International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 3, Issue 2, February 2016

Secure Key Generation Using Contextual Audio with Intruder Detection

[1]Aparna A, [2]Ajish S
[1] PG Scholar, [2] Assistant Professor In CSE
College of Engineering, Perumon, (CUSAT)
[1]aparna0126@gmail.com, [2] ajishs2014@gmail.com

Abstract- **The implementation of a secure communication channel for the transfer of confidential data is one of the primary concerns in the field of security. The most important entity in a secure communication process is the cryptographic key generation used for encryption and decryption purposes among communicating partners. The key used in such a process must be extremely strong and secure enough so that it is untraceable to attackers. Many methods have been proposed for generating a secure key ranging from biometric techniques to that of the modern quantum cryptography. An approach to generate a shared key between two users is to exploit the contextual information present in the environment in which they are present. Many schemes are exists there for generating the shared cryptographic keys from recorded contextual information by each of the devices. Keys generated from contextual possess a greater level of robustness and uniqueness. A common audio fingerprinting approach in connection with error correction codes and fuzzy cryptography provides a predefined noise threshold level. The system is vulnerable to attackers from the same scenario even though we are considering the contextual audio. In this paper, we propose a new key generation scheme to build a secure communication channel between two users in the same environment.**

Index Terms— **Contextual Information, Encryption, Fingerprinting, Key Generation, Noise**

I. INTRODUCTION

One of the major problems in the field of communication is that the establishment of a secure communication channel. Nowadays a large number of people have enough knowledge about the underlying technology which increases the chance for human impact on security. There exists many possible ways to establish authentication between devices in a communication scenario. One of the most widely used primitive approach, password based authentication causes many possible threats on security. Such problems can be eliminated up to certain extend by choosing contextual information as the seed for key generation. The environmental stimuli such as light, audio, RF Channel, proximity and temperature can be used as contextual information. Among that ambient audio is most widely used since it is a spatially centered context and that's why it offers more uniqueness among keys generated by each of the devices in a communication scenario.

Human impact is one of the major threats in the set of security risks. One of the efficient ways to establish a secure communication channel among devices is based on similar audio patterns recorded from the context. Features from ambient audio are used to generate a shared cryptographic key between devices without exchanging information about the ambient audio itself. The features extracted will be utilized for the key generation process. This paper explores a common audio-fingerprinting approach and account for the noise in the derived fingerprints by employing a suitable error correction scheme. The proposed scheme constitutes a totally unobtrusive but cryptographically strong security mechanism based on contextual audio recorded by each of the devices. An intruder who may present in the same environment is also able to generate the same set of keys. Such a situation can be prevent up to a certain extend by incorporating an attribute based encryption scheme or techniques like face recognition.

The paper points to develop a method for secure communication based on contextual information. Contextual audio is used as the seed for key generation since it offers higher uniqueness for cryptographic keys compared to other contextual information. The unbiased nature of keys offers more security than the existing methods. Most of the existing audio cryptographic techniques rely on the concept of considering the music specific properties of audio such as pitch, contour etc. For the secure shared key generation between devices the energy difference between successive bands taken as the parameter from the recorded audio by each of the devices. A fuzzy cryptographic scheme utilizes

by the system tolerates a certain predefined amount of noise also. The researches which are related to this system is restricted to its initial stage of implementation and problem finding. There exist a major problem like when an intruder who present in the same environment can also generate the same set of keys and thereby the decryption of secretly shared message is feasible for the intruder. For the intruder detection in such a situation, an efficient attribute based encryption scheme can be utilized. For the intruder detection, an efficient attribute based encryption scheme like face recognition can be incorporate with the existing system.

A framework associated with any of the context based device authentication consists of mainly five modules:

- ***Device Synchronization:*** The system desires synchronization among participating devices for sharing a common secret without any fluctuation based on considered context.
- ***Feature Extraction:*** The contextual features are extracted from the contextual information using a feature extraction method.
- ***Context Processing:*** Preprocessing techniques like smoothing and noise removal are applied on the extracted features.
- ***Key Generation:*** Using the compact notation of extracted feature representation, a key is generated by using any key generation algorithms.
- ***Communication:*** The generated secure key is used for secure encrypted communication.

II. RELATED WORK

Contextual data is in some aspects similar to biometric data. In contrast to contextual data for the use of cryptographic applications, biometric data has some unfavorable properties such as limited availability of biometric information. The security provided by biometric data can be broken by a determined adversary by taking high quality photographs of an Iris or face of the particular user. Biometric data can't be changed significantly. In order to increase the burden for an adversary to break a security system, it is beneficial to periodically change the secret utilized. If we use biometric features for security, this requirement cannot be met.

Quantum cryptography also considers the aspect of contextual information. Audio as a context for key generation [1] never requires an additional or change in the existing infrastructure. The greatest challenge in mobile device pairing for information transfer is the authentication of participant devices in a similar environment. During the pairing process, it is difficult to validate the intended mobile device to communicate.

Ngu Nguyen et al., 2012, [2] proposed an un-obstructive mechanisms to establish synchronization between devices based on the environmental audio for establishing a secure key generation. With the help of inbuilt microphones, each mobile device willing to communicate capture synchronized audio samples from the environment in which they are present. Each device then extracts the perceptual features of the recorded audio and from that computes a binary characteristic sequence. This sequence is unique for each of the captured audio samples.

This unique binary code will be fall on to a code-space of an error correcting code. In general, a fingerprint will not match in every aspect with any other fingerprint that is generated from the same environment due to the presence of noise. In the considered context, audio is spatially centered at different instants. Fingerprints generated from similar ambient audio may not be considerably similar, due to noise and inaccuracy in the audio sampling process. It is unlikely that two fingerprints are flawlessly identical to each other. Devices utilize their error codes for mapping fingerprints to the corresponding code words. The fingerprints having a hamming distance within a predefined threshold are considered as secure keys. The hamming distance between the fingerprints is directly proportional to the distance between participating devices.

Sigg, S et al., 2012, [3] constructed an Ad hoc Pairing application; an audio based secure system for Android mobile devices. It utilizes the concepts of Fuzzy cryptography for codeword mapping. Synchronization in audio samples is achieved with the help of an approximate pattern matching. For establishing a secure communication key between unacquainted devices require explicit user input to provide a shared piece of information. Devices which are willing to communicate each other are placed at a distance 'd' from an audio source. Each device records ten audio samples, the creation of fingerprints begins with the detection of the top 10 matching positions of a common pattern extracted from each audio sequence[4]. Each device can function either as a sender or as a receiver. As a result, each device can have 10 keys. Senders use these keys for encrypting the original data chunk, and on the other side the corresponding receiver use the key with best matching result for decryption purposes. In case the attempt for decryption fails, the receiver can use the next top similar key, and this process will be continued up to 10 trails until the decryption process is successful.

Guido Stromberg et al., 2007, [5] proposed a system and proved that it is also possible to analyze the accelerometer reading which provides a measure of the amount of shaking among devices. The extraction of characteristic features from simultaneous shaking processes

is extremely difficult. It requires repeated hash exchanges of key-sub-sequences until a common secret found.

Suhas Mathur et al., 2011, [6] Both the sender and the target can easily observe a typical laser beam (light). With the aid of high speed cameras, it is possible to capture the modulated signals with enough accuracy from surroundings, which can be used to generate a secret key among devices for a spontaneous device authentication

III. THE PROPOSED SCHEME

Key aspect of audio based encryption is the audio fingerprinting; there exist mainly two aspects of audio fingerprinting. Most of the researchers considering audio fingerprinting in the applications like finding duplicate tracks in a large database. In these applications we are considering the music specific properties of audio like pitch, contour etc. As per the contextual view it is not enough to consider the structured audio only for key generation, due to the presence of noise. An error correction scheme in combination with Fuzzy Cryptography tolerates a specific amount of noise in audio. The seed for shared key generation among devices are the ambient audio, an audio sequence that does not satisfy any formal parameters appropriately. For establishing a secure communication channel all devices should capture the instantaneous ambient audio and from that a fingerprint will be generated, which is a compact representation of the large audio stream recorded from the context. For fingerprints with a hamming distance within the error correction threshold utilized as secure keys. Devices exploit the error correction capabilities of the error correcting code utilized to map fingerprints to codewords. The steps involving are given below

A. Audio Acquisition and Device Synchronization

In the proposed scheme the first step involves the extraction of perceptual features from a piece of audio. Yang presented a method to identify energy peaks in signal spectrum to extract a unique pattern [7], this scheme supported by a general framework presented by Yang [8].

To create fingerprints split up the entire sequence 'S' into 'n' frames each of them have equal length. Then on each frame a DFT weighted by a Hanning window is applied. After that each frames divided into 'm' non overlapping frequency bands and computes sum of energy values on each band and stored in an energy matrix. Using this matrix a fingerprint 'f' is generated from the energy matrix 'E', where each bit describes the difference between two successive frames.

$$f(i,j) = \begin{cases} 1, & (E(i,j) - E(i,j+1)) - E(i-1,j) - E(i-1,j+1)) > 0 \\ 0, & \text{Otherwise} \end{cases}$$

The Accoustid Chromaprint, a tool which is used to extract the informal parameters of audio stream to generate a unique fingerprint, chroma features are extensively used in a number of music retrieval applications. Acoustic fingerprinting is a technique for identifying songs from the way they "sound" rather from their existing previously collected metadata. This plugin uses an open-source fingerprinting technology called Chromaprint and its associated Web service, called Acoustid. First, it can be trickier to set up the native fingerprinting library, whereas all of the beets core is written in pure Python. Also, fingerprinting takes significantly more CPU and memory than ordinary tagging, which means that imports will go substantially slower.

B. Reed Solomon Error Correction

A perfect match in fingerprints is unlikely since devices are spatially separated, not exactly synchronized with each other. Even though a number of devices which presented in a similar environment and records the contextual audio simultaneously, definitely there will be minute or large variations in fingerprints generated by each of the devices due to the presence of noise. The system should satisfy the requirement of a unique code word sequence for particular environments. The Reed-Solomon error correction scheme is most suitable for such a requirement. In coding theory, the Reed–Solomon code belongs to the class of non-binary cyclic error-correcting codes. The Reed–Solomon code is based on univariate polynomials over finite fields. It is able to detect and correct multiple symbol errors. By adding 't' check symbols to the data, a Reed–Solomon code can detect any combination of up to 't' erroneous symbols, or correct up to t/2 symbols. As an erasure code, it can correct up to 't' known erasures, or it can detect and correct combinations of errors and erasures. Furthermore, Reed–Solomon codes are suitable as multiple-burst bit-error correcting codes, since a sequence of b + 1 consecutive bit errors can affect at most two symbols of size 'b'. The choice of 't' is up to the designer of the code, and may be selected within wide limits. A Reed-Solomon code is specified as RS(n, k) with s-bit symbols. This means that the encoder takes k data symbols of s bits each and adds parity symbols to make an n symbol codeword. There are n-k parity symbols of s bits each. A Reed-Solomon decoder can correct up to t symbols that contain errors in a codeword, where 2t = n-k.

C. Fractal Encryption Scheme

Security is the major concern in the set of communication scenarios. In the proposed system we are considering the contextual information for the purpose of key generation among the participating devices. Due to the spatially centered behavior of contextual audio, it offers more uniqueness regarding the generated keys. It is

convenient to utilize the standard encryption schemes like DES, AES etc. The system demanding higher security requires much more complicated and efficient encryption scheme. A probability based key generation scheme, Fractal Encryption Key Generation associated with a set of keys instead of a single key generated by the standard encryption schemes. The major motivation of using fractal encryption is to reduce the computation cost and increase the security for the public-key systems, and this leads us to propose new public-key cryptosystem based on Fractal.

D. Face Recognition for Intruder Detection

In the proposed scheme there is a loophole of stealing the contextual audio sequence by an intruder. For offering much higher security, the incorporation of an advanced technology is required. The number of communicating partners in a similar environment is possible to be restricted by using the technique of face recognition during the time of user authentication. The faces of all users in particular scenario will trained in advance. The users face attributes compares against the previously trained attribute set to check whether he/she is a authenticated user or not.

Face Recognition is a term that includes several sub-problems. There are different classifications of these problems in the bibliography. The input of a face recognition system is an image or video stream. The output is an identification or verification of the subject or subjects that appear in the input image or video. Commonly, approaches define a face recognition system as a three step process – and is as shown in Figure 1.1. From this point of view, the Face Detection and Feature Extraction phases can run simultaneously.

Figure 1.1: A generic face recognition system

Face detection has many several well known challenges to deal with [9]. They are usually present in images captured in uncontrolled environments, such as surveillance video systems. These challenges can be attributed to some factors:

- ***Pose variation:*** The best scenario for a face detection would be one in which only we have frontal images were involved. But, as stated, this is very unlikely in general uncontrolled conditions. Moreover, the performance of face detection algorithms decreases rapidly when there are large pose variations. It is a major research issue. Pose variation can happen due to the subject's movements or camera's angle.
- ***Feature occlusion:*** The presence of elements like beards, glasses or hats introduces high variability. Faces can also be partially covered by objects or other faces.
- ***Facial expression:*** Facial features also vary greatly because of different facial gestures. Imaging conditions. Different cameras and ambient conditions can affect the quality of an image, affecting the appearance of a face.

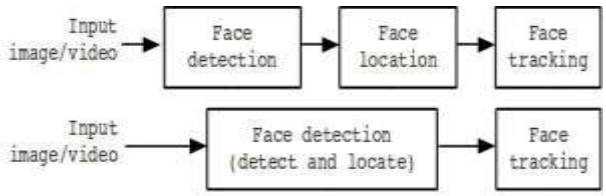

Figure 1.2: Face detection processes

Each possible face candidates is normalized by approximating the shirring angle due to head movement. Furthermore, the lighting is reduced by transforming their histograms into the histogram of a reference face image.

IV. CONCLUSION

A secure communication channel among devices is established by utilizing the audio as contextual information. The approach was exemplified for ambient audio and can be similarly applied to alternative contexts. In this paper, we have reviewed some of the techniques that make use of the contextual information to generate a safe and secure key. Adds face recognition as a solution for finding intruders in the scenario. Additionally, we analyzed the properties of fingerprints and estimated the entropy in statistical tests. There exists a major problem like when a third person who present as an intruder in the communicating environment he/ she can easily do the decryption of the secret message. Since he/ she is also generating the same set of keys in the environment as that of the other devices. This security issue is possible to be solved to a certain extend by incorporating an attribute based encryption scheme like face recognition also.

REFERENCES

[1] Ngu Nguyen, Stephan Sigg, An Huynh, and Yusheng Ji, "Using ambient audio in secure mobile phone communication", International Conference on Pervasive Computing and Communications, pp. 431-434, 2012

[2] Nguyen, Ngu, et al. "Pattern-based alignment of audio data for ad hoc secure device pairing.", Wearable Computers (ISWC), 2012 16th International Symposium on. IEEE, pp. 88-91, 2012.

[3] Sigg, Stephan, Dominik Schuermann, and Yusheng Ji. "Pintext: A framework for secure communication based on context.", Mobile and ubiquitous systems: Computing, networking, and services. Springer Berlin Heidelberg, pp. 314-325, 2012.

[4] D. Bichler, G. Stromberg, and M. Huemer, "Innovative Key Generation Approach to Encrypt Wireless Communication in Personal Area Networks," Proc. IEEE GlobeCom, pp. 177-181, 2007.

[5] D. Bichler, G. Stromberg, M. Huemer, and M. Loew, "Key Generation Based on Acceleration Data of Shaking Processes", Proc. Ninth Int'l Conf. Ubiquitous Computing, J. Krumm, ed., pp. 304-317, 2007.

[6] Mathur, Suhas, et al. "Proximate: proximity-based secure pairing using ambient wireless signals." Proceedings of the 9th international conference on Mobile systems, applications, and services. ACM, pp. 211-224, 2011.

[7] C. Yang, "MACS: Music Audio Characteristic Sequence Indexing for Similarity Retrieval," Proc. IEEE Workshop