Chapter 1

Introduction

1.1 Motivation

Advances in digital technology have led to widespread use of digital communication in various areas including government, legal, banking and military services. This in turn has increased the reproduction and re-transmission of multimedia data through both legal and illegal channels. However, the illegal usage of digital media causes a serious threat to the content owner’s authority or proprietary right. Thus, today’s information driven society places utmost importance on authenticating the information that is sent across various communication channels. In the case of digital audio communication schemes these disputes may be the denial of authorship of the speech signal, denial of sending or receiving the signal, denial of time of occurrence etc. Incorporating non-repudiation services in this context guarantees the occurrence of a particular event, the time of occurrence as well as the parties and the corresponding information associated with the event.

Typically, a non-repudiation service should produce cryptographic evi-
evidence that guarantee dispute resolution. In other terms, the service should hold relevant information that can achieve the goals against denying their presence or participation. Development of a non-repudiation service should have a service request, in the sense that, the parties involved should agree to utilize the service as well as to generate necessary evidence to support their presence. Evidence of this scheme should be transferred to the other party for the purpose of verification and storage. Separate evidence should be available for the originator as well as the recipient by considering the fact that, any one will not gain any extra benefit from this service and to ensure that the concept of fairness is applied. Timeliness and confidentiality are the other features of a non-repudiation service.

Currently most audio watermarking methods available are inclined towards copyright protection and copy protection. This is the key motive for the notion to develop a speaker verification scheme that guarantees non-repudiation services and the thesis is its outcome. Developing a non-repudiating voice authentication scheme is a challenging task in the context of audio watermarking. Our aim is to suggest a digital audio watermarking scheme that ensures authorized and legal use of digital communication, copyright protection, copy protection etc. that helps to prevent such disputes. Audio watermarking is the term coined to represent the insertion of a signal, image or text of known information in an audio signal in an imperceptible form. The embedded watermark should be robust to any signal manipulations and can be unambiguously retrieved at the other end.

1.2 Problem Statement

Evolution in digital technology led to widespread use of digital communication and illegal usage of digital media causes a serious threat to the content
owner’s authority or proprietary right. Recent copyright infringements in
digital communication make us believe that the stronger analytical tools
and methods need to be researched on.

In order to combat this malicious usage of digital audio communication
we need to:

- Understand the existing audio watermarking schemes especially that
  are proposed towards Intellectual Property Rights (IPR);
- Understand some of the best practices in existing watermarking schemes;
- Identify a differentiator for the new schemes which in turn results in
developing signal dependent watermarks;
- Classify the key acoustic characteristics that facilitate to uniquely
  identify the speaker by creating dedicated FeatureMarks;

1.3 Objectives of the Proposed Study

- Extract the key signal contingent features associated with the acoustic
  signals;
- Identify appropriate features that enable us to identify the speaker
  by employing artificial neural networks (ANN), k-nearest neighbors
  (k-NN) and support vector machine (SVM) classifiers;
- Craft the signal reliant watermark using the appropriate extracted
  features;
- Embed the new watermark or FeatureMark using Fast Fourier Trans-
  form (FFT);
Embed the new watermark or FeatureMark using Fast Walsh-Hadamard Transform (FWHT); 

- Evaluation of the proposed schemes in terms of imperceptibility, robustness and capacity; 

- Demonstrate speaker authentication as well as non-repudiation competency of the scheme. 

1.4 Scope of the Work

The work introduces three novel but diverse voice signal authentication schemes that assure non repudiation by utilizing the key acoustic signal features towards the preparation of the watermark. As part of this research ANN, k-NN and SVM classifiers are employed in tagging the appropriate acoustic features in the new FeatureMark. Acoustic characteristics such as Mel-frequency cepstral coefficients (MFCC), spectral roll-off, spectral flux, spectral centroid, zero-cross rate, energy entropy and short-time energy are vital to this research. This research also illustrate the watermark embedding algorithms which is central to this research. Experiments to determine the behaviors of the proposed schemes in terms of imperceptibility, robustness and capacity is also a component of this work. Main idea behind this work; realization of a non-repudiation service; is achieved in such a way that the speaker in the communicating group cannot subsequently deny their participation in the communication due to the signal-dependent dynamic watermark. 

- Scope 
  
  - Determination of apt audio features by conducting speaker recognition using different classifiers
1.5 System Framework

- Component-based FeatureMarking system with FFT, Barker code and data matrix
- Component-based FeatureMarking system with FWHT, Walsh code and quick response (QR) Code
- Evaluation of the FeatureMark strength
  1. Transparency Tests
  2. Robustness Tests
  3. Capacity Tests

The watermark technique introduced allows tracking the spread of illicit copies but does not do anything to limit the number of copies allowed or control its dissemination through computer networks or other digital media such as compact disks. This research doesn’t gaze on the impact of Human Language or impact of mimic sounds on the proposed watermark.

1.5 System Framework

The research involves iteration of steps from collection of acoustic samples from diverse speakers to the detection of embedded FeatureMark. The initial step is to collect different speech signals from people. Next step involves pre-processing of speech signal using the framing and windowing methods. Pre-processed signals are given into the feature extraction module. Once the features are extracted and stored in database, classification module starts functioning to determine some of the apt features that can identify speakers uniquely or in combination with other features.

The actual watermarking algorithm starts its functions only at this step and it needs the watermark developed using the extracted signal features as input. In order to prepare the signal dependent watermark, also termed
as FeatureMark in the suggested schemes; online data-code generators are employed. In some cases, a synchronization code is also generated that could guarantee robustness of the watermarking schemes. FeatureMark embedding is performed by either transforming the signal using FFT or FWHT transforms. Embedded signal is inverse transformed and send to the other end.

Overview of the proposed schemes is presented in the following figures - figure 1.1 and figure 1.2.

![Watermark embedding scheme](image)

Figure 1.1: Watermark embedding scheme

At the receiving end, presence of watermark is confirmed by performing proper signal transforms. Once the watermark has been detected, it should be extracted to confirm the authenticity of the signal. This guarantees the proof of ownership as the watermark itself holds information about the speakers. The watermark can be enhanced to hold the location, date and
1.6 Thesis Contributions

This dissertation contributes to the area of pure experimental computer science and introduces novel thinking and techniques to the fields of audio watermarking. The primary objective of this dissertation is to test the hypothesis that:

- digital communication require and should benefit from novel non-repudiation service designed to exploit the acoustic characteristics of the parties involved in the communication.

It should be eminent that it is not possible to formally prove the rightness or falsehood of this hypothesis. Instead, this dissertation is limited to
providing strong evidence for or against its validity. It does so by introducing three new FeatureMarking techniques and revelation of its experimental results. Proposed schemes were able to showcase improvement in terms of imperceptibility, robustness and capacity.

Major contributions of this work involve suggestion of a model for the generation of signal dependent dynamic watermarks that assures authenticity. Through this research three different audio watermarking schemes are offered in which the first one is an acoustic authentication scheme with FFT and barcode as the watermark, second one is a varying audio watermarking system with FFT and data matrix code as the watermark and the final scheme works with FWHT and QR code as the watermark that supports non-repudiation services.

- Proposed a model for the generation of signal dependent dynamic watermarks that assures authenticity of the signal rather than using the regular static ones.

- A speech signal authentication scheme is proposed that works with FFT and uses barcode as the watermark.

- A varying audio watermarking scheme is suggested by employing data-matrix code as the watermark and uses FFT for the marking/un-marking schemes.

- Another method that supports non-repudiation services to a great extent are implemented with the help of FWHT and QR code as the watermark.

- An encryption scheme is suggested that adds one more layer of security to the signal dependent dynamic watermark.
1.6.1 List of Research Papers

As part of the research work various papers were presented and published in peer reviewed International Journals as well as in Conference proceedings. They are listed below:


1.7 Thesis Outline

The thesis is divided into nine chapters and a brief description of each chapter is given below.

Chapter 1 is a general introduction on the importance of watermarking especially audio watermarking. The chapter concludes the significance of the present work.

Chapter 2 is a documentation of the background study conducted to understand the audio signals, an overview of human auditory system, the frequency component analysis of signals and finally the concept of watermarking and its evaluation strategies.

Chapter 3 comprises of a brief description of the existing works that are proposed in the field of audio watermarking. The existing schemes
are mainly classified under three categories such as time domain based algorithms, transform domain based algorithms and hybrid algorithms.

Chapter 4 focuses on the collection of speech data and how pre-processing is done to improve the result. Short-term processing of the signal manipulates the sound inputs appropriately and helps in improving the results of analysis and synthesis. It also guarantees a better quality for the extracted watermark. Feature extraction is another important step described in this chapter where some of the computational characteristics of speech signals are mined for later investigation. Features are extracted using the program code in Matlab by employing the FFT on the time domain signals. Features selected for this study includes physical features such as Mel-frequency cepstral coefficients (MFCC), spectral roll-off, spectral flux, spectral centroid, zero-cross rate, short-time energy, energy entropy and fundamental frequency, that directly correspond to the computational characteristics of the signal and not related to the perceptual characteristics.

Chapter 5 is dealing with the identification of exact features (from the features that we have chosen) that helps in speaker authentication to a great extent. This is achieved by employing three main classifiers such as ANN, k-NN and SVM for individual feature sets as well as different combinations of feature sets. This speaker recognition module reveals that MFCCs itself can help in identifying the speakers participated in the communication system. Different combinations of signal features such as MFCCs, spectral roll-off, zero-cross rate, spectral flux as well as spectral centroid are opted towards the creation of its signal dependent watermark.

Chapter 6 includes the first scheme that we have proposed towards authenticating each member who has participated in the communication system. This scheme works in the transform domain of an audio signal by employing FFT towards embedding and detection schemes. The pre-
pared watermark is a data code and employs Arnold/Anti-Arnold transform in embedding/extracting schemes for scrambling/de-scrambling the watermark. This two-dimensional watermark is transformed into a one-dimensional sequence of 1s and 0s (binary digits) to embed into the audio signal. To evaluate the efficiency of this method, subjective listening tests were conducted which demonstrates the transparency criteria. Robustness tests confirmed the strength against common signal manipulations and de-synchronization attacks and finally the capacity of this scheme was evaluated.

Chapter 7 demonstrates the second scheme which is a variation on the previous method with utilization of a 13-bit Barker code as synchronization code and a data matrix code as watermark in the embedding module. Embedding a synchronization code helps to locate the position of watermark in the modified signal which in turn reduces computational time of the system. FeatureMark embedding and thus its detection is achieved by transforming the signal using FFT. In this scheme also, efficiency tests were conducted to evaluate the transparency, robustness and capacity characteristics.

Chapter 8 introduces the third scheme that works with FWHT. In this scheme, a 16-bit Walsh code is generated and employed as the synchronization code and QR code is treated as its FeatureMark. A variation of this scheme is also suggested which incorporates an encryption scheme in the development of signal dependent watermark. Efficiency of this scheme is also tested by employing subjective listening tests that confirm transparency characteristics. Robustness tests are conducted to find out how the system is robust against common signal manipulations and de-synchronization attacks. Then capacity test is performed to identify the capacity of the proposed watermarking scheme.
Chapter 9 is summary of the work, where important conclusions such as the use of voice signal features, its classification and FeatureMarking towards the development of a secure, robust, voice authentication scheme that helps in guaranteeing the non-repudiation services are highlighted. A comparative study with the existing watermarking schemes is also presented. Towards the end of this chapter the future scope of the proposed works are given.

List of notations, abbreviations, publications, references and index are given at the end of this book.

1.8 Summary

The introductory chapter gives an idea about the work that we have done and the thesis contributions. With the existing audio watermarking algorithms we can guarantee copyright protection, copy protection and ownership to a great extent. But any of these schemes does not employ a signal dependent, dynamic watermark and this is the advantage that can be availed in the suggested schemes. And moreover embedding this FeatureMark helps in guaranteeing the ownership as well as non-repudiation service in a straight way.