganalytic method can be used to detect several kinds of steganography. Steganalysis is used to apply in information countermeasures, and prevent the illegal used of access information. This proposed system applied in this application

area to get more secure communication channel. In addition, this system is used to prevent the illicit use of steganography.

## 1.3. Objectives

The objectives of the proposed system are:

- To develop the steganography detection system based on the concept of audio features and physical model.
- To identify the statistical difference between the setgo and without stego.
- To analyze the audio feature based on Mutual Information (MI).
- To apply the statistical concept in audio steganalysis system.

## 1.4. Outlines of the Proposed System

The remainder of this system is organized into the following chapters:

Chapter 2 describes related works of the proposed system and explains proposed audio steganography detection system with process diagrams in both training and testing phases.

Chapter 3 discusses related methodologies of the proposed system and describes feature extraction process: which features have been generated from the audio and how they are generated is given in details. This chapter also explains steganography detection approach which is used in this proposed system.

Chapter 4 shows implementation results, two data sets used for training and test phases.

Chapter 5 describes the conclusion of the proposed system.

CHAPTER 2

# RELATED WORKS AND PROPOSED SYSTEM

Steganalysis is the mechanism of identifying the existence of hidden information in the stego media. Natarajan Meghanathan1 and Lopamudra Nayak [3] provided a critical review of the steganalysis algorithms available to study the physical characteristics of audio and image stego media compared to the corresponding original media and the procedure of embedding the information and its detection. This review offered a clear representation of the contemporary developments in steganography.

## 2.1. Related Works

Hamzeh Ghasemzadeh, Mehdi Tajik Khas andMeisam Khalil Arjmandi proposed a reliable audio steganalysis system is applied on the reversed Mel frequency cepstral coefficients (R-MFCC) which purposes to make available a model with maximum deviation from Human Auditory System (HAS) model. Genetic algorithm was used to improve dimension of the R-MFCC-based features [20].

Qingzhong Liu, Andrew H. Sung and Mengyu Qiao [1] presented an innovative stream data mining in detection of audio steganography, based on second order derivative of audio streams. The authors extracted Mel-cepstrum coefficients and Markov transition features on the second order derivative; a support vector machine was used to the features for detection of the presence of concealed message in digital audios.

Yuzhen Lin et al. proposed steganography detection system by modifying CNN appraoch. Precisely, a special convolutional layer was first designed, which could capture the minor steganographic noise. Then, a truncated linear unit was adjusted to activate the output of shallow convolutional layer. Additionally, the average pooling was applied to reduce the over-fitting risk [12].

Fatiha Djebbar and Beghdad ayad [4] presented a combined maximum entropy energy approach for audio steganalysis. First, the audio signal was separated into four energy-based regions: noise, low, medium and high; then entropy is calculated from each

region. Finally, a support vector machine was applied to the collected features for finding out the hidden data in audio signals. Active speech level algorithm was applied for capturing energy fluctuation in audio streams.

Yali Liu et al. [5] proposed a novel quality metric based on Hausdroff distance. In detail, the Hausdroff distance was used for measuring the distance between the wavelet coefficients of the stegoaudio object and the de-noised stego-audio object. Hamzeh Ghasemzadeh and Meisam Khalil Arjmandi [6] proposed features based audio steganalysis and try to get better performance.

Enas Wahab Abood et al. [7] proposed a combination approach of steganography and cryptography for producing a powerful hybrid securing stego-system. Initially, a text message was encrypted with a new method using a bits cycling operation to offer a cipher text. In the succeeding stage, an enhanced LSB method was used to hide the text bits randomly in an audio file of a wav format. Peak signal-to-noise ratio (PSNR), mean squared error (MSE) and structural similarity (SSIM) were employed to calculate the performance of the proposed system.

## 2.2. Steganography Detection

Detecting a steganographic file mean to analyze the cover media whether it has hidden message in it or not. Steganalyzer apply the various tools that are available on the Internet which use for creating steganographic files and moreover discriminate the characteristic of original and stego signal.

If we can get the original file source, we will compare the nature of questionable file with that. In the real world case, that is not possible to get the original format of cover file and steganographic tools.

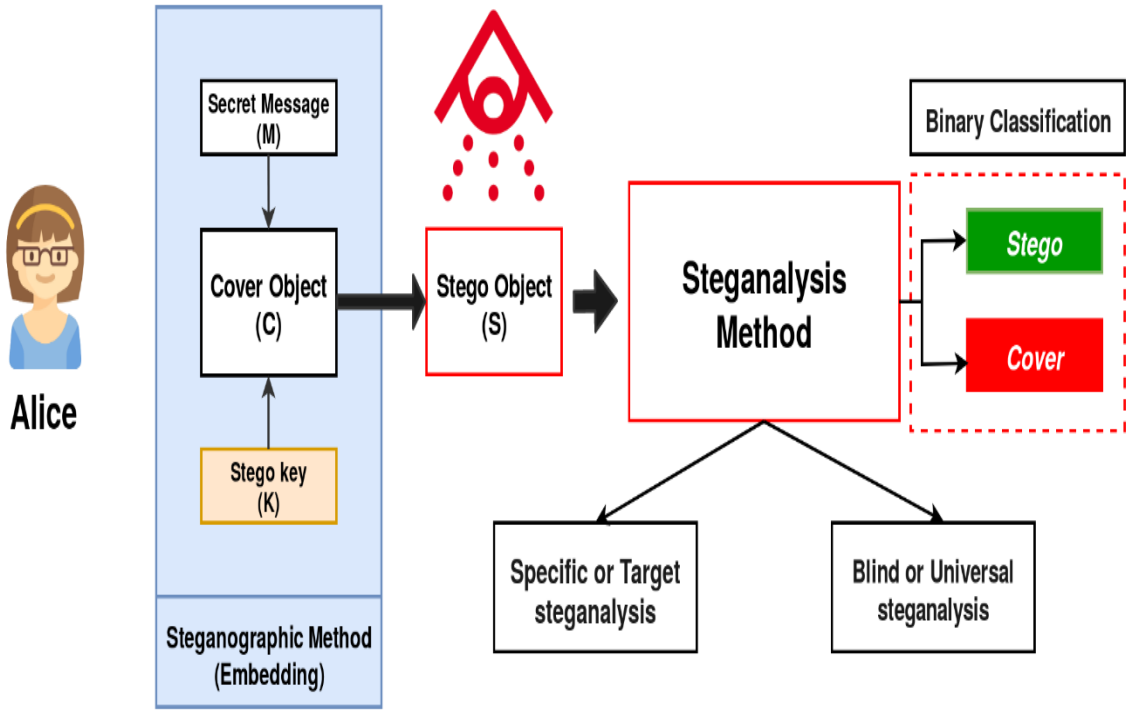



**Figure 2.1. Fundamentals Concepts of Steganalysis**

Figure 2.1 describes the sample process of steganography detection system. In the proposed system follow the procedure of this process and analyze the nature of audio signal based on mutual information approach. The final results take place the binary classification 0 for steganography and 1 for original file. Within the process of audio steganography detection system, specific steganalysis and universal steganalysis are applied in the proposed system.

## 2.3. Audio Steganographic Algorithms

The following section describes the methods of audio steganography which are explained with two parts technical and linguistic steganography approach. The most common methods are as follow:

Least Significant Bit Coding: LSB is the earliest method for audio steganography system which are finding the binary equivalent of hidden message are replaced in the sample of digitized audio file. Figure 2.2 shows the structure of steganographic methods.

Parity Coding: It is used to analyze the signal and separate the regions of audio samples and binary stream from hidden message in a sample regions of parity bit.

Phase Coding: It is used the substitution method for data embedding process. Initial audio segment is replaced by the reference phase of the hidden message.

Spread Spectrum: spread the hidden message over the cover media spectrum, using a code that is independent of the actual signal.

Echo Data Hiding: It evaluate the one echo from original signal and only one bit of information are encoded at a time.

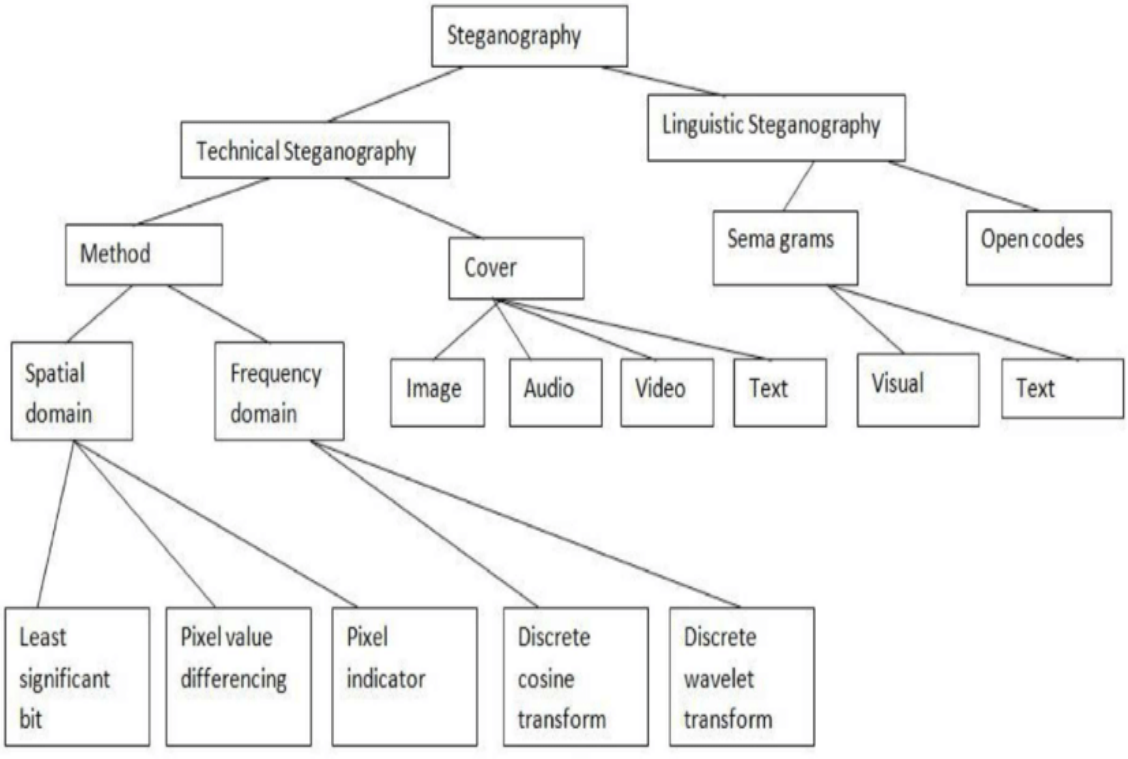



**Figure 2.2. Steganographic Methods**

## 2.4. Audio Steganographic Tools

There are eleven different products of steganographic tools which are shown in Table 2.1. Some are open source products that can be applied for free and some are commercial used. For audio data hiding, most of the product hide data in WAV file format. Hide4PGP and Steganos can be used for other format such as VOC and StegMark and StegHide embed secret in MIDI and AU format, respectively.

**Table 2.1. List of Audio Steganographic Tools**

| Audio Steganographic Tools | MP3 | WAV | Others | Production | License |
|---|---|---|---|---|---|
| Info Stego | Yes | | | Yes | Shareware |
| ScramDisk | | Yes | | Yes | Shareware |
| MP3Stego | Yes | | | Yes | Open Source |
| StegoWav | | Yes | | Yes | Open Source |
| Hide4PGP | Yes | | VOC | Yes | Open Source |
| Steghide | | Yes | AU | Yes | Open Source |
| S-Tool | | Yes | | Yes | Open Source |
| Invisible Secrets | | Yes | | Yes | Commercial |
| Paranoid | | | Yes | Yes | Commercial |
| Commercial Stegnos | | Yes | VOC | Yes | Commercial |

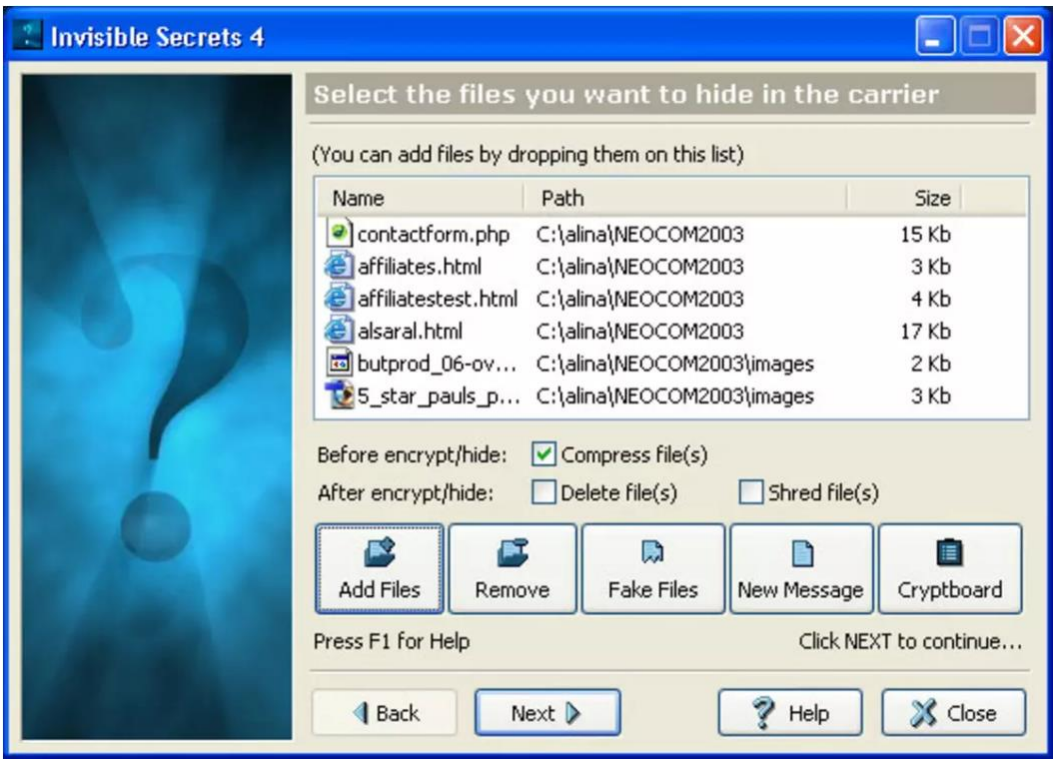
Different type of cover mediums (image, audio, and text) are used in Invisible Secret tools which is a commercial product. In the proposed system, hidden message are embedded within the audio file by using Invisible Secret tools for analysis process.

### 2.4.1. Invisible Secret

East-tec Invisible Secret is a revolutionary steganography and file encryption software that supposed to hide files and sensitive information. By using the Invisible Secret tools can encrypt file content and data hiding. File encryption can protect the confidential data from unauthorized user. Data hiding technique help to prevent data breaches by safely encrypting and embedding the data within cover media for data transmission and storage.

Password protection process is the powerful lock program that can protect the application from computer and prevent the unauthorized access. Password managing process which can manage your login data and generating process take place in the creating of highly secure passwords. Email encryption process can protect the confidential data from your email system and control the access to your personal information. This tool is suitable for various types of operating system and version 4.8 is released from 2014.


(a)




**(b)**

**Figure 2.3. User Interface of Invisible Secret Tools**

Figure 2.3 (a) and (b) describes the user interface of Invisible Secret tools which are available on the Internet for different type of operating system.

## 2.5. Proposed Audio Steganography Detection System

With the increasing attention on multimedia security, various MP3 steganographic and steganalytic algorithms have been offered more and more. The existing MP3 steganalysis is absence of good universality and detection performance. The aim of the proposed system is to apply one of the data analysis approaches in MP3 steganalysis.

The proposed system extracts the three different types of audio features from MP3 file, which are Mel-Frequency Cepstral Coefficient, Zero Crossing Rate, and Short Time Energy features. The extracted features are used to convert the dummy categorical variables for the next stage of MP3 steganography detection. In the process of steganography detection, Mutual Information (MI) is utilized based on the extracted categorical variables of audio features.

The proposed system tested the steganalysis technique on MP3 signals embedded with one of the MP3 steganographic tools which are available from the Internet. Steganographic tools will hide information in MP3 files during the compression process. The data is first compressed, encrypted and then hidden in the MP3 bit stream. This system used the input MP3 Signals as hidden message with MP3 signal which is embedded based on the stenographic tools and pure MP3 signal.

### 2.5.1. Process Diagram of the Proposed System

The proposed audio steganalysis system based on mutual information approach is divided into two phases: training phase (shown in Figure 2.4 and 2.5) and testing phase (shown in Figure 2.6).

In the training phase, it is needed to train input MP3 audio files with both original audio and stego audio. With an original audio, an original MP3 audio file is used as an input to the system. The input MP3 audio file is firstly converted to the audio WAV file format of 10 second length. After converting to the WAV file format, Mel-Frequency Cepstral Coefficient, Zero Crossing Rate, and Short Time Energy features are extracted. Then, the extracted features are converted to related categorical variables (statistical data type). Finally, data (converted categorical variables) are stored in the excel sheet.

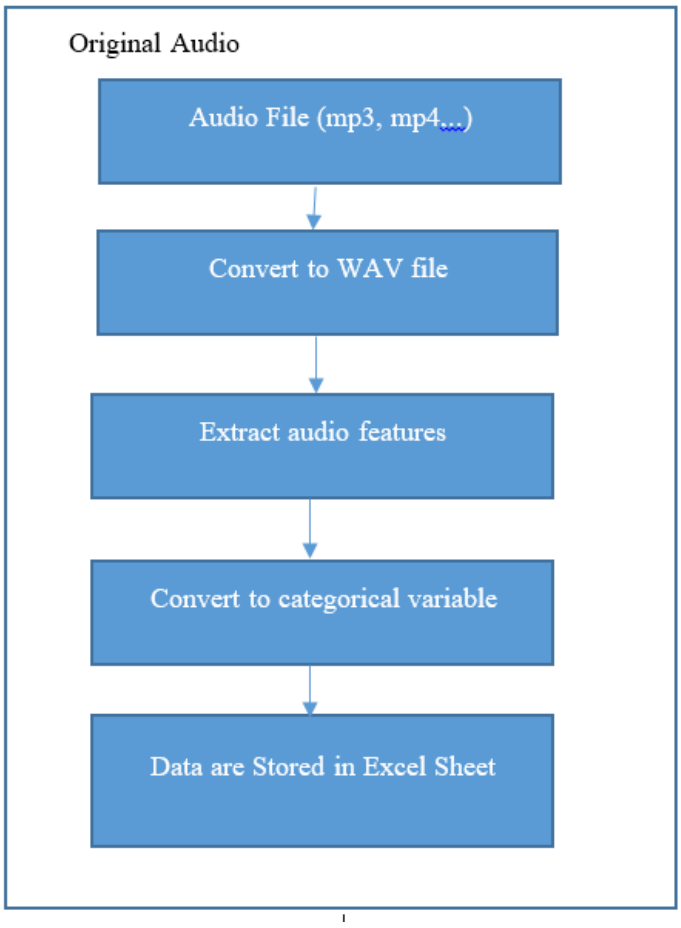



**Figure 2.4. Process Diagram of Audio Steganalysis System: Original Audio**

To train a stego audio, an original MP3 audio file is also used as an input to the system. The input MP3 audio file is firstly converted to the audio WAV file format of 10 second length. In training a stego audio, an additional step is needed to embed hidden message in the audio WAV file before the feature extraction process. After embedding hidden message, Mel-Frequency Cepstral Coefficient, Zero Crossing Rate, and Short Time Energy features are extracted from the stego audio with hidden message. Then, the extracted features are converted to related categorical variables (statistical data type) as in training an original audio file. Finally, data (converted categorical variables) are stored in the excel sheet.

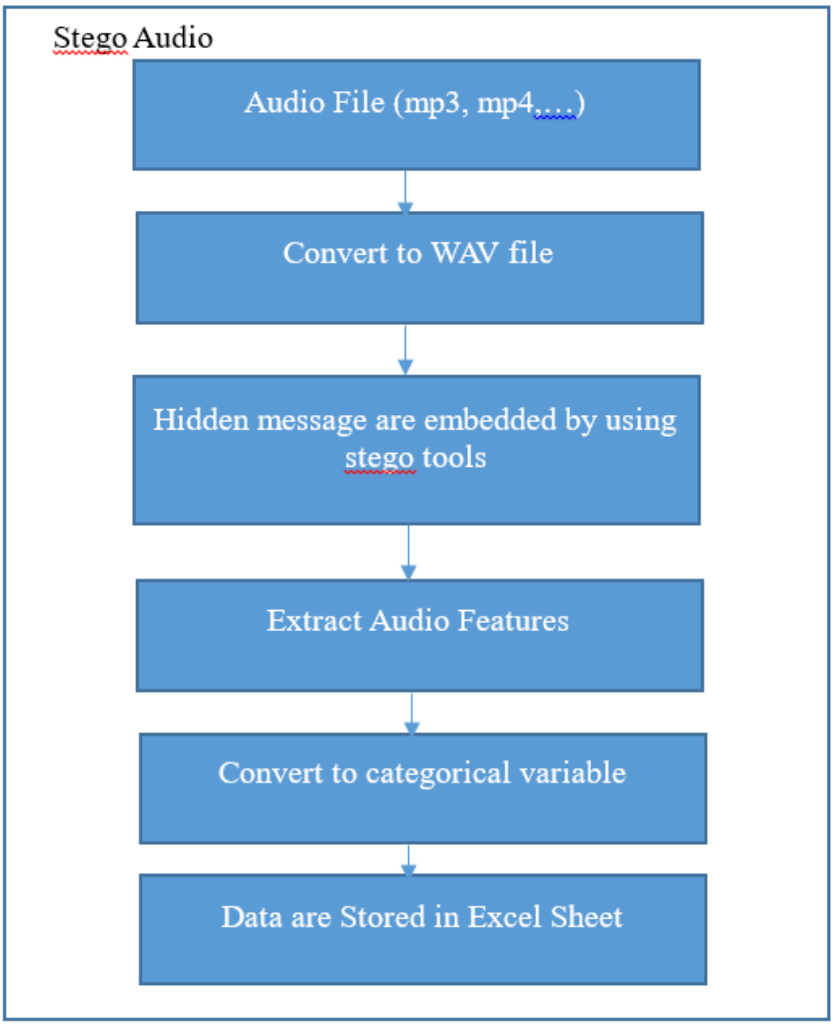



**Figure 2.5. Process Diagram of Audio Steganalysis System: Stego Audio**

In the testing phase, an MP3 audio file to be tested is used as an input to the system. The input MP3 audio file is firstly converted to the audio WAV file format of 10 second length as in the training phase. After converting to the WAV file format, Mel-Frequency Cepstral Coefficient, Zero Crossing Rate, and Short Time Energy features are extracted. Then, the extracted features are converted to related categorical variables (statistical data type). The next step is to find the relationship within the testing audio file and training audio files based on Mutual Information (MI). As a final point, the testing audio is labeled as “stego” or “without stego” based on the relationship within the testing audio file and training audio files obtained from the previous step. Mutual information can also be

computed for extracted features. Let be a variable regrouping k feature (k<p). The Mutual Information (MI) of Y with respect to X is computed as above, using natural extension of Mutual Information (MI) equations. Among all group of k features, there have the highest mutual relationship within stego or original features in training dataset. Bivariate mutual information is equal to zero if the two features are independent and there is no relationship each other.

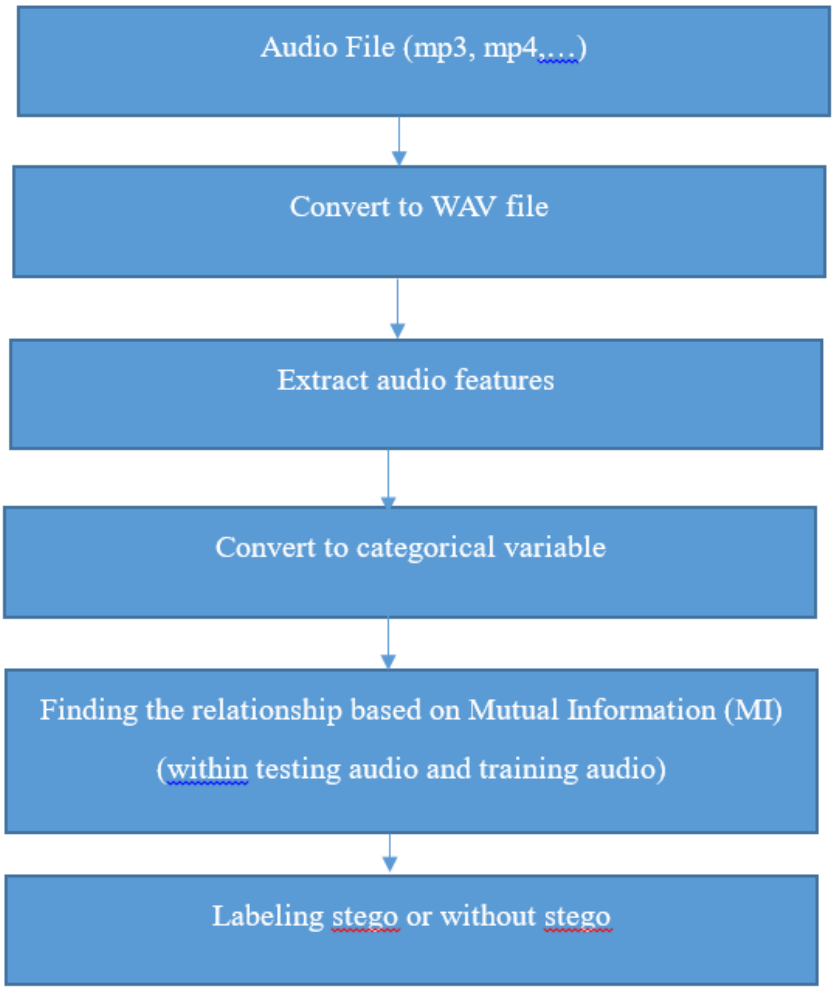



**Figure 2.6. Process Diagram of Audio Steganalysis System**

CHAPTER 3

# METHODOLOGIES

Steganography is a data hiding technique which embed hidden message in cover mediums (audio, video, image and text). The aim of steganography is for hiding the secret message from an intermediary. Usually, the least significant bits are used as a way of embedding message into image. Audio steganography also use least significant bits (LSB) algorithm. But, what makes audio steganography unlike from image steganography is audio steganography needs masking because human ears are sensitive to any change in audio. Masking can exploit the properties of human sensory organ for hiding information unnoticeably [2].

## 3.1. Related Methodologies of the Proposed System

Steganalysis is the science of identifying the occurrence of hidden data in the cover media files and is developing in competition with steganography. It has gained reputation in national security and forensic sciences since hidden messages detection can lead to the avoidance of disastrous security incidents [3]. In the earlier years, quite a lot of steganalysis techniques were offered for detecting the information-hiding behaviors in multiple steganography systems [1].

To identify the information-hiding in digital audios, X. Ru et al. introduced detection method by computing the features between the signal and a self-generated reference signal via linear predictive coding [2, 3]. Ismail Avcibas proposed the content-independent distortion measures as features in linear regression classifier design [8]. M. Johnson et al. introduced a statistical model by constructing a linear basis that captures certain statistical properties of audio signals with non-linear support vector machine classifier [9].

Fatiha Djebbar and Beghdad ayad [4] proposed continuous homogeneous energy-based audio stream segments to embed and produce a set of meaningful features and applied support vector machine classifier.

## 3.2. Features Extraction Process

Feature extraction is the procedure of transforming an audio signal into a sequence of feature vectors carrying representative information about the signal [17]. For the audio signal processing, two types of audio features are extracted from the questionable files. There are time domain based feature and frequency domain based features in Figure 3.1 which are mostly popular for classification and prediction system.

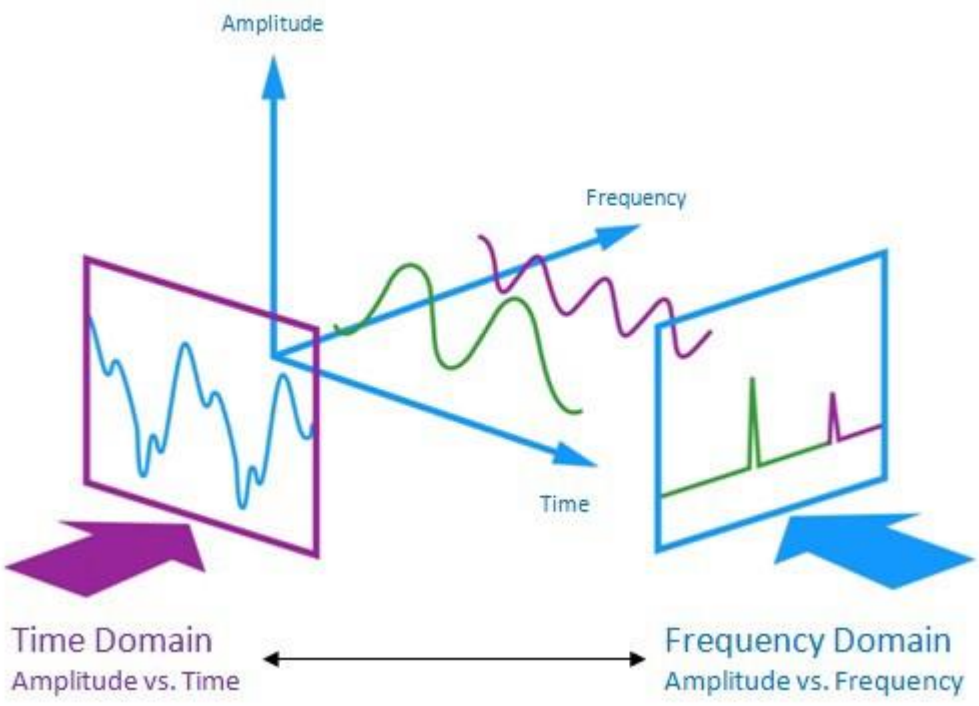



**Figure 3.1. Domain Based Features**

Audio features are categorized with level abstraction, temporal scope, musical aspect, signal domain and ML approach. In the proposed system, three type of audio features are extracted for covering the level of audio signal (high level, mid-level and low level).

Most audio detection and classification systems combine with two processing stages: feature extraction followed by detection and classification. In the proposed system, three types of features: “Mel-Frequency Ceptral Coefficient, Zero Crossing Rate, and Short Time Energy” which are computed from MP3 signal.

### 3.2.1. Mel-Frequency Cepstral Coefficient (MFCC)

Mel-frequency cepstral coefficients (MFCC) have been the principal features used for speech recognition. Their achievement has been due to their capability to characterize the speech amplitude spectrum in a compacted form [12].

The mapping between the frequency scale (Hz) and become aware of frequency scale (mels) is approximately linear below 1 kHz and logarithmic at higher frequencies. The following formula can evaluate the relationship

$$F_{mel} = 2595.log_{10}\left(1 + \frac{F_{Hz}}{700}\right) \quad 3.1$$

Where $F_{mel}$ is the collected frequency in mels and FHz is the frequency in Hz. The bandwidth and the spacing of these critical-band filters are invariable values, 300 mels and 150 mels within the mel-frequency domain. X(m) is the power spectrum of input signal, S[k] is the power in k-th critical band and M represents the number of the critical bands in mel scale, ranging usually from 20 to 24. Then,

$$S[k] = \sum_{j=0}^{f/2-1} W_k(j).X(j),\ k = 1, \ldots\ldots, M \quad 3.2$$

where $W_k$ is the critical-band filter. Let L denote the desired order of the MFCC. Then, we can find the MFCCs from logarithm and cosine transforms as follows:

$$C[n] = \sum_{k=1}^{M} \log(S[k]) \cos\left[(k - 0.5)\frac{n\pi}{M}\right],\ n = 1, \ldots L \quad 3.3$$

Figure 3.2 shows the procedure of generating MFCC features. To generate a cepstral feature vector for each frame, the earliest phase is to split the speech signal into frames, usually by applying a windowing function (removes edge effects) at fixed intervals. The succeeding phase is to take the Discrete Fourier Transform (DFT) of each frame.

The next phase is to smooth the spectrum and give emphasis (Mel-scaling) to perceptually meaningful frequencies. Mel scale is a scale that relates the perceived frequency of a tone to the actual measured frequency. It scales the frequency in order to match more closely what the human ear can hear [4]. The components of the Mel-spectral

vectors calculated for each frame are highly correlated. The latest phase of MFCC feature generation is to apply a Discrete Cosine Transform to the Mel-spectral vectors which decorrelates their components [12]. Figure 3.2 describes the generating process of MFCC features.

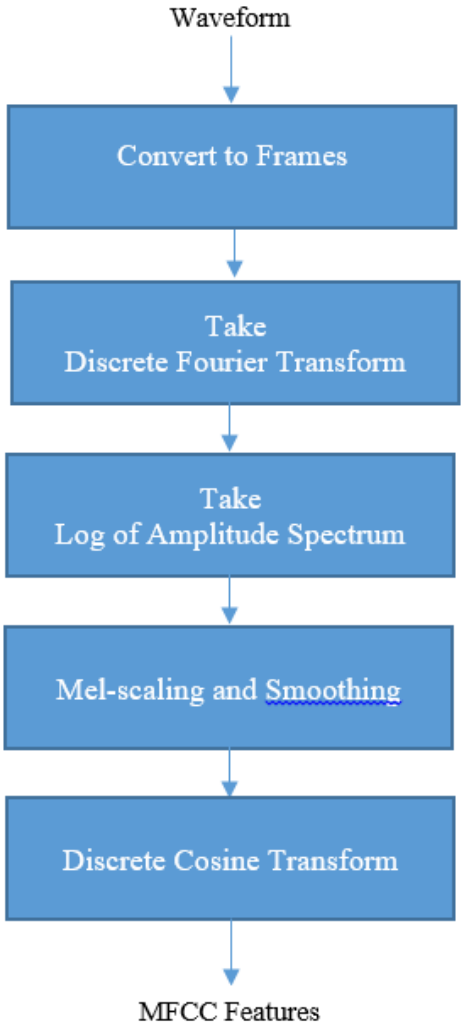



**Figure 3.2. Procedure of Generating MFCC Features**

Figure 3.3 shows an example of MFCC feature extraction from audio signals. The MFCC features are extracted from audio signal using Open SMILE toolkit [15]. The audio frames of 100 ms sampled are set at a rate of 50 ms using Hamming window. The frequency range of the Mel-spectrum is set from 0 to 8 kHz.

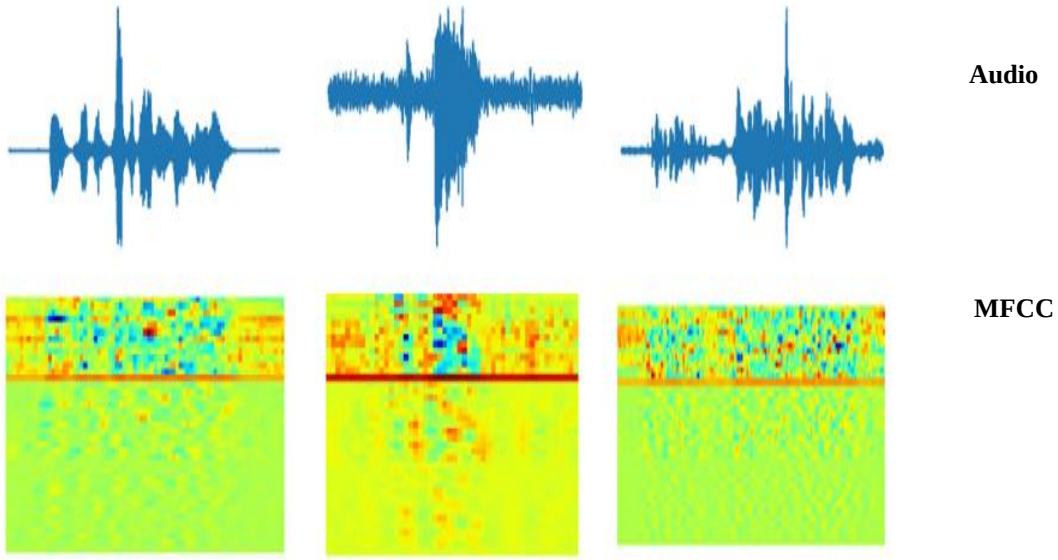



**Figure 3.3. Example of Audio Conversion to MFCC**

## 3.2.2. Zero Crossing Rate

There are different algebraic sign in a zero crossing rate. It is evaluate the frequency content of input audio signal. Moreover it used to analyze the signal passing through a value of zero which is varied in a given time and amplitude of the signal [16]. It can be used to measure the noisiness of a signal, as its values are higher for the noisy portions of the signal. The short time Zero Crossing Rate (ZCR) is defined according to the following equation,

$$ZCR = \frac{1}{N}\sum_{n-1}^{N}\frac{|sgn\{x(n)\}-sgn\{x(n-1)\}|}{2} \quad 3.4$$

where sgn() is the sign function and x(n) is the signal of length n, N is the window length. Figure 3.4 shows sample WAV audio signals of 10s length (all signals are sampled at 44.1 kHz) and their respective Zero Crossing Rate (ZCR) features.

Zero crossing are a basic property of an audio signal that is often employed in audio classification. Zero Crossing allow for a rough estimation of dominant frequency and the spectral centroid.

### 3.2.3. Short Time Energy

Short-term energy is the primary and most ordinary feature that has been applied. Substantially, energy is a measure of how much signal there is at any one time. Energy is used to determine voiced sounds, which have higher energy than silence/un-voiced, in a continuous signal. The energy of a signal is usually computed on a short- time basis, by windowing the signal at a particular time, squaring the samples and taking the average [18]. The short-time energy function of a signal frame with length N is defined as:

$$E = \frac{1}{N}\sum_{n=0}^{N-1} x^2(n) \qquad 3.5$$

where x(n) is the signal of length n and N is the window length.

## 3.3. Dummy Coding

This coding is developed by Cohen and Cohen in 1983, which is the simplest coding structure that supposed to examine group mean differences. A dummy variable is a numerical variable for representing group behavior. For a categorical variable with multiple levels (*n*), *n*-1 numbers of dummy variables are required to represent it. Steganography detection system assumes that the independent features are numerical data. These are converted to the categorical variable for reducing the effect of redundancy within the extracted features of audio signals.

**Advantages of Dummy Variables**

Dummy coding has the valuable advantages for representing categorical variable. This is suitable for nominal data and used to get dichotomous data. The most efficient one is "to avoid a biased assessment of the impact of an independent variable, as a consequence of omitting another independent variable that is related to it". A second benefit to such a coding structure is its ease of interpretation.

## 3.4. Mutual Information

The Mutual Information (MI) between two random variables or random vectors measures the "amount of information", that is the "loss of uncertainty" that one can bring

to the knowledge of the other, and vice versa. The theory of uncertainty of a random variable is stated by its entropy. Although the opinion of entropy has first been developed for discrete variables, it can be extended to continuous variables rather easily [p19].

The entropy $H(Y)$ of a random variable Y with probability density function (pdf) $p_Y$is defined by,

$$H(Y) = -\int p_Y(y) log p_Y(y) dy \quad 3.6$$

The entropy of a random variable or vector $Y$ when the value of some other random variable X is known is the conditional entropy:

$$H(Y|X) = -\int p_X(x) \int p_Y(y|X = x) log p_Y \quad 3.7$$
$$(y|X = x) dy dx$$

The mutual information is the difference between the entropy of a variable and the conditional entropy $I(X,Y) = H(Y) - H(Y|X)$. Given two random variables x and y, their mutual information is defined in terms of their probabilistic density functions $p(x), p(y) and p(x,y)$:

$$I(x;y) = \iint p(x,y) log \frac{p(x,y)}{p(x)p(y)} dx dy \quad 3.8$$

During the evaluation of maximum relevance features, to have the largest mutual information $I(x_i;c)$with the binary classification (stego or not), reflecting the largest dependency on the target class. In terms of sequential search, the $m$ best individual features, i.e., the top $m$ features in the descent ordering of$I(x_i;c)$ , are often selected as the m features [20].

In the mutual information approach, Mutual Information (MI) values of each of the tested audio features $X_i$ with dependent variables Y (stego or without stego) are computed. The features with highest mutual information of the dependent variable Y is then used and labeled as represent to the test audio file. Figure 3.4 shows Mutual Information Analysis for Steganographic Detection.

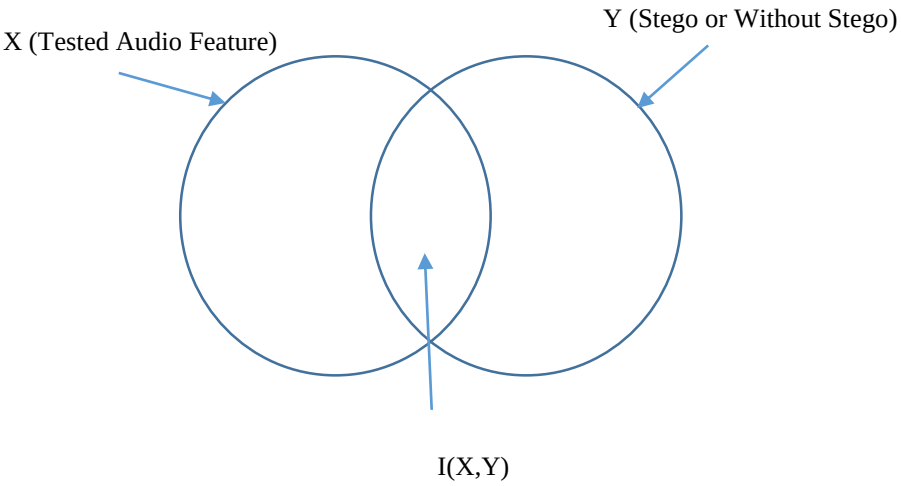



**Figure 3.4. Mutual Information Analysis for Steganographic Detection**

### 3.4.1. Estimation of Mutual Information

Mutual information is based on conditional entropy that can measure the statistical difference between two variables X and Y. $I_{X,Y}$ estimates the average amount of information on the actual value of Y supplied by using the information of explanatory features X: $I_{X,Y} = H_Y - H_{Y/X}$ . Normalized by the entropy of variable Y, the mutual information ratio (MIR), $R_{X,Y}$ , is a zero-to-one range measure of the dependency of X and Y : $R_{X,Y} = I_{X,Y}$ HY . For two explanatory features of X and Y, actual information of X doesn't support to any information on Y and $R_{X,Y} = 0$.

CHAPTER 4

# THE IMPLEMENTATION OF THE PROPOSED SYSTEM

In the process of audio steganalysis system, choosing the features of audio signal and selecting the natures of audio domain are the main challenge for that. There are different types of audio features are available in steganalysis tasks.

This study showed that the optimal features depend on the domain and steganalysis techniques. Features based audio steganalysis system has been developed for this thesis, which takes several audio quality measures namely Mel Frequency Cepstral Coefficient (MFCC), Short Time Energy (STE) and Zero Crossing Rate (ZCR) are applied in audio steganalysis.

## 4.1. Preprocessing

There are different characteristics of audio features which are not suitable to put into a feature vector. Each feature component should be normalized to get the better analysis with similar scale. The extraction of numerical feature vectors that characterize the audio content which are the fundamentals step of audio steganalysis system. In order to get high accuracy for audio steganalysis, it is important to apply the most relevance features that need to extract the temporal and spectral characteristics of audio signal.

In this study, audio clip level features are computed based on the frame level features and a clip is used as the basic analysis unit. The audio quality with the sampling frequency of 44.1 KHz, bit rate of 128 kbps and stereo channel. Analyzing into 10 sec audio file with original signal and hidden message are embedded in the audio file by using Invisible Secret tools, two datasets are produced: training and testing (Tr and Ts). Each data set contains 180 WAV audio signals. Each dataset contains 90 audio sample with hidden message and 90 audio sample without hidden message. All stego-audio signals are generated by hiding data from different text files. The stego signals produced by Invisible Secret tools.

## 4.2. Analysis Results for Steganalysis

Feature-based steganography detection system performs well for Invisible Secret tools. The proposed system can detect the audio signal which modifies the characteristics of audio signal. From the analysis it can be seen that Mutual Information (MI) perform well for the presence of secret message of features based steganalysis.

The proposed Mutual Information (MI) approach for evaluating the stego and without setgo ones which constructs the superiority of this approach compared to the passive steganalysis of Invisible Secret. Experimental results tested in three types of music (POP, ROCK, COUNTRY) which are shown in Table 4.1.

**Table 4.1. Description for Data Analysis**

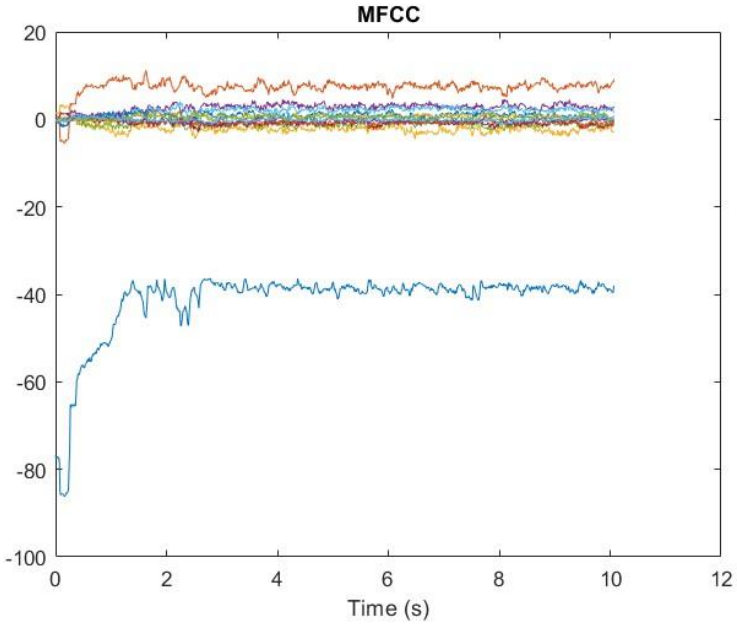
| Types of Song | Original Audio file | | Stego Audio file (embedded with Invisible Secret tools) | | Hidden Message (.txt file) |
|---|---|---|---|---|---|
| | Training (10 sec) | Testing (10 sec) | Training (10 sec) | Testing (10 sec) | |
| Pop | 10 | 10 | 30 | 10 | Different types of sentence/ file size |
| Rock | 10 | 10 | 30 | 10 | |
| Country | 10 | 10 | 30 | 10 | |

### 4.2.1. Mel Frequency Cepstral Coefficients

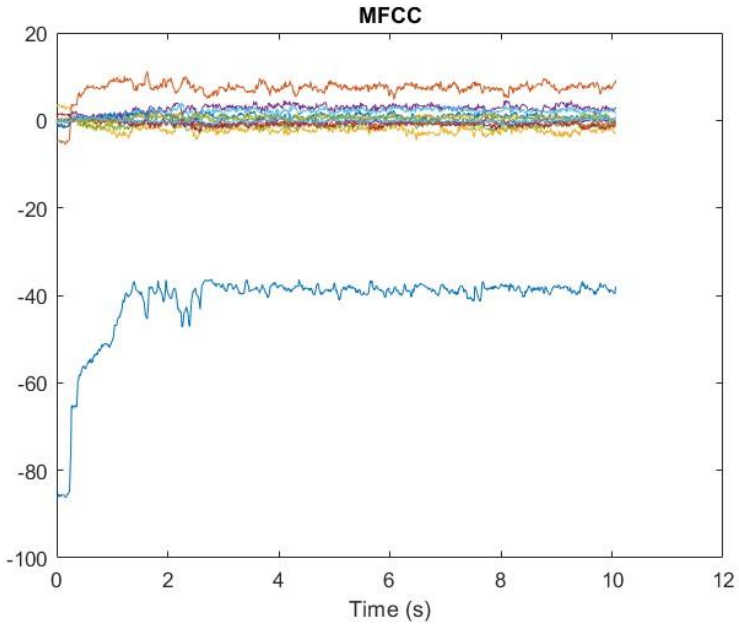
A time series of MFCC vectors are computed by iterating over the audio file resulting in fourteen coefficients. Figure 4.1 describes the extracted MFCC features from one of the audio file which has (14) MFCC features. The actual features used for mutual analysis task were the means of the MFCCs taken over a clip. In this way a very compact data set was created. The following Figures (4.2, 4.3 and 4.4) show the plots of the stego and original signals as a function of time together with their respective MFCCs.

| MFCC1 | MFCC2 | MFCC3 | MFCC4 | MFCC5 | MFCC6 | MFCC7 | MFCC8 | MFCC9 | MFCC10 | MFCC11 | MFCC12 | MFCC13 | MFCC14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3.5655 | -6.9866 | 2.7054 | -0.30815 | 0.38273 | -0.0434 | 0.38681 | -0.23894 | 0.14944 | 0.13464 | -0.15459 | -0.15086 | 0.29556 | 0.32516 |
| 4.4845 | -6.599 | 3.1027 | -0.09792 | 0.68427 | -0.0582 | 0.58715 | -0.01705 | 0.17109 | 0.29254 | 0.33346 | 0.13757 | 0.18008 | 0.45939 |
| 4.8336 | -6.4121 | 2.5173 | -0.2607 | 0.3355 | -0.05086 | 0.93223 | 0.13392 | 0.13587 | 0.46617 | 0.093572 | -0.04049 | -0.01984 | 0.52342 |
| 4.6952 | -5.9635 | 2.3886 | -0.17294 | 0.2442 | -0.17792 | 0.82332 | 0.22993 | 0.079082 | 0.22667 | 0.004223 | 0.13893 | -0.1315 | 0.59382 |
| 4.2913 | -6.0681 | 2.2461 | -0.18507 | 0.37288 | -0.26851 | 0.64284 | -0.11976 | -0.07634 | 0.25648 | 0.14611 | 0.059892 | -0.37164 | 0.20121 |
| 4.076 | -5.7743 | 2.1831 | -0.20081 | 0.3045 | -0.2794 | 0.64177 | -0.29874 | 0.11602 | 0.39438 | 0.12809 | -0.02608 | -0.2827 | 0.46173 |
| 4.2414 | -5.5057 | 2.2914 | -0.19383 | 0.3785 | -0.32872 | 0.83409 | -0.37801 | 0.098823 | 0.42159 | 0.076434 | -0.20991 | -0.22081 | 0.6159 |
| 4.336 | -5.3965 | 2.443 | -0.28394 | 0.089635 | -0.44755 | 1.117 | -0.17081 | 0.016285 | 0.7128 | 0.028854 | -0.32275 | -0.26118 | 0.58144 |
| 4.3109 | -5.1312 | 2.5226 | -0.482 | -0.0193 | -0.25155 | 0.97627 | -0.01471 | 0.018644 | 0.69457 | -0.00049 | -0.03449 | -0.21384 | 0.73893 |
| 4.2851 | -5.3566 | 2.3864 | -0.41661 | -0.03397 | -0.16672 | 1.0332 | -0.05404 | -0.10093 | 0.69808 | -0.00019 | -0.08302 | -0.3917 | 0.73236 |
| 4.1032 | -5.5315 | 2.1349 | -0.48696 | 0.011132 | -0.25927 | 0.90275 | -0.01226 | -0.0254 | 0.69307 | 0.072698 | -0.05035 | -0.33574 | 0.75344 |
| 3.9453 | -5.6883 | 2.0864 | -0.50099 | 0.13128 | -0.14431 | 0.98746 | -0.01513 | -0.0632 | 0.62453 | -0.29605 | -0.25129 | -0.32494 | 0.69484 |
| 3.7746 | -5.8508 | 2.2054 | -0.26306 | 0.24445 | -0.41348 | 1.1187 | 0.05891 | 0.19931 | 0.8607 | -0.02622 | -0.25587 | -0.27556 | 0.65276 |
| 3.7958 | -6.054 | 2.036 | -0.34916 | 0.15267 | -0.49879 | 0.82244 | 0.059065 | 0.024287 | 0.61761 | 0.11591 | -0.14375 | -0.52774 | 0.41761 |
| 3.8367 | -6.0096 | 2.0592 | -0.33347 | -0.16703 | -0.6313 | 1.1151 | 0.31376 | -0.00369 | 0.66234 | 0.24663 | -0.02016 | -0.46939 | 0.58098 |
| 3.9713 | -5.7094 | 2.2282 | -0.51659 | -0.19654 | -0.54898 | 1.0656 | 0.15883 | 0.1014 | 0.51617 | 0.32078 | -0.01601 | -0.34204 | 0.64041 |
| 3.9882 | -5.6147 | 2.1269 | -0.36673 | -0.01456 | -0.47866 | 1.0505 | 0.10274 | 0.032646 | 0.63434 | 0.19708 | -0.19554 | -0.29058 | 0.65861 |
| 3.9478 | -5.6884 | 2.0449 | -0.46901 | 0.077561 | -0.16877 | 1.2605 | -0.1166 | -0.24927 | 0.41502 | 0.085145 | -0.28175 | -0.38328 | 0.55852 |
| 4.0789 | -5.5957 | 2.1727 | -0.41205 | 0.20635 | -0.36727 | 0.88553 | -0.34033 | 0.017822 | 0.42507 | -0.03729 | -0.09564 | -0.43155 | 0.36364 |
| 4.1235 | -5.4973 | 2.3479 | -0.31897 | 0.23343 | -0.37114 | 1.0138 | -0.1854 | -0.13153 | 0.49143 | 0.084932 | -0.19448 | -0.27409 | 0.48625 |
| 4.1225 | -5.5269 | 2.0964 | -0.28121 | 0.166 | -0.64327 | 0.91627 | -0.23069 | -0.13619 | 0.52647 | 0.14934 | -0.34491 | -0.31337 | 0.59493 |
| 3.9496 | -5.6252 | 2.0882 | -0.17458 | -0.03001 | -0.4654 | 0.83456 | -0.26943 | -0.10106 | 0.50628 | -0.08421 | -0.57377 | -0.61913 | 0.6267 |
| 3.9524 | -5.5042 | 1.9879 | -0.24527 | 0.060441 | -0.51472 | 0.71656 | -0.30911 | -0.09249 | 0.41824 | -0.35611 | -0.55705 | -0.55247 | 0.77636 |
| 4.1419 | -5.3035 | 2.1004 | -0.21711 | -0.03674 | -0.50723 | 0.76493 | -0.18496 | 0.11943 | 0.53351 | -0.26121 | -0.46045 | -0.46426 | 0.88589 |
| 4.2387 | -5.2118 | 2.1456 | -0.34615 | -0.04169 | -0.44508 | 0.83586 | -0.21771 | -0.08804 | 0.32072 | -0.09304 | -0.36654 | -0.63588 | 0.74489 |

**Figure 4.1. Sample Extracted Training Features for MFCC**

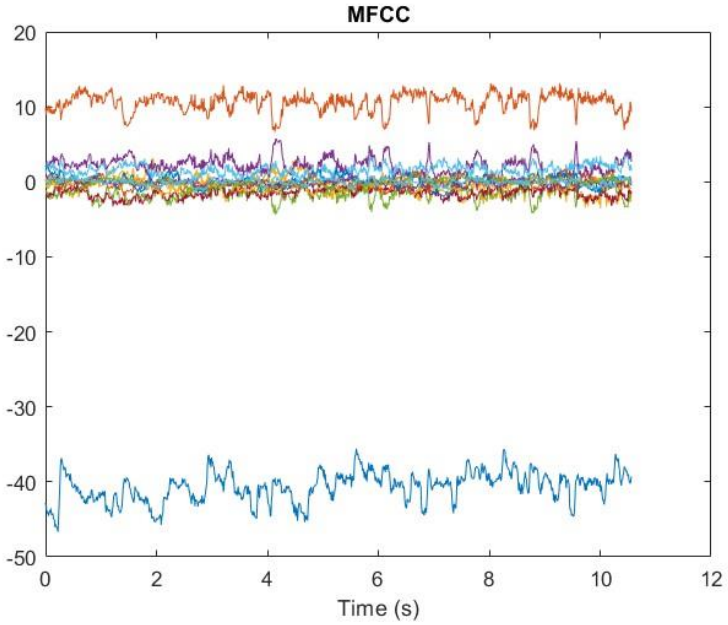



**(a) Original**

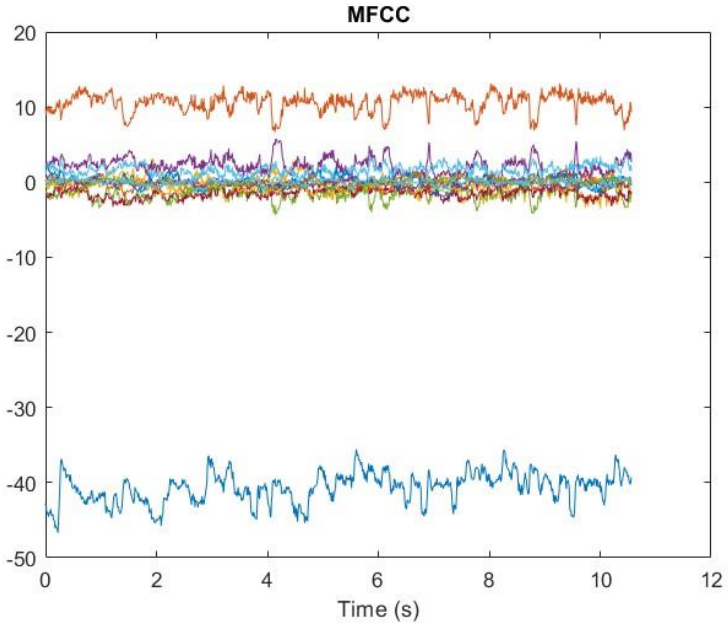



**(b) Stego**

**Figure 4.2. MFCCs for Rock Song**

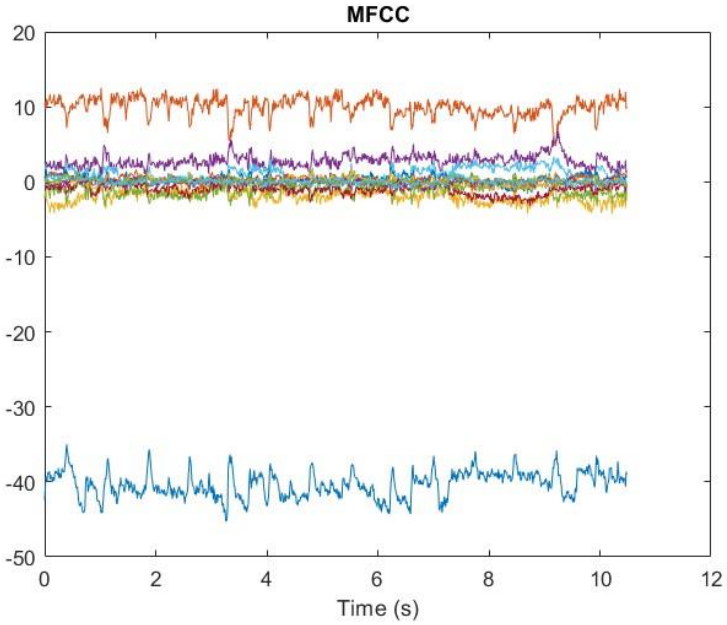



**(a) Original**

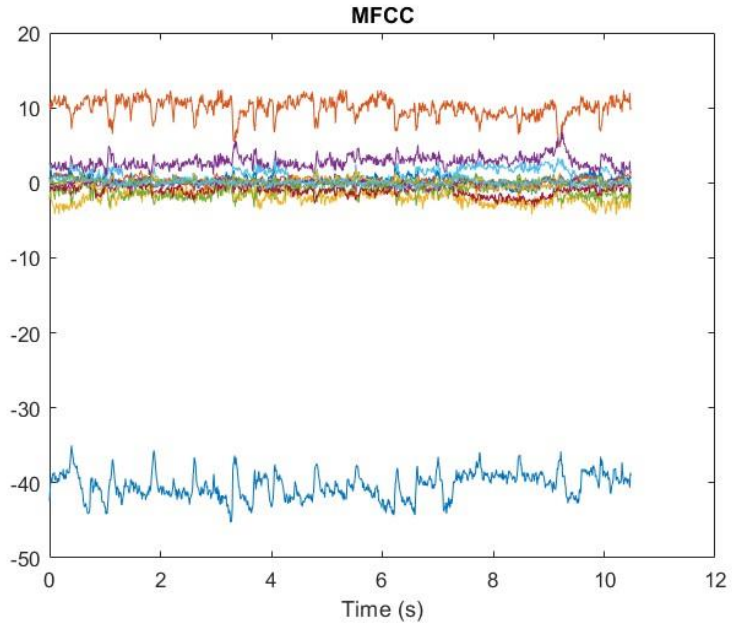



**(b) Stego**

**Figure 4.3. MFCCs for Pop Song**




**(a) Original**




**(b) Stego**

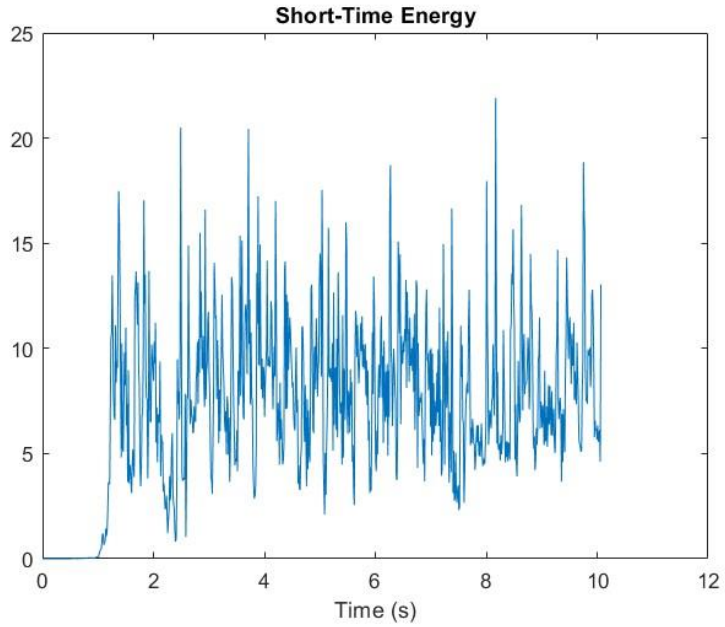
**Figure 4.4. MFCCs for Country Song**

### 4.2.2. Short Time Energy

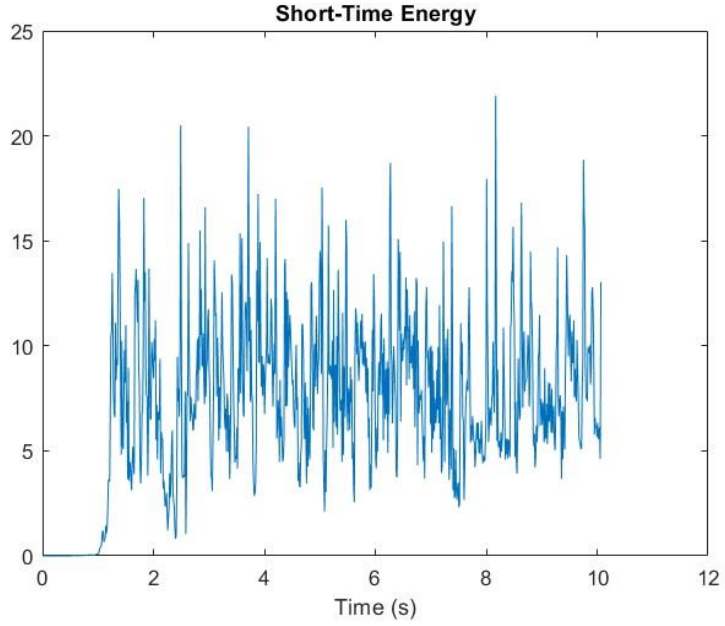
The STE measurement of an audio signal can be used to determine stego and original. The amplitude of the audio signal varies with music types. Generally, the amplitude of pop music is much lower than the amplitude of rock music. The energy of the signal provides a representation that reflect these amplitude variations. In the ratio of energy method, original is decided when the energy is high. Usually, stego has a low short term entry but a high zero crossing rate. Figures (4.6, 4.7 and 4.8) describe the plots of STE features which are extracted from different types of audio song. It can be seen that nature of this two signals are slightly varied. The following Figure 4.5 shows the sample extracted training features for STE. Two types of STE features are extracted from one of the audio song.

| | BUD | BUE | BUF | BUG | BUH | BUI | BUJ | BUK | BUL | BUM | BUN |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | STE | STE | STE | STE | STE | STE | STE | STE | STE | STE | STE |
| 2 | 0.002616 | 0.006941 | 0.003037 | 0.006941 | 0.003037 | 0.006941 | 0.003037 | 0.00824 | 0.014103 | 0.00824 | 0.014103 |
| 3 | 0.014621 | 0.006913 | 0.002878 | 0.006914 | 0.002878 | 0.006914 | 0.002878 | 0.014393 | 0.013531 | 0.014393 | 0.013531 |
| 4 | 0.010963 | 0.013495 | 0.007427 | 0.013495 | 0.007427 | 0.013495 | 0.007427 | 0.008288 | 0.004964 | 0.008288 | 0.004964 |
| 5 | 0.004215 | 0.004441 | 0.001537 | 0.004441 | 0.001537 | 0.004441 | 0.001537 | 0.028313 | 0.025788 | 0.028313 | 0.025788 |
| 6 | 0.004325 | 0.003306 | 0.003837 | 0.003306 | 0.003837 | 0.003306 | 0.003838 | 0.011717 | 0.014586 | 0.011717 | 0.014586 |
| 7 | 0.005353 | 0.004805 | 0.002868 | 0.004805 | 0.002868 | 0.004806 | 0.002868 | 0.006477 | 0.006098 | 0.006477 | 0.006098 |
| 8 | 0.002078 | 0.004608 | 0.00219 | 0.004608 | 0.00219 | 0.004608 | 0.00219 | 0.007408 | 0.008783 | 0.007408 | 0.008783 |
| 9 | 0.00189 | 0.003571 | 0.002072 | 0.003571 | 0.002072 | 0.003571 | 0.002072 | 0.002459 | 0.007872 | 0.002459 | 0.007872 |
| 10 | 0.006408 | 0.00536 | 0.003203 | 0.00536 | 0.003203 | 0.00536 | 0.003203 | 0.006761 | 0.006819 | 0.006761 | 0.006819 |
| 11 | 0.006088 | 0.005614 | 0.00311 | 0.005614 | 0.00311 | 0.005614 | 0.00311 | 0.002597 | 0.008307 | 0.002597 | 0.008307 |
| 12 | 0.008508 | 0.002925 | 0.002104 | 0.002925 | 0.002104 | 0.002925 | 0.002104 | 0.007206 | 0.007999 | 0.007206 | 0.007999 |
| 13 | 0.003844 | 0.005482 | 0.002325 | 0.005482 | 0.002325 | 0.005482 | 0.002325 | 0.011124 | 0.005289 | 0.011124 | 0.005289 |
| 14 | 0.004311 | 0.00623 | 0.002788 | 0.00623 | 0.002788 | 0.00623 | 0.002788 | 0.009598 | 0.006607 | 0.009598 | 0.006607 |
| 15 | 0.002545 | 0.002665 | 0.001609 | 0.002665 | 0.001609 | 0.002665 | 0.001609 | 0.005786 | 0.0051 | 0.005786 | 0.0051 |
| 16 | 0.002146 | 0.003421 | 0.002104 | 0.003421 | 0.002104 | 0.003421 | 0.002104 | 0.007424 | 0.006855 | 0.007424 | 0.006855 |
| 17 | 0.00195 | 0.005102 | 0.002492 | 0.005102 | 0.002492 | 0.005102 | 0.002492 | 0.00487 | 0.003204 | 0.00487 | 0.003204 |
| 18 | 0.003216 | 0.00334 | 0.002485 | 0.00334 | 0.002485 | 0.00334 | 0.002485 | 0.007971 | 0.010621 | 0.007971 | 0.010621 |
| 19 | 0.003773 | 0.005451 | 0.002662 | 0.005451 | 0.002662 | 0.005451 | 0.002662 | 0.008041 | 0.008915 | 0.008041 | 0.008915 |
| 20 | 0.002529 | 0.003583 | 0.00166 | 0.003583 | 0.00166 | 0.003583 | 0.00166 | 0.008198 | 0.009985 | 0.008198 | 0.009985 |
| 21 | 0.002332 | 0.003308 | 0.001946 | 0.003308 | 0.001946 | 0.003308 | 0.001946 | 0.008233 | 0.006765 | 0.008233 | 0.006765 |
| 22 | 0.002397 | 0.004177 | 0.002541 | 0.004177 | 0.002541 | 0.004177 | 0.002541 | 0.006649 | 0.009953 | 0.006649 | 0.009953 |
| 23 | 0.001538 | 0.003502 | 0.002674 | 0.003502 | 0.002674 | 0.003502 | 0.002674 | 0.00684 | 0.007421 | 0.00684 | 0.007421 |
| 24 | 0.001493 | 0.003809 | 0.002305 | 0.003809 | 0.002305 | 0.003809 | 0.002305 | 0.009856 | 0.005469 | 0.009856 | 0.005469 |
| 25 | 0.00278 | 0.005535 | 0.002221 | 0.005535 | 0.002221 | 0.005535 | 0.002221 | 0.006872 | 0.007173 | 0.006872 | 0.007173 |
| 26 | 0.00211 | 0.004748 | 0.002146 | 0.004748 | 0.002146 | 0.004748 | 0.002146 | 0.009898 | 0.006688 | 0.009898 | 0.006688 |
| 27 | 0.003146 | 0.004215 | 0.001311 | 0.004215 | 0.001311 | 0.004215 | 0.001311 | 0.009187 | 0.009618 | 0.009187 | 0.009618 |

**Figure 4.5. Sample Extracted Training Features for STE**

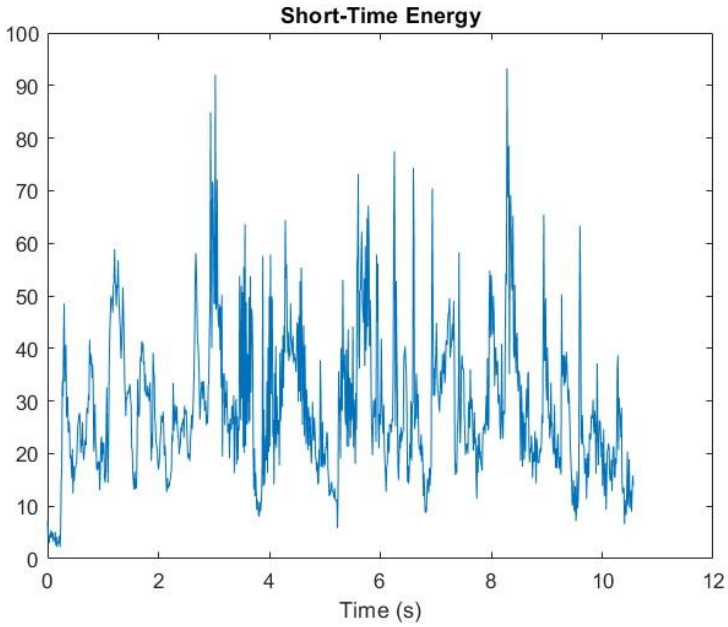



**(a) original**

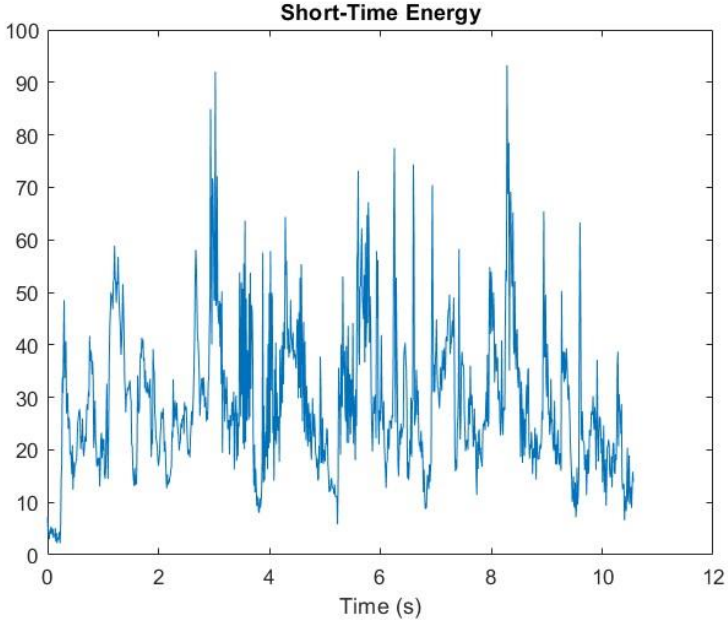



**(b) Stego**

**Figure 4.6. STE for Rock Song**

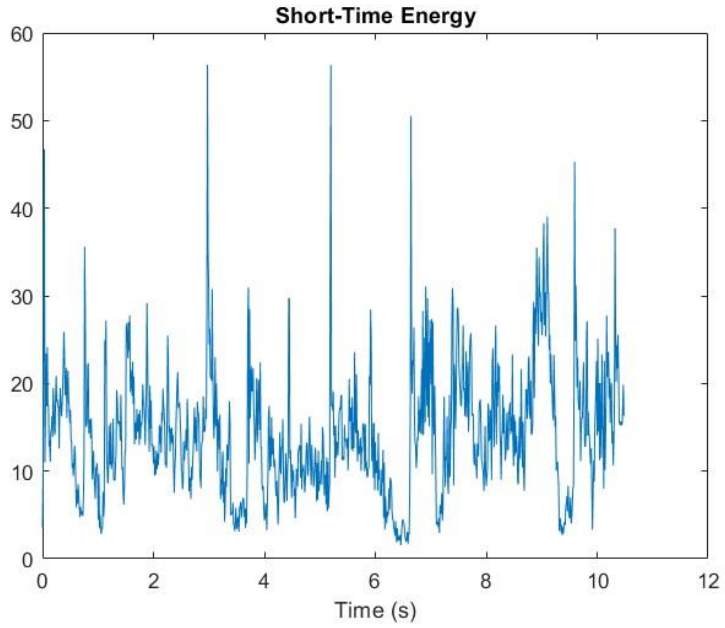



**(a) Original**

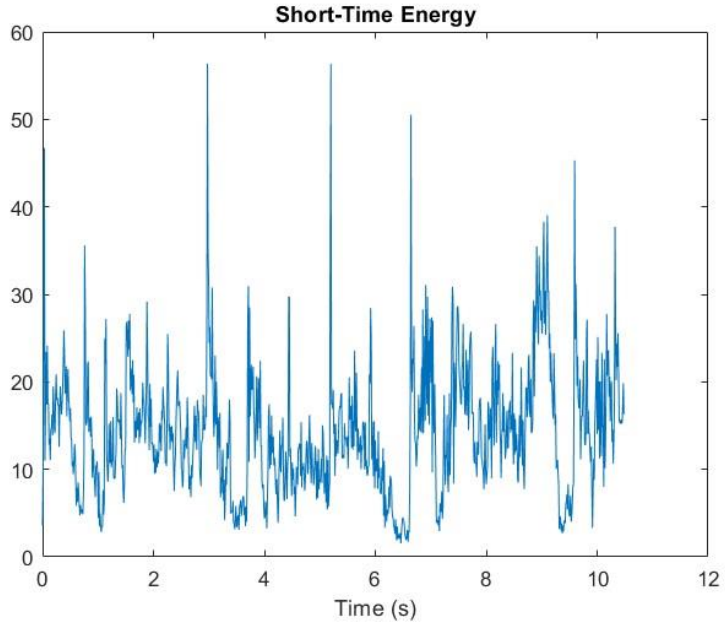



**(b) Stego**

**Figure 4.7. STE for Pop Song**




**(a) Original**




**(b) Stego**

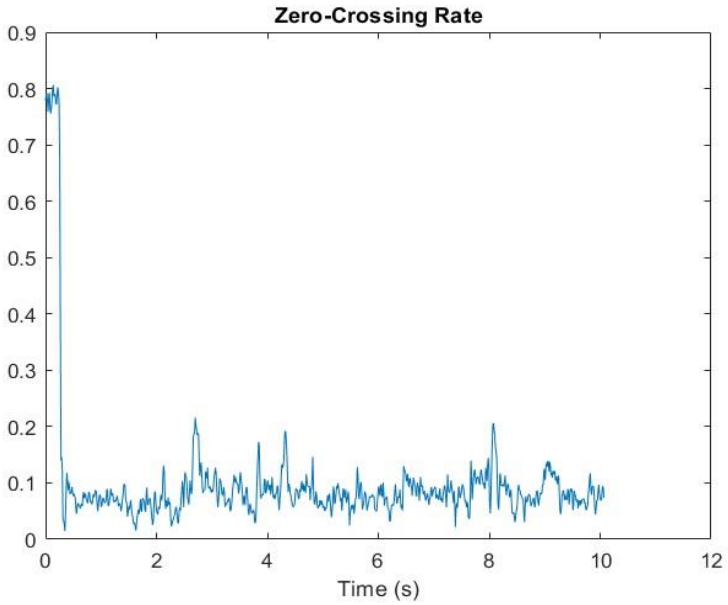
**Figure 4.8. STE for Country Song**

### 4.2.3. Zero Crossing Rate

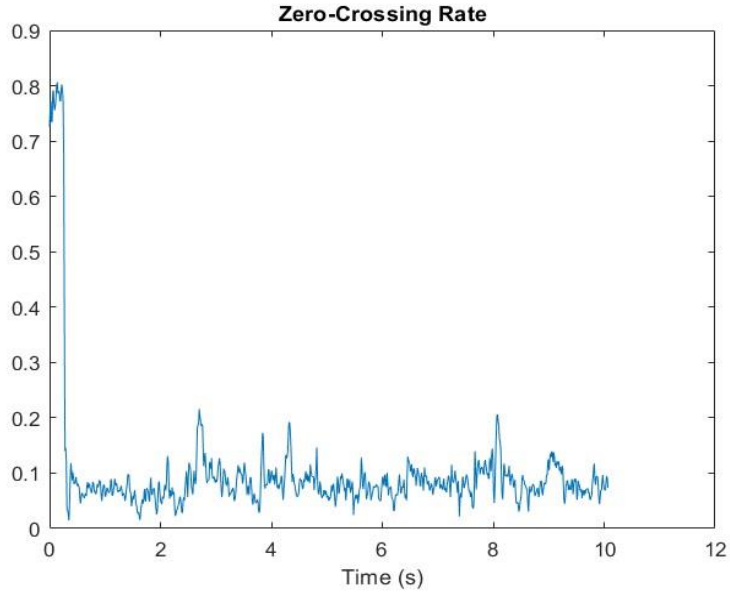
Zero Crossing Rate (ZCR) is computed as the total number of times that the waveform crosses the zero axis. It is useful to discriminate stego signal from detected signal. The Zero Crossing Rate (ZCR) feature was intended to be used either in conjunction with the MFCCs or independently. The following Figures (4.10, 4.11 and 4.12) show the plot of Zero Crossing Rate (ZCR) for stego and original signals extracted from 10 sec clips. Figure 4.9 describes the sample extracted training features of ZCR from original and stego files.

| BUW | BUX | BUY | BUZ | BVA |
|---|---|---|---|---|
| ZERO | ZERO | | | |
| 0.046896 | 0.047068 | | | |
| 0.074133 | 0.073255 | | | |
| 0.057154 | 0.057179 | | | |
| 0.037437 | 0.036029 | | | |
| 0.052936 | 0.052118 | | | |
| 0.048961 | 0.046912 | | | |
| 0.049673 | 0.043038 | | | |
| 0.054764 | 0.055174 | | | |
| 0.054709 | 0.048399 | | | |
| 0.064236 | 0.069081 | | | |
| 0.03494 | 0.038863 | | | |
| 0.062352 | 0.071594 | | | |
| 0.049105 | 0.050803 | | | |
| 0.065595 | 0.067677 | | | |
| 0.042018 | 0.040018 | | | |
| 0.025514 | 0.026434 | | | |
| 0.055946 | 0.056972 | | | |
| 0.035159 | 0.034196 | | | |
| 0.06202 | 0.058812 | | | |
| 0.10612 | 0.10413 | | | |
| 0.10053 | 0.10378 | | | |
| 0.072119 | 0.076815 | | | |
| 0.070119 | 0.073859 | | | |
| 0.046561 | 0.049805 | | | |
| 0.054096 | 0.054092 | | | |
| 0.063805 | 0.063563 | | | |

**Figure 4.9. Sample Extracted Training Features for ZCR**

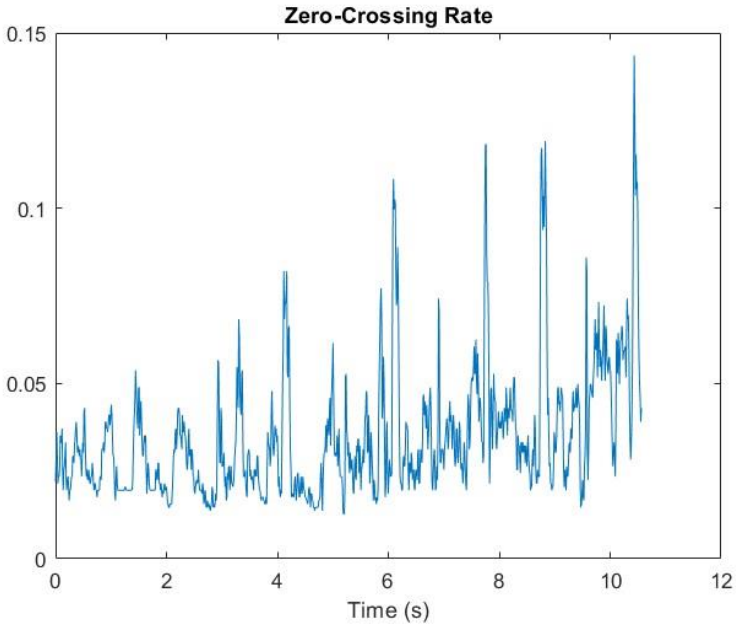



**(a) Original**

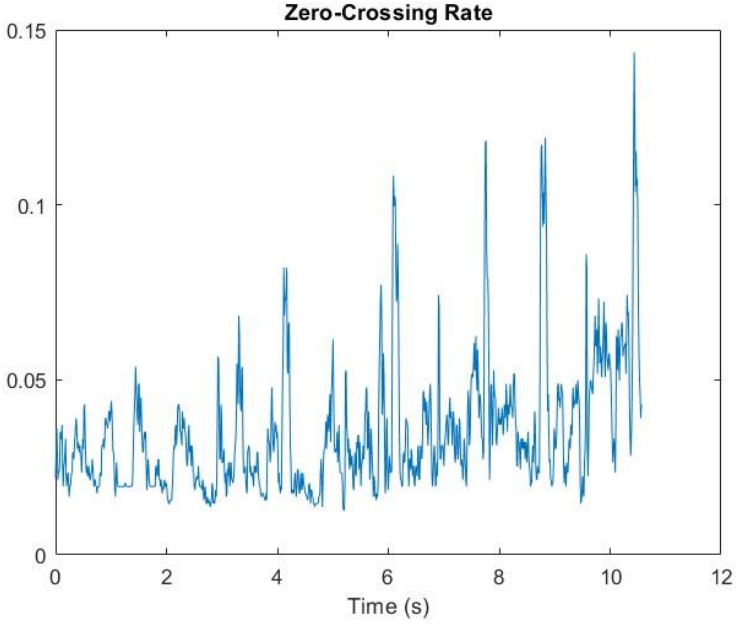



**(b) Stego**

**Figure 4.10. ZCR for Rock Song**

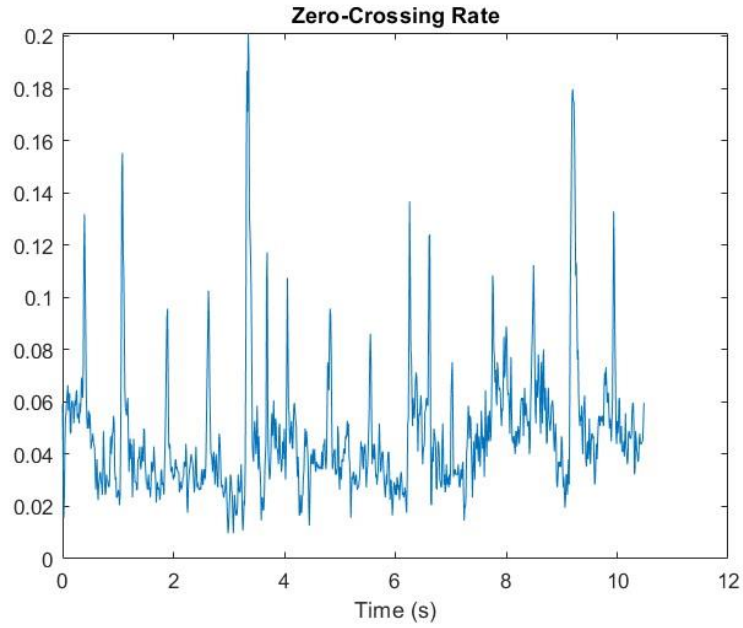



**(a) Original**

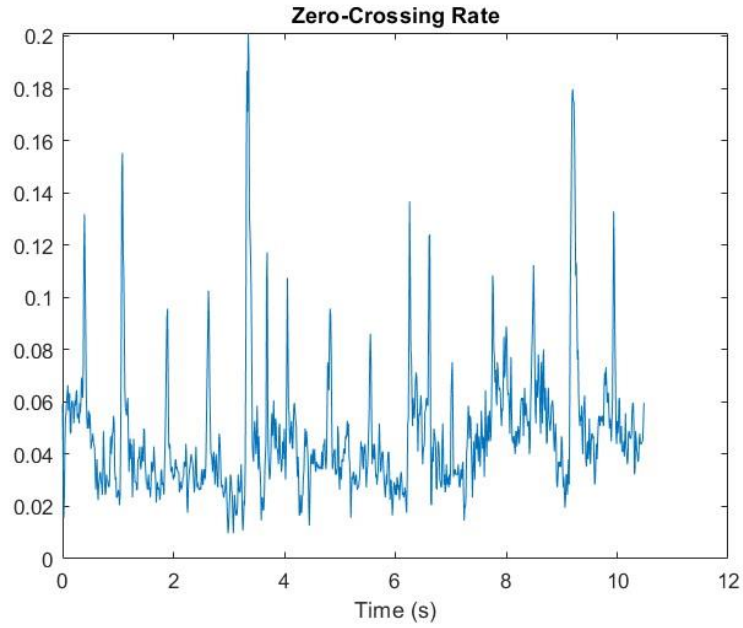



**(b) Stego**

**Figure 4.11. ZCR for Pop Song**




**(a) Original**




**(b) Stego**

**Figure 4.12. ZCR for Country Song**

## 4.3. Dummy Coding

This coding is developed by Cohen and Cohen in 1983, which is the simplest coding structure that supposed to examine group mean differences. A dummy variable is a numerical variable for representing group behavior.

Steganography detection system assumes that the independent features are numerical data. These are converted to the categorical variable for reducing the effect of redundancy within the extracted features of audio signals. Table 4.2 illustrates the converting extracted features to categorical based on the appropriate interval. Figure 4.13 shows the sample dummy categorical variables of extracted features.

**Table 4.2. Dummy Coding for Sample Audio Features**

| MFCC | MFCC (Coding) | ZCR | ZCR (Coding) | STE | STE (Coding) |
|---|---|---|---|---|---|
| -1.33204 | 1 | 0.053951 | 3 | 0.000671 | 4 |
| -1.49417 | 1 | 0.05396 | 3 | 0.000389 | 2 |
| -1.64355 | 1 | 0.035514 | 2 | 0.000664 | 4 |
| -1.81382 | 1 | 0.036092 | 2 | 0.000247 | 2 |
| 1.573448 | 2 | 0.06811 | 4 | 0.000234 | 2 |
| 2.042204 | 3 | 0.066647 | 4 | 0.000178 | 1 |
| 2.279364 | 3 | 0.028541 | 2 | 0.000092 | 1 |
| 1.972043 | 2 | 0.026904 | 2 | 0.000102 | 1 |
| 1.668896 | 2 | 0.06135 | 4 | 0.000474 | 3 |
| 1.425379 | 2 | 0.0603 | 4 | 0.000139 | 1 |

For MFCC: -2 to 0 = 1

0 to 2 = 2

2 to 4 = 3

For ZCR: 0 to 0.02 = 1

0.02 to 0.04 = 2

0.04 to 0.06 = 3

0.06 to 0.08=4

For STE: 0 to 0.0002= 1

0.0002 to 0.0004 = 2

0.0004 to 0.0006 = 3

0.0006 to 0.0008=4

| MFCC | MFCC | MFCC | MFCC | MFCC | MFCC | MFCC | MFCC | MFCC | MFCC | MFCC | MFCC | MFCC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 2 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 1 |
| 4 | 4 | 3 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 |
| 4 | 4 | 3 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 3 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 |
| 3 | 4 | 3 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 |
| 3 | 4 | 3 | 1 | 2 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 |
| 3 | 4 | 3 | 1 | 1 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 1 |
| 3 | 4 | 3 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 |
| 3 | 4 | 3 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 1 | 1 |
| 3 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |
| 3 | 4 | 3 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 |
| 3 | 4 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 |
| 4 | 4 | 3 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 1 | 1 | 1 |
| 4 | 4 | 3 | 1 | 1 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 1 |
| 4 | 4 | 3 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 2 | 1 | 1 |

**Figure 4.13. Sample Dummy Coding Result for Extracted Features**

## 4.4. Experimental Results of Proposed System

This section describes the graphical user interface of audio steganalysis system. The proposed audio steganography detection system is implemented by using MATLAB programming. The following Figure 4.14 shows the step by step process for detection of audio steganography.

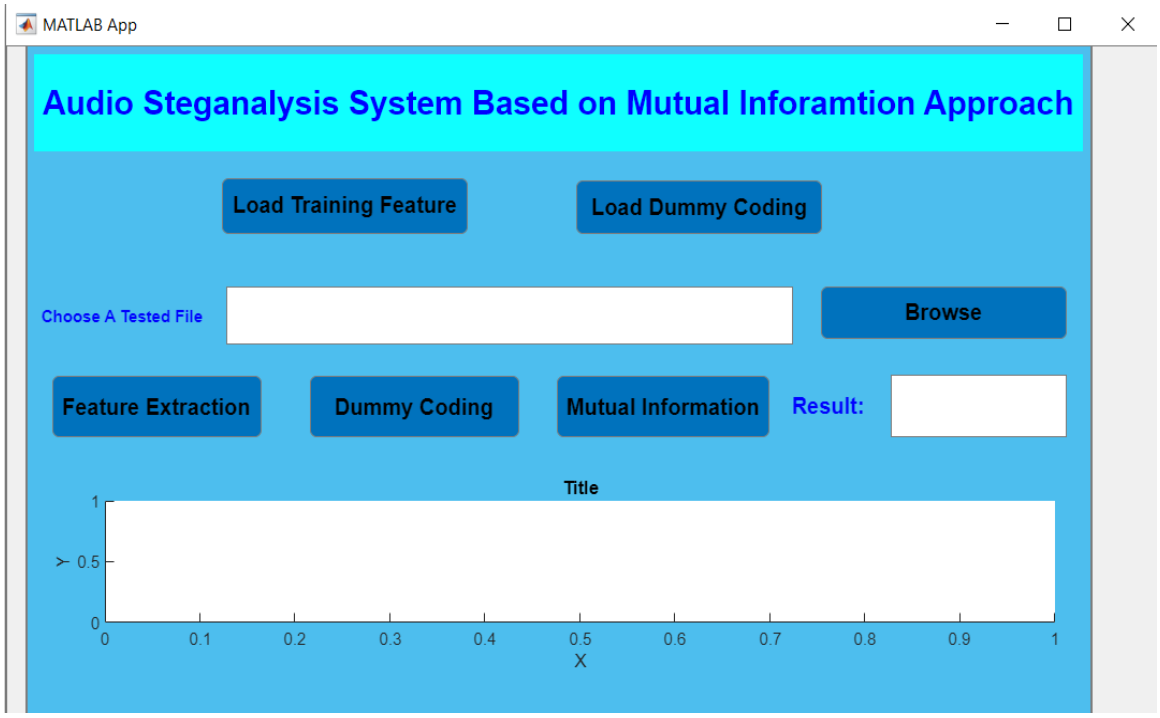



**Figure 4.14. User Interface for Proposed System**

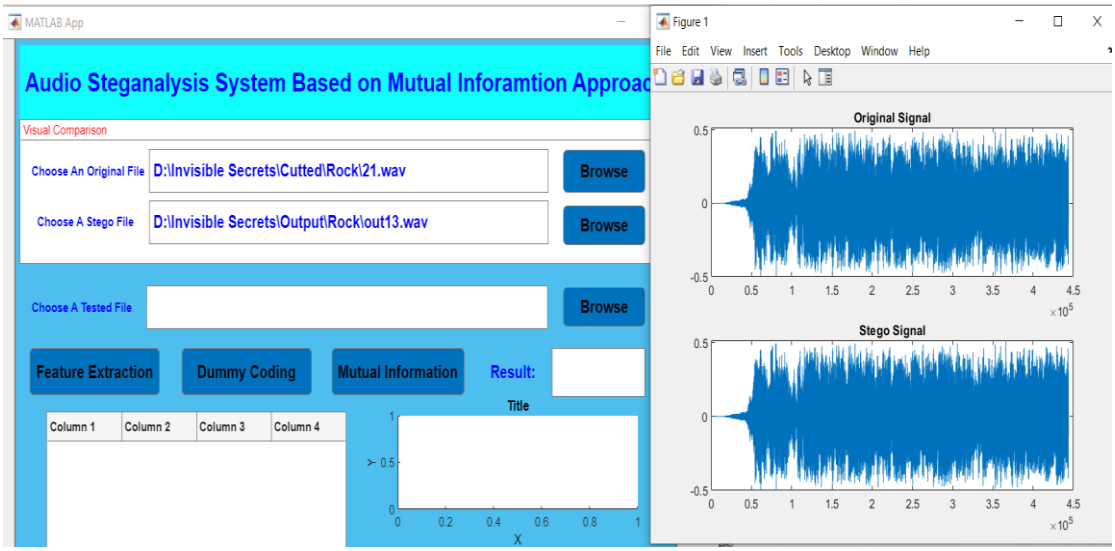



**Figure 4.15. Frequency Sample Analysis for Difference Audio Signal**

Figure 4.15 shows the comparison of original audio file with stego file which has hidden message. There is no difference in sampling frequency which cannot be identified with visual analysis. Figure 4.16 describe the loading of training features from the original audio and audio file with hidden message.

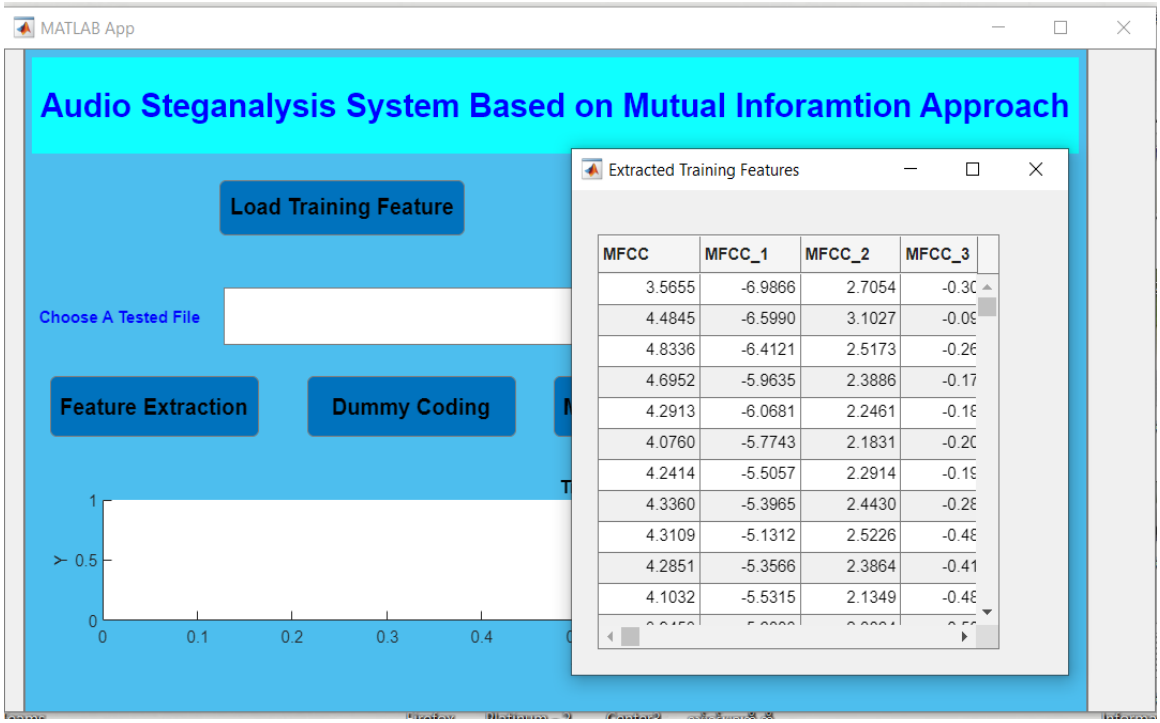



**Figure 4.16. Loading the Extracted Training Features**

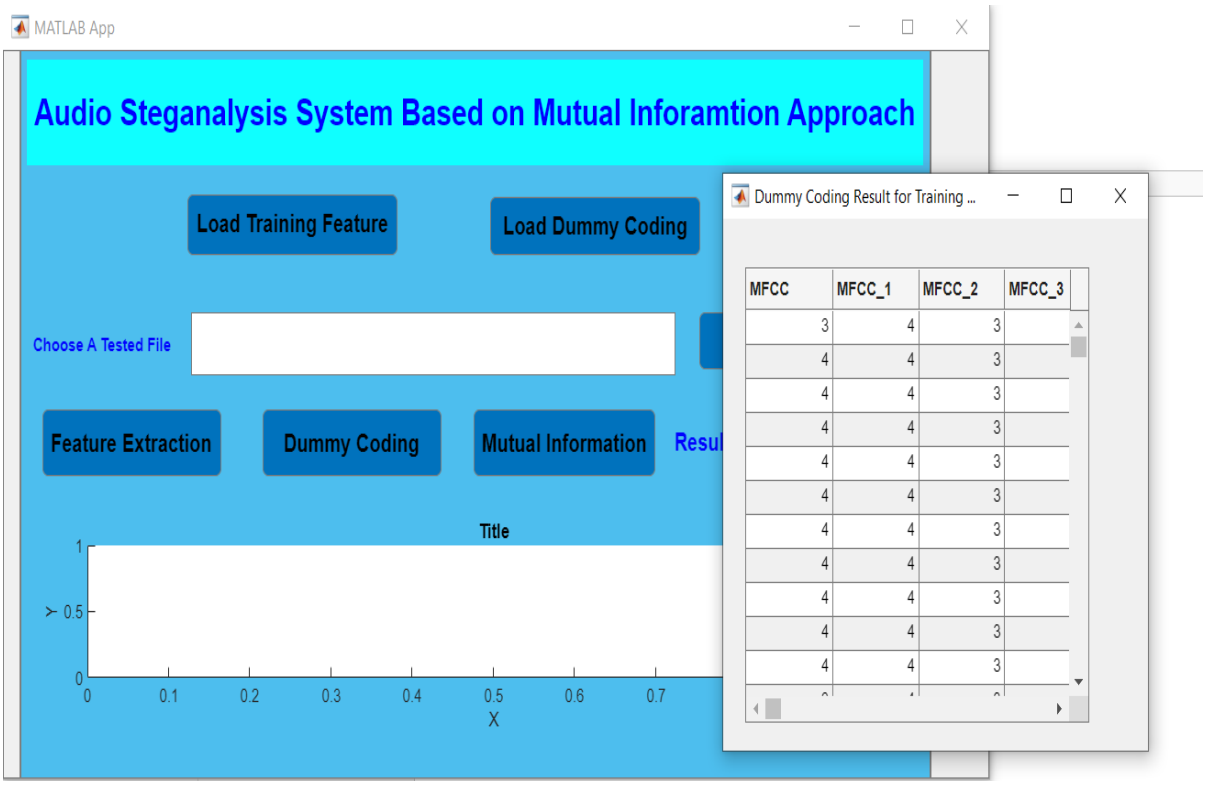



**Figure 4.17. Loading the Dummy Categorical Value for Training Feautres**

Figure 4.17 shows the results of dummy categorical value from the extracted audio features. Categorical values are classified based on the histogram bin ratio and which are defined for getting the highest prediction results for steganography detection.

Figure 4.18 describes how to choose the audio file which will be tested with proposed steganography detection system based on mutual information.

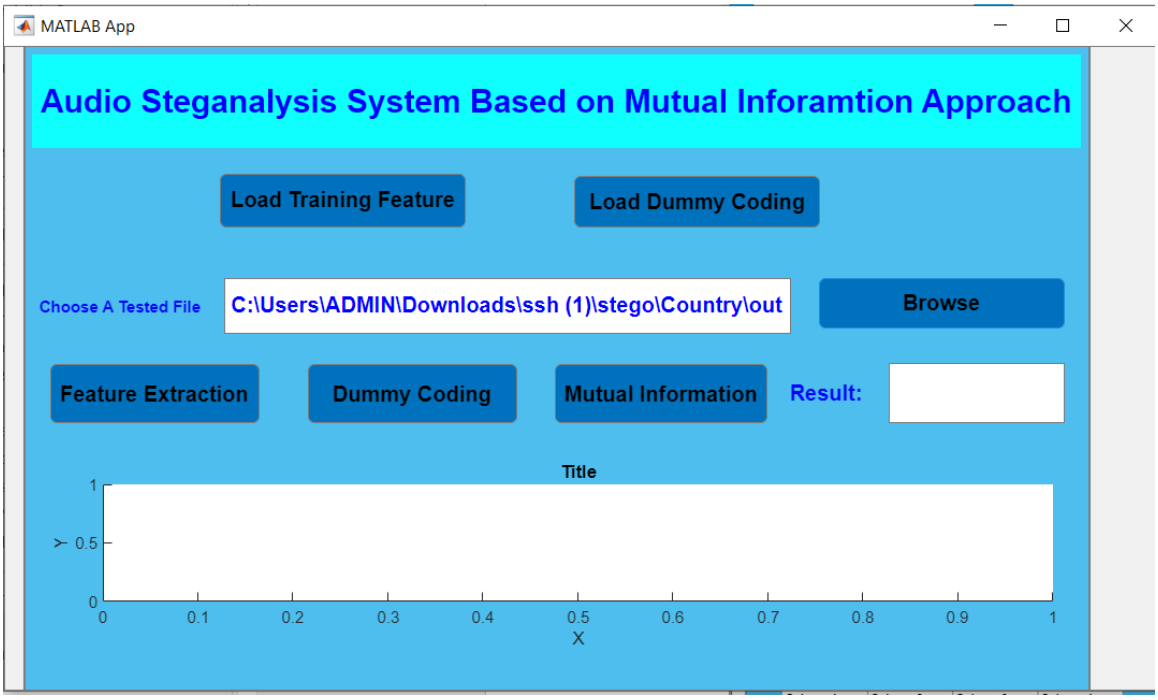
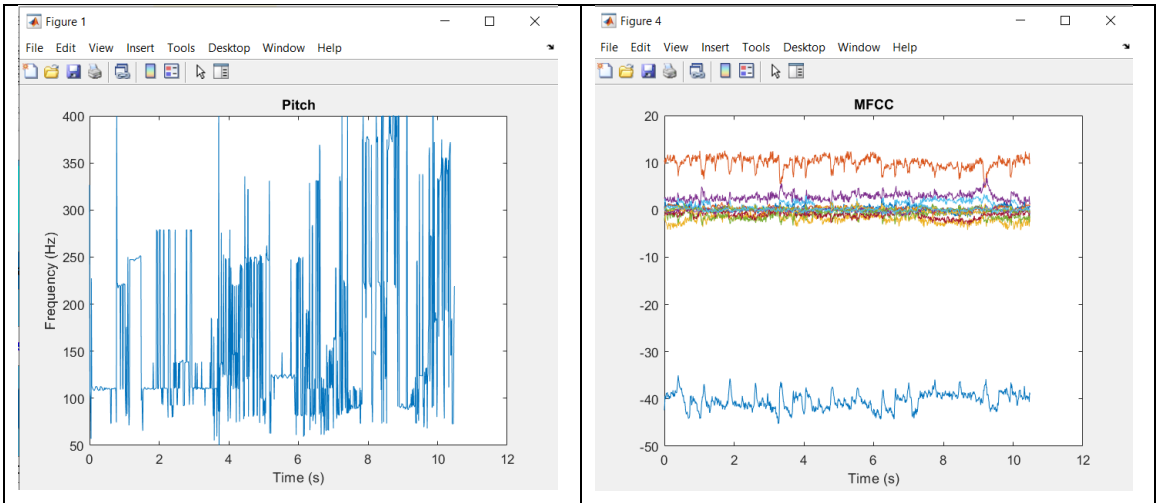
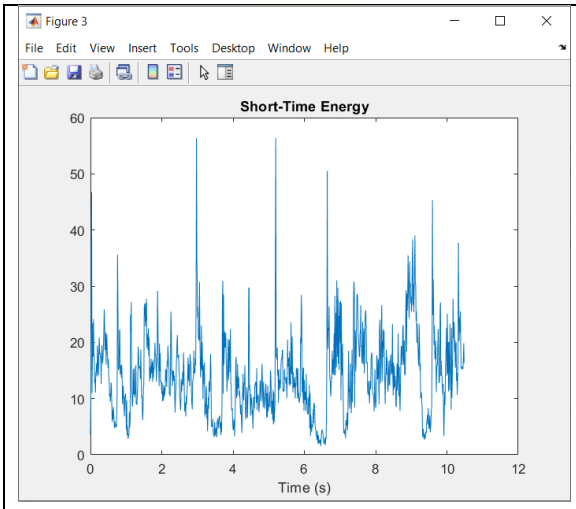


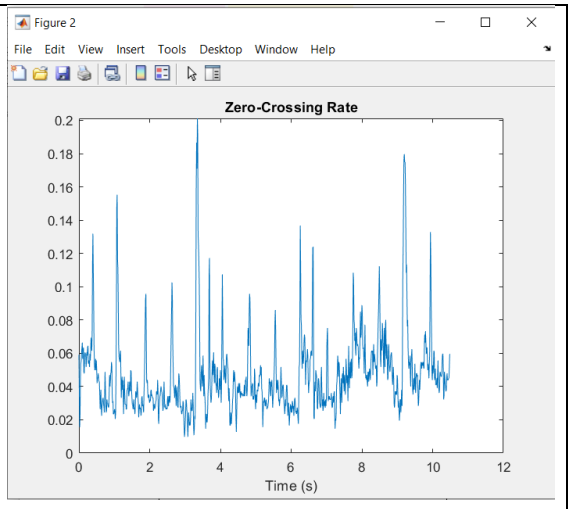

**Figure 4.18. Choose a File for Testing**




| (a) Sampling frequency of testing audio file | (b) Plotting the MFCC features of testing audio file |
|---|---|




**(c ) Plotting STE features of testing audio file**




**(d ) Plotting the ZCR features of testing audio file**

**Figure 4.19. Feature Extraction Results of Testing Audio File**

Figure 4.19 describes the extracted audio features of tested audio file which can be seen. There is some difference for original and stego audio file with MFCC, STE and ZCR features. The following Figure shows the process for converting the extracted features to categorical variable based on dummy coding. The final step, the proposed system used Mutual Information (MI) for finding the relationship between the tested audio features with training audio features from dataset. There is a significance relationship between these features which are classified as stego or original signal based on Mutual Information (MI) results. The results of converting dummy categorical variable from extracted audio features of testing audio file which are shown in Figure 4.20.

Figure 4.21 shows the final result of audio steganography detection system. In this step, evaluate the mutual information value between the testing and training features. Moreover, find the maximum value in the results of mutual information. Maximum value refer to the Stego categories of training features. There is a relationship between testing file with stego features. Therefore, the result concluded that this tested file was “Stego” or audio file with hidden message.

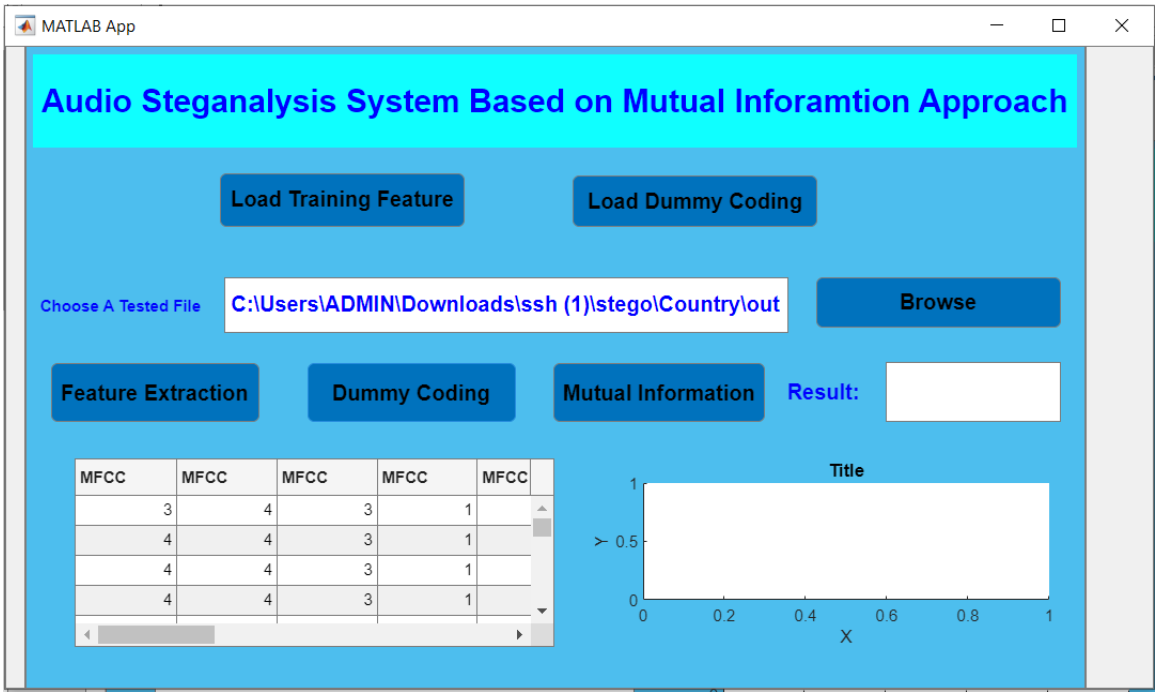



**Figure 4.20. Convert Categorical Variable for Extracted Features**

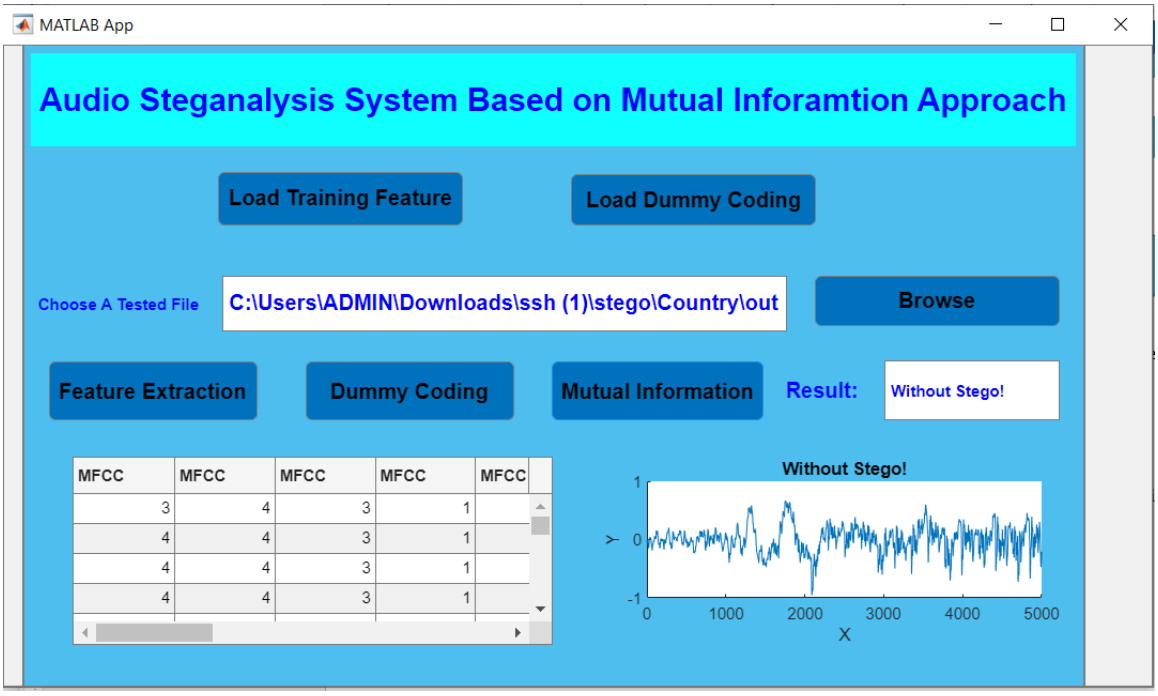



**Figure 4.21. Mutual Information Analysis for Testing Audio File**

## 4.5. Performance Analysis for Proposed System

Figure 4.22 describes the detection accuracy under difference number of bits. In the stenographic tools, hidden messages are embedded by different encoding scheme so this causes to vary detection accuracy over different bit numbers. Maximum number of bits are embedded in audio signal that is more accurate than signal with minimum number of bit.

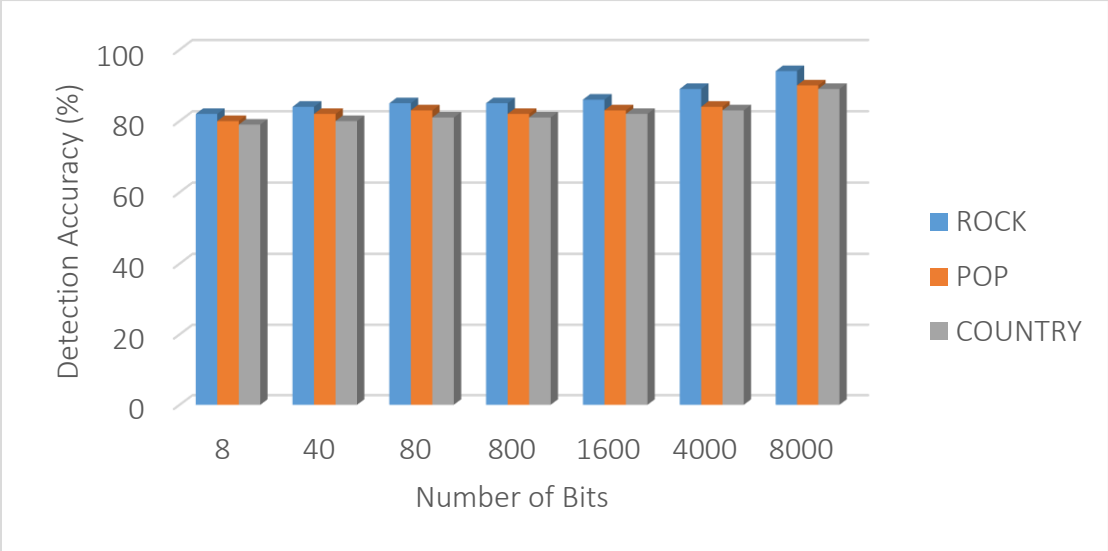
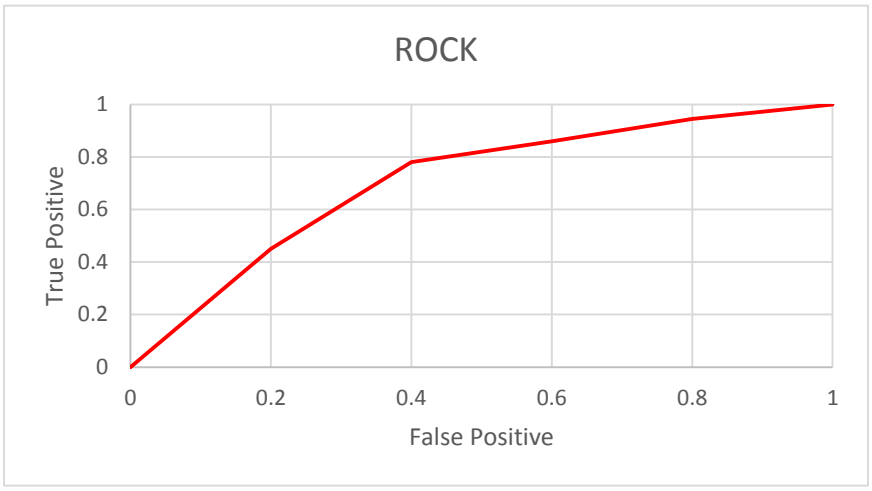
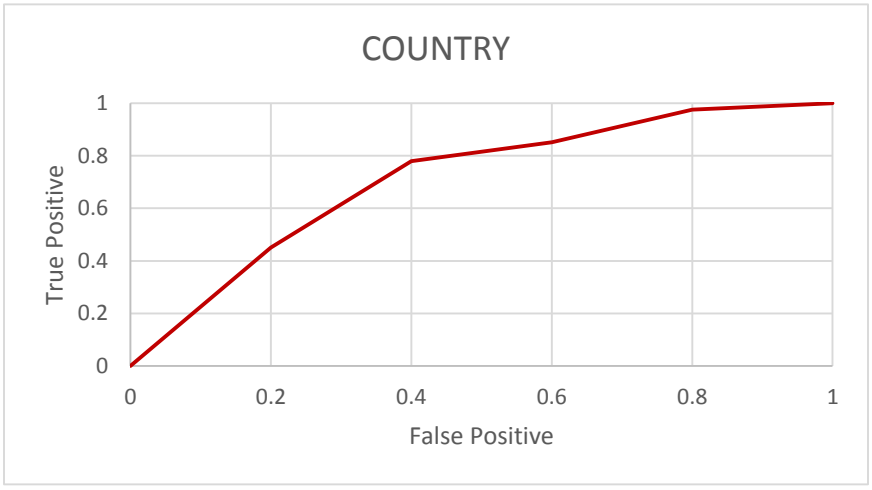



**Figure 4.22. Detection Accuracy with Difference Number of Bits**

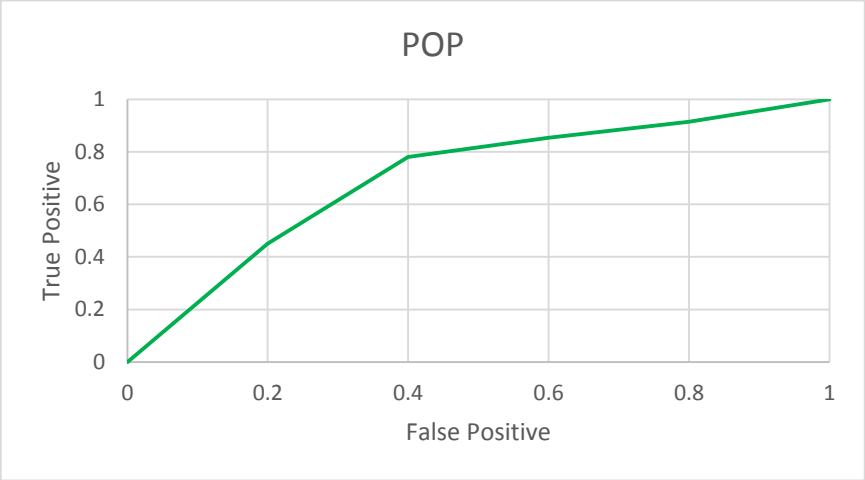
In machine learning, True Positive Rate (TPR) is used to calculate sensitivity and recall and False Positive Rate (FPR) can be evaluated as one minus the more well-known specificity


ROCK
True Positive
False Positive
0
0.2
0.4
0.6
0.8
1

(a)


COUNTRY
True Positive
False Positive
0
0.2
0.4
0.6
0.8
1

(b)




(c)

**Figure 4.23. ROC Curve over Different Types of Audio Signal**

In these experiments, ROC curve has been used to verify the effectiveness of the proposed method. Figure (4.23) gives the ROC curves as the detection threshold is varied. It can be seen that maximum amount of bits are embedded in audio signal in Rock song, true positive rate is nearly one and false positive rate is decreased.

CHAPTER 5

# CONCLUSION

This Audio steganalysis system based on Mutual Information approach has been proposed. The proposed system applied to extract time domain and frequency domain feature of audio signal. There are Mel Frequency Cepstral Coefficient (MFCCs), Short Time Energy (STE) and Zero Crossing Rate (ZCR). This system showed the efficacy and efficiency in applying audio feature extraction based Mutual Information (MI) approach for performing steganography detection. The steganography is readily detected by Mutual Information (MI) approach and the performance of the proposed system can be evaluated in different bit numbers with different audio files. Mutual information is a lot like correlation in that it measures a relationship between two quantities.

One of the advantages of applying mutual information approach is that it can detect any kind of relationship, while correlation only detects linear relationships. A theoretical strong advantage of Mutual Information (MI) analysis is that it does not require to specify a functional form of dependency such as correlation or Cramer indicator. On the contrary, due to probabilistic properties of Mutual Information (MI), this mutual information are very efficient to detect non-linear data. The proposed audio steganography detection system used different types of audio features which are continuous and non-linear signal.

The experimental result showed that the proposed system is useful for steganalysis of Invisible Secret tools and to improve the accuracy of the detectors. Mutual Information can be employed very effectively and systematically in analyzing the categorical data. Dummy variables are used to make the resulting application easier to implement and distinguish in steganography detection system.

As the limitation, the proposed system can only detect original audio signal or audio signal with hidden message. For future work, the place of hidden message in audio signal can be identified by using the combination of different types of audio features with machine learning algorithm.

Mutual information is a powerful method for identifying the strength of the relationships between variable of different natures without any constraint on the distribution laws. It is a very useful way to select the most pertinent variables that may be included in predictive models.

# REFERENCES


[1] A.Hyvärinen and E.Oja (2000) Independent Component Analysis: Algorithms and Applications. Neural Networks Research Centre Helsinki University of Technology Neural Networks, 13(4-5):411-430.

[2] A.H. Sung Novel Stream Mining for Audio Steganalysis. ACM 978-1-60558-608-3/09/10 19th International Conference on Pattern Recognition.

[3] C. Kraetzer and J. Dittmann, (2007) Mel-Cepstrum based steganalysis for voip stegography. Proc SPIE, vol 6505, p. 650505.

[4] C. Platt (2004) UnderMP3Cover, http:// sourceforge.net/ projects/ump3c.

[5] C. Kraetzer, J. Dittmann. (2008) Pros and Cons of Mel-cepstrum Based Audio Steganalysis Using SVM Classification. Lecture Notes in Computer Science, vol. 4567, pp. 359-377.

[6] D.Yan, R.Wang Steganalysis for MP3Stego using differential statistics of quantization step. Digital Signal Processing http:// dx.doi.org/10.1016/j.dsp.2013.02.013.

[7] F. A. P. Petitcolas, MP3Stego, http://www.cl.cam.ac.uk/fapp2/steganography/mp3stego.

[8] Hyv¨ arinen, A., Oja, E.: (1997) A fast fixed-point algorithm for independent component analysis. Neural Computation 9 1483–1492.

[9] H. Ozer, B. Sankur, N. Memon, I. Avcibas, (2006) Detection of audio covert channels using statisticalfootprints of hidden messages. Digital Signal Processing, 389 – 401.

[10] Lee, T.W (1998) Independent Component Analysis: Theory and Applications. Kluwer Academic Publishers.

[11] M. Qiao, A. Sung, Q. Liu, (2009) Steganalysis of MP3Stego. In: Proceedings of International Joint Conference on Neural Networks, Atlanta, Georgia, pp. 2566 – 2571.

[12] M. Qiao, A. Sung, Q. Liu, (2009) Feature mining and intelligent computing for MP3 steganalysis. In Proceedings of International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing, Shanghai, China, pp. 627 – 630.

[13] N. F. Johnson, S. Jajodia. (1988) Exploring steganography: Seeing the unseen. IEEE Computer 31 (2), 26 – 34.

[14] Noch,Stego-Lame, (2002) http:// sourceforge.net/ projects/ stego-lame.

[15] Q.Ding X. Ping (2010) Steganalysis of Analysis-by-synthesis Compressed Speech. 978-0-7695-4258-4/10 $26.00 © 2010 IEEE DOI 10.1109/MINES. 148.

[16] R. Chandramouli, M. Kharrazi, N. Memon. (2004) Image steganography and Steganalysis: Concepts and practices. International workshop on Digital Watermarking, 204 – 211.

[17] R, Krenn, (2004) Steganography and Steganalysis, An article, January.

[18] S.S. Agaian, (2004) Two algorithm in digital audio steganography Using quantized frequency domain embedding and Reversible integer transforms, Non-Linear Signal Processing Lab, University of Texas at San Antonio, Texas.

[19] S.X. Guang, (December,2010) A steganalysis method based on distribution of first letters of words. In procedding of the 2010 international conference on intelligent information hiding and multimedia signal processing, pp 369-372.

[20] S. Mahajan, (October 2012) A review of methods and approach for secure steganography. International journal of advanced research computer science and software engineering. Vol 2, no.10. pp 67-70.

[21] Sajedi, (May 1998) On the limits of steganography. IEEE journal of selected areas in communication, vol 16, no 4, pp 474-481.

[22] Y.Huang, H.Song, Detecting MP3Stego and Estimating the Hidden Size. In Proceedings of the 20th International Joint Conference in Artificial Intelligence (IJCAI). 2808–2813.

[23] Y.Tint, K.T.Mya (2013) Steganalysis for MP3Stego Using Independent Component Analysis. International Conference of Information and Communication Technologies for Education.

# LIST OF PUBLICATIONS

[1] Su Su Hlaing, Yawai Tint, Audio Steganalysis System based on Mutual Information Approach, Parallel and Soft Computing, UCSY, 2022.