D.Ambika, V.Radha / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 5, September- October 2012, pp.1044-1049

Secure Speech Communication – A Review

D.Ambika*, V.Radha**

*(Department of Computer Science, Avinashilingam Institute for Home Science and Higher education for women, Coimbatore-43)

** (Department of Computer Science, Avinashilingam Institute for Home Science and Higher education for women, Coimbatore-43)

ABSTRACT

Secure speech communication has been of great importance in civil, commercial and military communication systems. As speech communication becomes widely used and even more vulnerable, the importance of providing a high level of security becomes a major issue. The main objective of this paper is to increase the security, and to remove the redundancy for speech communication system under the global context of secure communication. So it deals with the integrating of speech coding, with speaker authentication and strong encryption. This paper also gives an overview and techniques available in speech coding, speaker Identification, Encryption and Decryption. The primary objective of this paper is to summarize some of the well known methods used in various stages for secure speech communication system.

Keywords: Acoustic Environment, Speaker Authentication, Speech coding, Speech encryption with decryption, Speaker Identification

1. INTRODUCTION

The increasing demand of multimedia applications in communication system has tiled the way for secure communication. This is necessary to overcome unauthorized modifications and unwanted disclosure while transmitting speech and other data, especially in wireless channels. In secure speech communication systems the redundancy of the language plays an important role. The more redundant the language, the easier it is for an intruder to decipher the information with ease and convenience. That is why many real-world cryptographic implementations use a compression program to reduce the size of the signal before encryption. [1]. Compression of signal to lower rates with good speech quality not only eliminates the redundancy issue but also provides a lower bandwidth signal, which solves multiple problems in communication and multimedia applications [1]. The possible threats which could attack in passive or active way includes eavesdropping, modification, replay, masquerading, penetration and repudiation [2]. Speech is used by human to convey information to one another and this paper concentrates on the integrating of speech coding, speaker identification and strong encryption for providing secure communication.

This paper is organized as follows. Section 2 gives details about the Speaker Identification Section 3 presents the Speech Coding. Section 4 deals with the Speech Encryption and Decryption. Section 5 explains about the Literature Survey Section 6 gives about the Proposed Methodology. Finally, the Conclusion is summarized in Section 7 with future work.

2. SPEAKER IDENTIFICATION

During communication the enemy may attempts to act as the authorized user and also tries to gain access in it, so to avoid it the sender have to prove that he is an authenticate speaker against masquerade. A masquerade is a kind of attack where the enemy attempts to act as the authorized user and also tries to gain access in the communication .There by using control one may pass the operational order favorable to him which may leads to vulnerability. In such cases a security alert is needed which will save information. So using speech recognition algorithm, the speaker is identified. All the authorized speakers' names are indexed and by using the hashing algorithm the speaker can be retrieved from the database. The identified speaker can be authenticated and by combining secure hashing algorithm bits with encoded speech the message digest is recovered. The hash functions are used to compress an n (arbitrarily) large number of bits into a small number of bits (e.g. 512) secure information is transmitted towards the receiver through different channels. If there is any mismatch the destination recipient will get alert. There are many ideal cryptographic hash functions available [3]. Some of the hashing functions are Message Digest, Secure Hash Algorithm (SHA), and RIPEMD, GOST and HAVAL. Although there is a long list of hash functions many of the functions are found to be vulnerable. The SHA-0 and SHA-1were developed by National security Agency. Still there is a competition for the replacement for SHA-2[4], and also to ensure the long term toughness of applications that uses hash function.

3. SPEECH CODING

The need for elimination, reduction of the redundancy or irrelevant information from the analog

signals and gave birth to an area of speech coding. Commercial systems that rely on efficient speech coding include cellular communication, voice over internet protocol (VOIP), videoconferencing, electronic toys, archiving, and digital simultaneous voice and data (DSVD), as well as numerous PC-based games and multimedia applications. [5].The properties of speech coders include low bit rate, high speech quality, low coding delay, robustness in the presence of channel errors [6]. The speech coding can be Lossy or Lossyless. In lossy compression the actual data can be recovered from the compressed file. In Lossless compression, the actual data cannot be retrieved from the compressed file even it gives best possible quality for the given technique [7]. The quality of speech drops drastically if the encoding bit rate is reduced beyond a limit. Different speech coding schemes have resulted into various speech codecs and it can be broadly classified depending upon the bit rate at which they operate [8]. There is another way of classifying the speech coding techniques which is based on the concepts utilized and it can broadly be categorized into:

- Waveform Coding
- Parametric Coding
- Hybrid Coding

3.1 Waveform Coding

Waveform coding is used to analyze code and reconstruct original speech sample by sample. It includes time domain coding and frequency domain coding. The method such as Pulse Code Modulation (PCM), Differential PCM (DPCM) [9], Adaptive DPCM (ADPCM), Delta Modulation (DM), and Adaptive PCMID are some of the popular time domain waveform coding techniques and Transform Coding (IC), Sub band Coding (SBC) are a few spectral domain waveform coding techniques. The Pulse Code Modulation (PCM) [9] which is used to digitize the signals through signal conversion. Differential Pulse Code Modulation (DPCM) can be analog signal or a digital signal. It uses the baseline of PCM but it adds some functionality based on the prediction of the samples of the signal. In DPCM, first an estimate of each sample is found based on prediction from past few samples and then the difference of estimate from the original. The DPCM can provide PCM quality of speech at 56kbps. The ADPCM Adaptive Differential Pulse Code Modulation (ADPCM) [10] which is used to provide much lower data rates by using a functional model of the human speaking mechanism at the receiver end [11]. The frequency domain includes sub band coding and the transform coding. Advantages with SBC is that quantization noise in each band gets isolated from others and also bit rate optimization can be achieved by assigning more number of bits to speech signal in lower frequency bands(that is responsible for intelligibility) than in higher frequency bands. Variants of sub band coding are capable of providing speech at 9.6-32 kbps with speech quality comparable to that of ADPCM and ADM. In transform coding the signal is transformed to its representation in another domain in which it can be compressed well than in its original form.

3.2 Parametric Coding Methods

The parametric coding methods are capable of providing good quality of speech. Linear Predictive Coding (LPC), Residual Excited Linear predictive Coding (RELP) and Mixed Excitation Linear Predictive Coding (MELP) are the popular example under this class. The LPC method can produce intelligible speech at 2.4 kbps and it is one of the earliest speech coders proposed in literature. LPC at bit rates 600 BPS were given by Kang et al [12]. Atal and Remde established a Multipulse Excitation Model and given the improvement related to classical LPC, and then Self Excitated Vocoders and Residual Excited Linear Predictive (RELP) coders were introduced. The encoded prediction residual in RELP is used to excite the synthesis filter. Speech quality offered by RELP coders at 4.8 kbps is higher than that of two-state excited LPC coders. To remove the annoying artefacts in LPC such as buzzes, tonal noises etc., the MELP uses sophisticated excitation and improved filter model with additional parameters to capture signal dynamics with improved accuracy. MELP utilized vector quantization for LSF parameters and achieved improvement in speech quality with naturalness, smoothness and adaptability to diverse signal conditions, in comparison with 2.4 kbps LPC [13] without elevating the bit rate.

3.3 Hybrid Coders

Hybrid coding combines strengths of waveform coding and parametric coding techniques. Like a parametric coder, it relies on speech production model. Hybrid coders are used to encode speech, and the bandwidth requirements lies between 4.8 and 16 kbps. The hybrid coders include CELP, MPE and RPE coders. Multipulse Pulse Excited coding (MPE) and Regular Pulse Excited coding (RPE) techniques try to improve the speech quality by giving a better representation of the excitation signal and it produces high quality speech at 9.6 kbps. Codebook Excited Linear Prediction (CELP) technique can be also called as analysis-by-synthesis (Abs) technique with the bit rates of 4.8 kbps. CELP and its variants are the most outstanding representatives of this class which dominates medium bit rate coders. Idea of CELP was born as an attempt to improve on LPC coder. It basically covered a wide range of bit rates 4.8-16 kbps.

4. SPEECH ENCRYPTION AND DECRYPTION

Cryptography is the science of converting information from the comprehensible form to incomprehensible form for its secured

communication over the insecure channel. Encryption is a very common technique for promoting the security and it is a much stronger method of protecting speech communication than any form of scrambling. The various advantages of encryption includes that [14] it can protect information stored on the computer from unauthorized access and while it is in transit from one computer system to another it can protect information. In general, encryption techniques are classified into two broad categories [15] such as Symmetric and Asymmetric Encryption

4.1 Symmetric Encryption

The symmetric encryption also called as single-key encryption, one-key encryption or private key encryption. It uses the same secret key to encrypt and decrypt the information. It is essential that the sender and the receiver should know the secret key, which will be helpful to encrypt and decrypt all the information's. Fastness is the major advantage of using symmetric encryption. Some of the commonly used symmetric encryption algorithms are listed in table 2[16]

4.2 Asymmetric Encryption

In asymmetric encryption, different keys can be used to encrypt and decrypt the data. The encryption key is public whereas the decryption key is private. However, it has two major disadvantages such as it is based on nontrivial mathematical computations, and it is very slower than the symmetric ones. Some of the popular examples of asymmetric encryption algorithm include RSA, DSA and PGP [17].The RSA encryption is the best known public key

Symmetric Encryption Algorithm	Developer	Block size	Cryptanalysis resistance	Security
Advanced Encryption Standard	Vincent Rijmen and Joan Daemen in 2000	128-, 192- or 256-bit	It is very Strong against truncated differential, linear, interpolation and square attacks	More secure
Data Encryption Standard	IBM in 1977	64bit block	Vulnerable to differential and linear crypt analysis	Proven inadequate
Triple Data Encryption Standard	1978	64bit block	Vulnerable to differential, Brute force attacker could be analyze plain text using differential crypt analysis	One only weak which is exit in DES
CAST	Carlisle Adams and Stafford Tavares in1996	64bit block		Very fast and efficient [18]

algorithm, named after its investors: Rivest, Shamir and Adleman. The key used for encryption is a public key and the key used for decryption is a private key. The Digital Signature Algorithm (DSA) is a United States Federal Government standard or FIPS for digital signatures. It was proposed by the National Institute of Standards and Technology (NIST) in August 1991 for use in their Digital Signature Standard (DSS). PGP stands for Pretty Good Privacy which is a public-private key cryptography system allows for users to easily integrate the use of encryption in their daily tasks, such as electronic mail protection and authentication, and protecting files stored on a computer. It was originally designed by Phil Zimmerman. It uses idea, cast or triple des for actual data encryption and RSA (with up to 2048-bit key) or DH/DSS (with 1024-bit signature key and 4096-bit encryption key) for key management and digital signatures.

5. REVIEW OF LITERATURE

W.W Chang et al., in the "automated cryptanalysis of DFT-based speech scramblers" [19] presented an automated method for cryptanalysis of DFT based analog speech scramblers through statistical estimation treatment. E.V.Stansfield et al in Speech processing techniques for HF radio security [20] explains the techniques used to provide secure conversational speech communications over HF radio channel. An efficient implementation of multi-prime on DSP Processor proposed by K. Anand etal in 2003[21] implemented Montgomery squaring reduction method which speed ups by 10.15% for various key sizes Texas Instrument TMS320C6201 DSP Processor.

Cryptanalysis of adaptive arithmetic coding encryption scheme presented by J Lim et al in 1997[22] carried an analysis on different plaintext and cipher text and subsequent results were evaluated accordingly. Speaker Recognition from Coded Speech and the Effects of Score Normalization was proposed by R.B. Dunn et al [23]. This paper explains about the effect of speech coding on automatic speaker recognition where training and testing conditions are matched and mismatched. In the recognition performance there is little loss in the toll quality speech coders and more loss when lower quality speech coders are used. Both types of score

normalization considerably improve performance, and it can eliminate the performance loss when there is a mismatch between training and testing conditions.

Jan Silovsky et al presented their work in the paper titled Assessment of Speaker Recognition on Lossy Codecs used for Transmission of Speech in 2011[24].This paper investigates the effect of lossy codecs used in telephony on text-independent speaker recognition. Here the speaker recognition performance is degraded due to the bandwidth usage, transmission packet loss and utilization of discontinuous transmission techniques. There is only little loss in recognition performance for codecs operating at bit rates of about 15 kb/s and the best overall performance was observed for the SILK codec. R.B. Dunn et al present their work on" Speaker Recognition from coded Speech and the Effects of Score Normalization", in [25] MIT Lincoln Laboratory, Lexington. The author investigates the effect of speech coding on automatic speaker recognition when training and testing conditions are matched and mismatched. This paper use standard speech coding algorithms (GSM, G.729, G.723, MELP) and a speaker recognition system based on gaussian mixture models adapted from a universal background model for experimentation.

Aman Chadha, Divya Jyoti, M. Mani Roja, reviewed their work in the paper titled" Text-Independent Speaker Recognition for Low SNR Environments with Encryption" in [26]. The main objective of this paper is to implement a robust and secure voice recognition system using minimum resources offering optimum performance in noisy environments. The proposed text-independent voice recognition system makes use of multilevel cryptography to preserve data integrity while in transit or storage. The experimental results show that the proposed algorithm can decrypt the signal under test with exponentially reducing Mean Square Error over an increasing range of SNR. Further, it outperforms the conventional algorithms in actual identification tasks even in noisy environments.

A. R. Stauffer, A. D. Lawson **,"** Speaker Recognition on Lossy Compressed Speech using the Speex Codec" in 2009 in[27]. This paper examines the impact of lossy speech coding with Speex on GMM-UBM speaker recognition (SR). Results show that Speex is effective for compression of data used in SR and that Speex coding can improve performance on data compressed by the GSM codec .J. D. Gibson, A. Servetti focus on Selective Encryption and Scalable Speech Coding for Voice Communications over Multi-Hop Wireless Links in [28]. This paper proposes and investigates a combination of scalable speech coding and selective encryption for secure voice communication over multi-hop wireless links that addresses both the efficient use of network, node resources and security against unwanted eavesdroppers. It is shown that when the Shannon lower bound is satisfied with equality for rate distortion optimal scalable coding, transmission of the enhancement layer in-the-clear provides no information regarding the core layer.

6. PROPOSED METHODOLOGY

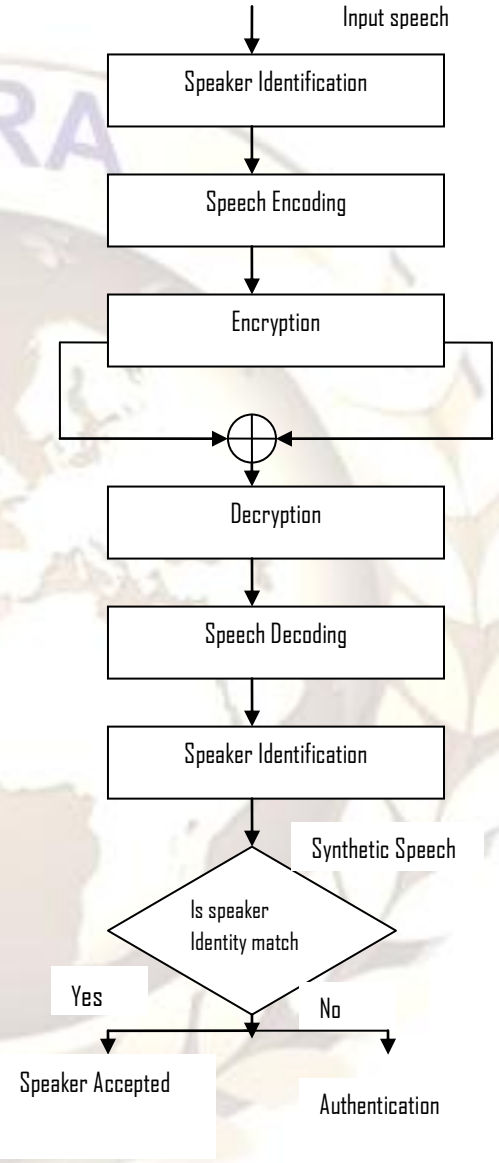

Fig. 1 Secure Speech Communication

The overall proposed methodology for secured speech communication is given in figure 1 which is illustrated as follows:

- First of all the input speech signal is given by the sender.
- But before transmitting to the receiver the speaker is identified by using the speech recognition algorithm to prove that he is an authorized speaker.

- Along with the speaker identification the original speech is compressed using the compression algorithm
- As the next step the compressed speech is encrypted and the secure information is transmitted towards the receiver through different channels
- The receiver decompresses the data and decrypts the information to get the original speech and also he checks by using the hashing algorithm that the original message came from the authorized user.
- If the speaker identity matches then the receiver follows the instructions according to the sender
- If there is any mismatch the destination recipient will get alert.

7. CONCLUSION AND FUTURE WORK

Speech processing for secured communication has been in development for more than 50 years. This paper gives the various techniques in the field of Speech Coding, Speaker Identification with Encryption and Decryption. The various approaches available for developing a secured communication are clearly explained. In recent years, the need for secured communication research based on Speech Coding with cryptography has highly increased. This paper is based on compression technologies that exploit this technology with encryption and decryption for an environment that promotes and facilitates the use of safe communication. More specifically, the application takes responsibility for the speaker identification using the cryptographic hash functions so that the unauthorized speakers will find difficult to trace out the original data. The future work is to find out the optimal method suitable for this environment to provide better secured communication.

REFERENCES

[1] A. Jameel et al, " A robust secure speech communication system using ITU-T G.723.1 and TMS320C6711 DSP", Microprocessors and Microsystems, volume 30,2006,Pages 26-32

[2] A. Jameel, "Transform-domain and DSP based secure speech communication", Microprocessors and Microsystems,2007, 335–346

[3] http://en.wikipedia.org/wiki/Cryptographic_hash_function

[4] http://csrc.nist.gov/groups/ST/hash/sha-3/index.html

[5] Mark Hasegawa-Johnson et al," Speech Coding: Fundamentals And Applications", Wiley Encyclopedia of Telecommunications

[6] Akella Amarendra Babu et al, ,"Robust speech processing in EW environment", International Journal of Computer Applications (0975 – 8887) Volume 38–No.11, January 2012

[7] Dr.V.Radha et al," Comparative Analysis of Compression Techniques for Tamil Speech Datasets, IEEE, ICRTIT, June 3-5, 2011

[8] Venkatesh Krishnan," A Framework For Low Bit-Rate Speech Coding In Noisy Environment", A school of Electrical and Computer Engineering

[9] P.Cummiskey et al,"Adaptive Quantization in Differential PCM Coding of speech, "Bell sys.Tech.J.Vol.52.No.7.p.1105.sept.1973

[10] G.Kang and D.Coutler,"600 Bit-per-second voice digitizer(linear predictive format coder),"NRL Report 8043,Nov 1976

[11] Jorgen Ahlberg," Speech & Audio Coding" TSBK01 Image Coding and Data Compression Lecture 11, 2003

[12] G.Kang,"Application of linear prediction to a narrow band voice digitizer,"NRL Report 7774,Oct.1974

[13] Benesty,Sondhi, Huang, "Springer Handbook of Speech Processing"

[14] Nehaluddin Ahmad,,"Privacy and the Indian Constitution: A case study of Encryption", Communication of the IBIMA volume 7,2009 ISSN:1943-7765

[15] S.Rajanarayanan and A. Pushparaghavan, "Recent Developments in Signal Encryption – A Critical Survey", International Journal of Scientific and Research Publications, Volume 2, Issue 6, June 2012, ISSN 2250-3153

[16] Hamdan.O.Alanazi etal, "New Comparative Study Between DES, 3DES and AES within Nine Factors", Journal of computing, volume2, issue3, march 2010

[17] http://zybersene.blogspot.in/2012/06/symmetric-encryption-vs-asymmetric.html

[18] Http://www.omnisecu.com/security/public-key-infrastructure/symmetric- encryption-algorithms.htm

[19] W.W Chang et al., "The automated cryptanalysis of DFT-based speech scramblers," IEICE Trans .Information and system,E83,pp.2107-2112,2000

[20] E.V Stansfield et al., "Speech processing techniques for HF radio security," IEEE Proce.,vol.136,1989

[21] K.Anand,et al.,"An efficient implementation of multi-prime on DSP Processor,"inProc.ICASSP,2003,pp.413-416

[22] J.Lim et al.,"Cryptanalysis of adaptive arithmetic coding encryption scheme," in Proc. ACISP,1997 pp.216-227

[23] R.B. Dunn, T.F. et al, "Speaker Recognition from Coded Speech and the Effects of Score Normalization", IT Lincoln Laboratory, Lexington, MA

[24] Jan Silovsky, Petr Cerva, Jindrich Zdansky,"Assessment of Speaker Recognition on Lossy Codecs Used for Transmission of Speech", 53rd International Symposium ELMAR-2011, 14-16 September , Zadar, Croatia

[25] R.B. Dunn, T.F. Quatieri, D.A. Reynolds, J.P. Campbell, "Speaker Recognition from Coded Speech and the Effects of Score Normalization", IT Lincoln Laboratory, Lexington, MA

[26] Aman Chadha, Divya Jyoti, M. Mani Roja, "Text-Independent Speaker Recognition for Low SNR Environments with Encryption", International Journal of Computer Applications (0975 – 8887) Volume 31– No. 10, October 2011

[27] A. R. Stauffer, A. D. Lawson," Speaker Recognition on Lossy Compressed Speech using the Speex Codec" 2009 ISCA 6-10 September, Brighton UK

[28] J. D. Gibson, A. Servetti ," Selective Encryption and Scalable Speech Coding for Voice Communications over Multi-Hop Wireless Links "