International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No. (02), January 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

A Predictive Approach for Speaker Verification By Machine learning and MFCC

Shipra Gupta[1], Dr. Amit Sharma[2]

[1]M.Tech Scholar, [2]Professor,Department of Computer Science & Engineering,
Vedant College of Engineering & Technology, Bundi, Rajasthan (India)

ABSTRACT

Speech processing is one of the popular application area of digital signal processing. There are several parts of speech processing as speech recognition, speech synthesis, speaker recognition, speech coding etc. Speaker Recognition is identification of a person based on his voice-prints. There are two forms of speaker recognition: text dependent (constrained mode) and text independent (unconstrained mode). The project proffers an intelligent two factor speaker authentication system. The word intelligent is used to combine Artificial Intelligence with Feature Extraction Techniques (MFCC). Feature extraction is the key process for speaker recognition. In this work, Mel Frequency Cepstral Coefficients and Vector Quantization performs Speaker Identification based on an audio signal. Verification and authentication of speaker takes place using algorithm of artificial intelligence (AI System) namely decision tree.

Keywords: Mel Frequency Cepstral Coefficients (MFCC), Vector Quantization, Artificial intelligence.

I. INTRODUCTION

In the current era of massive information transfer and storage, security is a primary concern. In order to protect the sensitive information we usually rely on user-names and passwords or two factor authentication services for superior security. The vast number of applications demands a simpler and more efficient security protocol. Biometric security systems are convenient and efficient solutions to the scenario. The word biometric is derived from Greek word 'bios' and 'metrics' meaning 'life measurement'. Biometric security devices utilize a person's unique physical traits for user authentication and access control. The biometric techniques include both physiological and behavioral traits. The major physiological biometric techniques are face recognition, fingerprint recognition; palm print, hand geometry retina scan etc. Voice Recognition, typing rhythm and gait are some examples of behavioral biometric techniques. Biometric security systems provide immense benefits over conventional security systems. The benefits includes better accuracy, convenience, security and authenticity. The biometric systems eradicates the need to remember passwords and are more user friendly and ambiguous.

International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No. (02), January 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

II BIOMETRIC APPLY AND USE FOR SECURITY

Biometric security devices are presently applied in various sectors such as banking security, smart-phone security, network security, business and residential control, automobile security etc. However it should be noted that biometric security always acts as an added security layer as biometric systems are based on approximate match based on a threshold value. Thus even biometric devices are error prone. The recognition errors[1] are of two types as mentioned below:

1) False Accept Rate(FAR):

The False Accept Rate is a measure of chances of accepting a malicious user.

2) False Reject Rate(FRR) :

The False Reject Rate is a measure of chances of rejecting an authorized user.

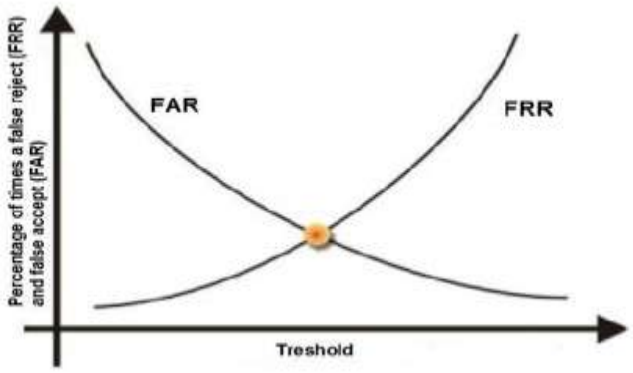

Fig. 1 Representation of False Accept Rate(FAR) and False Reject Rate(FRR)

The mentioned errors are complimentary which means that efforts to minimize one can lead to maximization of other. Therefore a balance between the two is achieved by applying a decision threshold. Comparison of different biometrics is summarized in Table 1. It also conveys that the voice biometrics has the highest potential for growth due to high user acceptance and easier implementation as voice biometrics do not require separate complex hardware.

Table 1. Comparison of biometrics

Characteristic	Fingerprints	Hand geometry	Retina	Iris	Face	Signature	Voice
Ease of Use	High	High	Low	Medium	Medium	High	High
Error incidence	Dryness, dirt, age	Hand injury, age	Glasses	Poor lighting	Lighting, age, glasses, hair	Changing signatures	Noise, colds, weather
Accuracy	High	High	Very high	Very high	High	High	High
Cost	*	*	*	*	*	*	*
User acceptance	Medium	Medium	Medium	Medium	Medium	Very high	High
Required security level	High	Medium	High	Very high	Medium	Medium	Medium
Long-term stability	High	Medium	High	High	Medium	Medium	Medium

* The large number of factors involved makes a simple cost comparison impractical.

International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No. (02), January 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

III. SPEAKER RECOGNITION SYSTEM

Speaker Recognition is identification of a person based on his voiceprints. There are two forms of speaker recognition: Text dependent (constrained mode): In a system using "text dependent" speech, the individual presents either a fixed password or prompted to recite a pre programmed phrase for instance, "Please say the numbers „31-52-93‟ .This mode of speaker recognition is applied with cooperative users. Text independent (unconstrained mode): A "text independent" system is a form of speaker recognition that has no advance knowledge of the speaker's phrasing and is used in situations where the individuals are either unwilling to cooperate or are kept unaware of being scrutinized. A text-independent system is more difficult to design than a text dependent system but offers more protection against fraudulent activities.

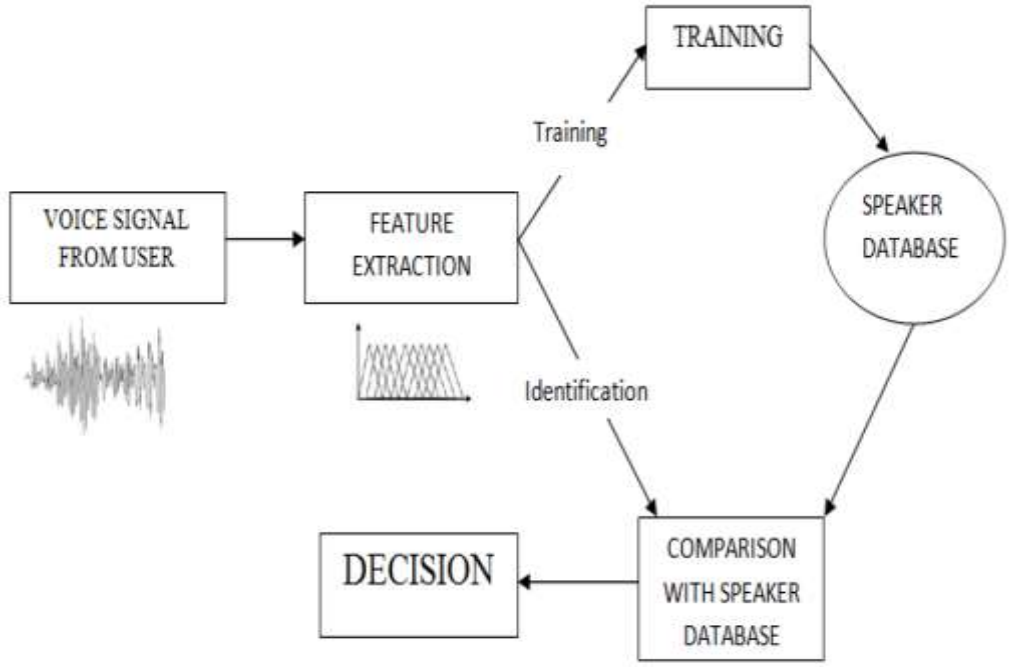

fig 2. Speaker Recognition System

The figure demonstrates the basic block diagram of a speaker recognition system. The process starts with user‟ s voice signal. The feature Extraction toolbox converts the voice signal into a sequence of vectors and these forms a unique set for each user. The extracted data is stored in the speaker database along with a user id during the training period. During the identification phase the extracted features are compared with the speaker database and if the feature match upto a threshold value then user is allowed access otherwise the access is denied.

IV. ALGORITHMS

The project proffers an intelligent two factor speaker authentication system by combining Artificial Intelligence with Feature Extraction Techniques (MFCC). In first stage voice authentication system that incorporates Mel

International Journal of Advance Research in Science and Engineering
Volume No.07, Special Issue No. (02), January 2018
www.ijarse.com
IJARSE
ISSN: 2319-8354

Frequency Cepstral Coefficients and Vector Quantization performs Speaker Identification and Verification based on an audio signal. In second stage further verification and authentication of speaker takes place using algorithm of artificial intelligence (AI System) namely decision tree.

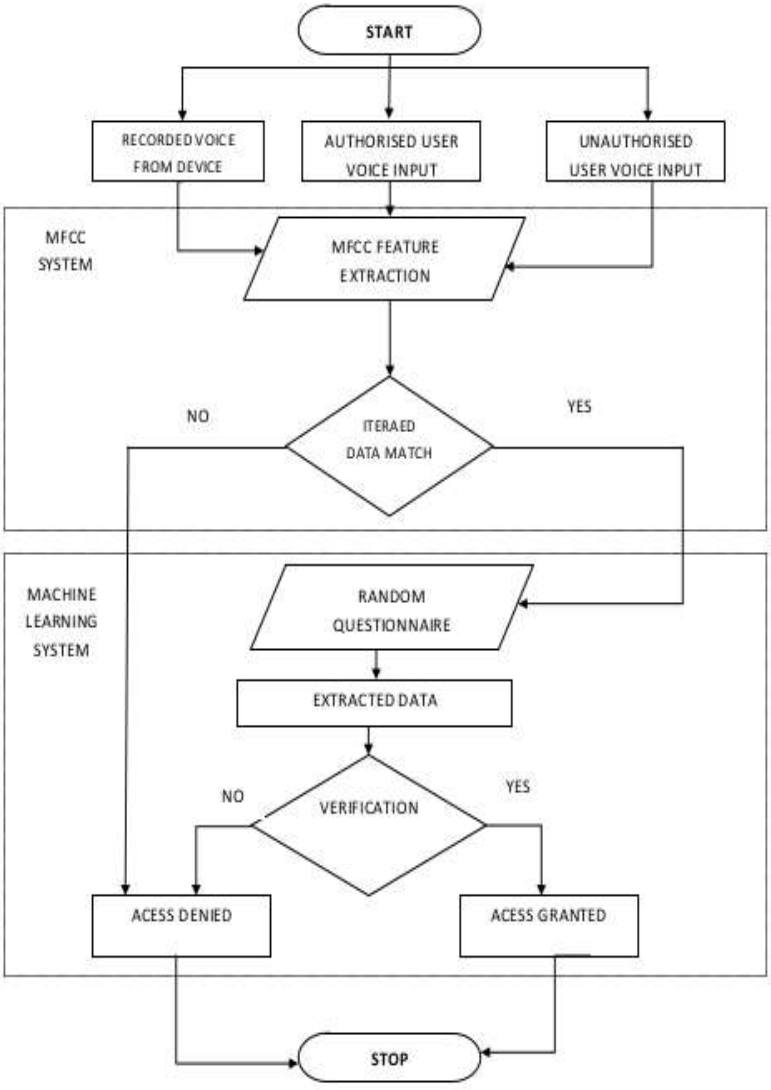

Fig 3 . Algorithms Implementation flow chart

The AI systems asks random questions related to user and verifies that the reply is relevant and came from the same user as determined by the voice authentication system i.e. AI system checks that the voice signal is not breached. The project proposes an improved as well as intelligent system for preventing privacy and security breaches in speaker verification systems.

V. MOTIVATION

As a sole form of security voice authentication system poses potential risks and can lead to inconvenience and thefts as voice prints may get stolen from the system. Voice prints can be easily compromised as we expose our voice in recordings, voice messages, phone calls etc. For instance, if an application requires voice of the user to

unlock it, an attacker with audio recording of the user's voice can easily unlock it and cause potential damage to privacy and security of the user. Advances in technology, specifically those that automate speech synthesis such as voice morphing, allow an attacker to build a very close model of a victim's voice from a limited number of samples. Voice morphing can be used to transform the attacker's voice to speak any arbitrary message in the victim's voice . Federal Communications Commission (FCC) and Better Business Bureau (BBB) issued warning about spam callers who records user voice and use the same to authorize fraudulent utility charges or credit activities. In an experiment the HSBC's voice recognition system was cracked by a journalist using his twin brother. Thus voice recognition as a standalone security solution is not sufficient and must be used as a hybrid technology.

VI. CONCLUSION

The paper is a comprehensive study of currently available algorithms for a speaker verification system. Our main observation is that the best combination varies depending on the input-number of speakers and the combination that performs best for fewer samples does not always give the best performance with larger number of samples In conclusion, the best combination of algorithm must be chosen depending on the end requirement. Throughout the course of study a random dataset of male and female voices were used to train the network and an enhance study can be done by using either male voices or female voices or a combination of male and female voices in a definite proportion.

REFERENCES

[1] Raghvendra Priyam, Rashmi Kumari and Dr. Prof Videh Kishori Thakur, "Artificial Intelligence Applications for Speech Recognition " in Conference on Advances in Communication and Control Systems 2013 (CAC2S 2013).

[2] Trevor R. Agus, Simon J. Thorpe, Clara Suied and Daniel Pressnitzer, "Characteristics of human voice processing" in Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium.

[3] Shaik Shafee and Dr.B.Anuradha,"Speaker Identification and Spoken word Recognition In Noisy Background using Artificial Neural Networks" in International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) – 2016.

[4] D.M.Chandwadkar and Dr. M.S.Sutaone, "Role of Features and Classifiers on Accuracy of Identification of Musical Instruments" inComputational Intelligence and Signal Processing (CISP), 2012 2nd National Conference.

[5] Chao Wang, Ruifei Zhu, Hongguang Jia, Qun Wei, Huhai Jiang, Tianyi Zhang and Linyao Yu, "Design of Speech Recognition System " in Third International Conference on Information Science and Technology March 23-25, 2013; Yangzhou, Jiangsu, China".

[6] Niladri Sekhar Dey, Ramakanta Mohanty andK. LChugh ,"Speech and Speaker Recognition System using Artificial Neural Networks and Hidden Markov Model" in International Conference on Communication Systems and Network Technologies 2012.

[7] Md. Afzal Hossan, Sheeraz Memon and Mark A Gregory,"A Novel Approach for MFCC Feature Extraction"in Signal Processing and Communication Systems (ICSPCS), 2010 4th International Conference.

[8] Qingyang Hong, Caihong Zhang, Xiaoyang Chen and Yan Chen, "Embedded Speech Recognition System for Intelligent Robot" in Mechatronics and Machine Vision in Practice, 2007. M2VIP 2007. 14th International Conference.

[9] Yuan Xue Luping Wang, Linxuan Li Zhiqi Liu and Jialin Liu, "Matlab-based Intelligent Voiceprint Recognition System" in Sixth International Conference on Instrumentation & Measurement, Computer, Communication and Control 2016.

[10] Rozeha A. Rashid, Nur Hija Mahalin, Mohd Adib Sarijari, Ahmad Aizuddin and Abdul Aziz, "Security System Using Biometric Technology: Design and Implementation of Voice Recognition System (VRS)" in Proceedings of the International Conference on Computer and Communication Engineering 2008.

[11] D. A. Reynolds and L. P. Heck, "INTEGRATION OF SPEAKER AND SPEECH RECOGNITION SYSTEMS" in Acoustics, Speech, and Signal Processing, 1991. ICASSP-91,1991 International Conference .