Secure E- Commerce Transaction using Noisy Password with Voiceprint and OTP

Komal K. Kumbhare
Department of Computer Engineering
B. D. C. O. E.
Sevagram, India
komalkumbhare27@gmail.com

Prof. K. V. Warkar
Department of Computer Engineering
B. D. C. O. E.
Sevagram, India
kanchan.22warkar@gmail.com

Abstract**—** E-Commerce application is used for trading products by using communication technology. To protect customer's privacy and against fraud, special attention must be given to the issues related to security of e-commerce transactions. Web application uses traditional passwords which are vulnerable to replay attack. To overcome this problem OTP mechanism is used. Biometric technique measures unique individual features of user for personal recognition. In this paper, we have implemented a new password technique, i.e. Noisy Password to protect against attacks like shoulder surfing, key loggers, etc. The proposed idea is to use biometric with cryptography to enhance security of OTP.

Keywords*- Noisy Password; Voiceprint Biometric; Vocal Tract Feature Extraction; One-Time-Password; hash function.*

*****

I. INTRODUCTION

Today everyone uses e-commerce application for trading the products like in online shopping. E- Commerce application provides the platform for selling and buying of items by using the computer networks and communication technology. Hence it becomes essential that we provide high security to E-Commerce transactions. There is need to protect against attacks like eves dropping, Trojan attack, Key logging, etc.

Most of the websites or such type of applications depends on relatively weak protection mechanism for user authentication by using plaintext of password and user identification. This password can be transferred by using secured channel but still this approach is non-resistant to some attacks. We can see, this authentication method is vulnerable to attacks like shoulder surfing or eves dropping. Hence, instead of using traditional password method, the Noisy Password, proposed by K. Alghathbar and H. Mahmoud [], have been used in proposed system to resist these types of attacks. This new password method is made of three different sets of alphanumeric characters. This method is resistant to key loggers, Trojan program, shoulder surfing and eves dropping etc.

There are mainly three different authentication areas out of which biometric technique assures high degree of security. There are different types of biometric techniques that can be used to authenticate the legitimate user like fingerprints, face etc. But biometric scheme, developed using image processing methods, require scanning devices of high cost. In this paper, voiceprint biometric is implemented because the implementation cost of this technique is less as compared to the other biometric techniques like fingerprint, iris scan, face, palm-vein, etc.

The traditional text based passwords are vulnerable to reply attack. To overcome this problem, many applications use two factor authentications. For that they commonly use One Time Passwords. One Time Password is string of numeric or alphanumeric characters that are automatically generated by using different schemes. There are mainly three approaches to generate the string of OTP. They may be based on time-synchronized, using mathematical function which uses previous OTP string to generate new OTP or the mathematical function that chooses random number. Most of the mobile application uses the OTP mechanism for two factor authentication. But over the time it has been proved that OTP mechanism is no longer that much secure. It is necessary to encrypt the OTP message. In the proposed system, the OTP is encrypted by using the voiceprint template.In the work reported in this paper, we aim to achieve a high level of security of E-Commerce application and transaction.

This paper is organized as follows. In Section 2, we describe the related work and literature review. In Section 3, we propose a system that enhances the security of E-Commerce application and transaction. In Section 4, we discuss the techniques we use for the implementation of the proposed system. Section 5 shows the experimental results of the proposed system. Finally, Section 6 concludes our work.

II. RELATED WORK

The OTP is used as multi factor authentication method. But providing the security to OTP message is also necessary. D. Mahto and D. K. Yadav [1] have proposed an approach that increases the security level of OTP. They have used the ECC algorithm and the palm vein biometric. Since the implementation cost of palm vein biometric is high, it limits the number of users of the proposed system. The proposed system can only be used at office level.

L. Muda, M. Begam and I. Elamvazuthi [4] have presented the methodology for the voice recognition system. In which they have used MFCC for feature extraction and for feature matching they have used the DTW algorithm. They have briefly discussed about the steps involved in the feature extraction process of voice signal. They have experimented with one female and one male speaker using Mono microphone and Gold Wave software. Since they have used the microphone to acquire the voice sample of the speaker, they sampled the signal at the rate of 16 kHz.

Again K. Chakraborty, A. Tabele S. Upadhya [5] have implemented MFCC for voice identification. They studied about the MFCC coefficients extracted from voice samples of two speakers uttering the word HELLO at two different time. They found that the MFCCs are unique for every speaker, but certain variations can be observed depending on the environment of recording area.

E. Kalaikavitha and Juliana Gnanaselvi [6] proposed the idea to enhance the security level of OTP by using encryption algorithm. In their proposed approach, after entering the username and password, server encrypts the generated OTP using the AES algorithm and sends it to user. Since OTP is in encrypted form user cannot read it. So user forwards that OTP with password to the system. The system decrypts the OTP and verifies the OTP, password and mobile number of the user. But the system load is increased by encryption and decryption of OTP.

Pattern classification techniques are used for speaker recognition which includes DTW, GMM, SVM, HMM etc. Loh Mun Yee and Abdul Manan Ahmad [7] presented the overview of pattern classification techniques. They particularly focused on comparing the performance of DTW, GMM and SVM algorithms. They have used the TIMIT speech database for implementation purpose. It can be seen that SVM gives the worst result.

Parwinder Pal Singh and Pushpa Rani [8] have presented an approach for feature extraction of voice signal by using Mel-Scale Frequency Cepstral Coefficients. This approach is a nonparametric frequency domain. They took the voice samples of isolated words and then the noise removal is done by using Praat tool. For feature extraction MFCC is used.

S. B. Dhonde and S. M. Jagade [9] have given the brief study of Linear Predictive Cepstral Coefficients (LPC), Perceptual Linear Prediction Coefficients (PLPC) and Mel - Frequency Cepstral Coefficients (MFCC). According to survey, LPC is an effective technique for speech recognition but it is not good for speaker recognition.

In [10], a brief analysis of MD5 and SHA is elaborated. They analyzed the MD5, SHA1, SHA256, SHA384 and SHA512 by using different parameters which includes key length, block size, cryptanalysis, total steps and rounds required to execute the hash function. Though MD5 is faster as compared to SHA, it provides less security. Also further study shows that MD5 is not resistant to collision attack and pre-image attack.

E. Sediyono, K. I. Santoso and Suhartono [13] have combined the two concepts to develop the authenticated, non-memorized and secured OTP. They have concatenated the some attributes of user ID and then the generated OTP is sent to the user with the help of SMS. Again they encrypted the concatenated attributes of user ID using MD5 algorithm.

A system proposed in [13] has high degree of security. They combined the OTP, SMS Gateway and MD5 hash function. Authors have minimized the delay time to 3 min that hackers will not get time to attack or eavesdrop. To overcome the drawbacks of one-time-password and traditional static text-based passwords, the new password scheme is proposed by [2], called noisy password. Noisy password proposed by [3], consists of four parameters, actual password or fixed set, terminator, safeguard number and the variable set i.e. noise. All the parameters are of alphanumeric variable length. The noisy password is resistant to the shoulder surfing attack. But the system was not user friendly. The error rates of the system were high. To reduce the error rate they proposed same scheme include time delay. It seems that the percentage error in password entry of noisy password but less than PEN technique as shown in [3].

III. PROPOSED SYSTEM

The proposed system consists of three parts, registration, authentication and OTP generation. At the time of registration, users have to enroll themselves by providing their personal information along with that they have to provide the part of noisy password and three voice samples. The voice samples are acquired using the smart phone. The part of noisy password gets saved in MySQL database. Then the feature extraction process is done on the respective three voice samples of the user. The output of this process is the feature vector which will get stored in the mat file. The Mel-Frequency Cepstral Coefficient technique is used for the feature extraction. The feature extraction process is implemented by using the MATLAB Simulation tool.

Next two phases are authentication and OTP generation. So, in the proposed system, user is authenticated with help of Noisy Password and their real time voice sample. The input voice sample is again goes through the feature extraction process and then matched with the existing voice feature vector of that user.

The traditional text based passwords are vulnerable to reply attack. To overcome this problem, many applications use two factor authentications. For that they commonly use One Time Passwords. One Time Password is string of numeric or alphanumeric characters that are automatically generated by using different schemes. In the proposed system, the OTP is generated using the mathematical function that chooses the random number as OTP. The OTP is used as multi factor authentication method. But providing the security to OTP message is also necessary. The security of the OTP is enhanced by encrypting the OTP message. The OTP is decrypted by using the SHA256 hash value of voice feature vector as private key. The user then input the plaintext of OTP into the application. The server verifies the plaintext of OTP and authenticates the user.

IV. IMPLEMENTATION

This section describes the authentication methods used in the proposed system.

A. *Noisy Password Method*

Today attacker can intrude by simply shoulder peeping the user because of wide view angles of laptop screens. To overcome the problem of shoulder surfing or noisy password is used. The noisy password consists of three parts i.e. the terminator set, the actual password and the noise. Every time whenever the users want to access their account, they have to enter different string of alphanumeric characters, actual password embedded in it. Noisy password consist of actual password (F), Terminator set (X) and the Variable set (V). All the three sets are of variable alphanumeric type. The following section describes the algorithms for selection of password and extraction of actual password from the password string.

1) Selection of Password: In conventional authentication system, userID and the corresponding password are used for the authentication. In the proposed system the traditional password is replaced by the noisy password. Same as the traditional password, user have to register themselves by providing the part of noisy password, i.e. actual password (F) and terminator set (X) in such a way that the length of both, actual password and terminator set must be same. Figure 5.1 shows the flowchart for selecting the noisy password. At the time registration, user will be asked to enter their userID, actual password and terminator set along with their personal information. These information will get store in database. Steps for selection of password is as follows:

Algorithm:

Step 1: Read user ID, actual password and terminator from user but length of actual password and terminator must be same.

Step2: Store the credentials in database and initialize counter.

2) *Extraction of Password:* The extraction of the actual password is the sole responsibility of the server. The subset of terminator set denotes the end of subset of noise or variable set (V) and the beginning of the subset of actual password. Since every time the user access his account using different password, it makes the system resistant to the shoulder surfing or peeing attack and key loggers, etc. Now if user wants to change his password, he may change either actual password (F) or terminator set (X) or both. The Flowchart of algorithm for extraction process is given in Fig. 1.

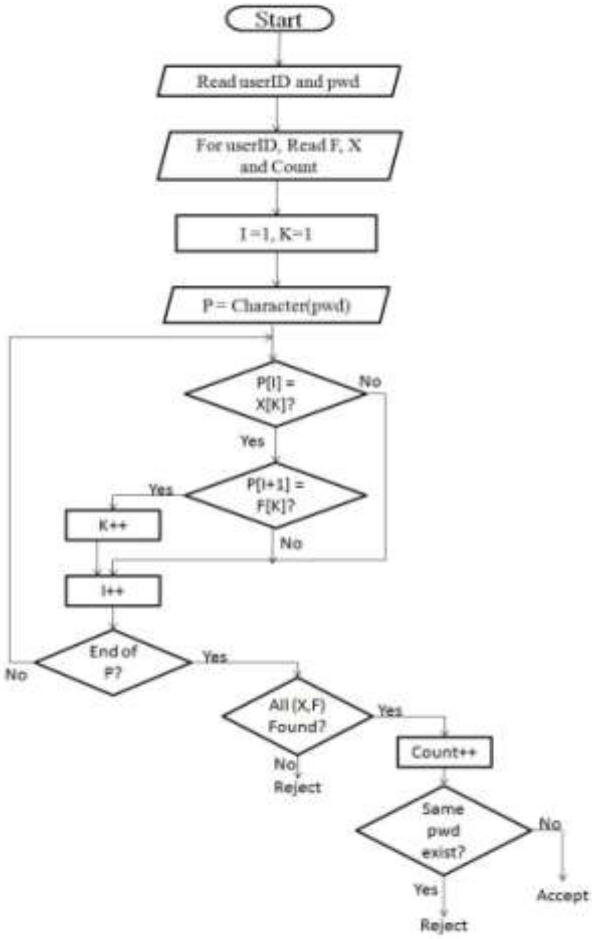

Figure 1. Extraction of Actual Password from Password String

Figure 1 describes how the actual password is extracted from the variable length password string. Let us consider the following example.

Example:

Actual Password (F) = 6,3,5,1
Terminator set (X) = 4,2,7,8
Variable set (V) =
634623756924696362585698888455475321223754325667574564322367674325568798654328534851808855322568997655444 5837438

Now at the time of log in, user provides variable length noisy password, containing variable noise, actual password and terminators embedded in it.

At the server side, program traverse the string and check for first character of terminator set and whether the immediate character is the corresponding character of actual password but in sequence. The gray colored italic character shows the terminator set and the bold character shows the subset of actual password. If the pair (X, F) is found, the server will then check for next character pair. Once all the pairs are found, the server will permit the user to log in to the system otherwise reject it.

63462375692*4***6**9636258569888845547 5*2***3**122375432566 7*7***5**
456432236767432556879865432853485*8***1**088553225689976
554445837438

B. *Voiceprint Biometric*

The speech contains many levels of information. First is message in the form of words. At other levels, speech contains information about emotion, gender and the identity of the speaker. The aim of speaker recognition is to identify the speaker by extraction and recognition of the information contained in the signal. Every person has differences in voice depending on the construction of their articulator organs, such as length of the vocal tract, characteristics of vocal chords etc. Speaker recognition system uses feature derived only from the vocal tract. Speaker identification consists of the process of converting voice samples into features. These features are in the form of acoustic vector. The Mel-Frequency Cepstral Coefficients and Gaussian Mixture model are used for feature extraction and recognition. Fig. 2 shows the block diagram of computation of MFCC.

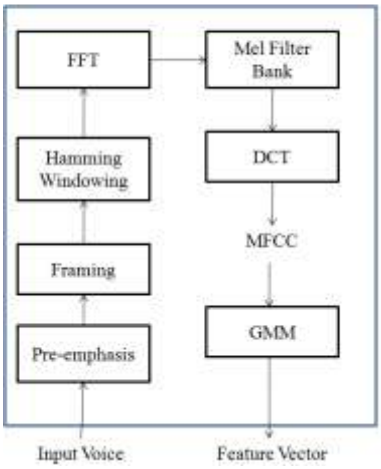

Figure 2. Computational Steps for Speaker Recognition

The steps involved in this process are given below.

1) *Pre- emphasis:* This is the first step for calculating the MfCC. In this step, the voice signal is passed to the high pass filter. The pre-emphasis phase compensate the high frequency part which was suppressed during the sound production process of humans.

2) *Framing and windowing:* In the next step,the input voice signal is segmented into overlapping frames of 20 ms. In this implementation, the sample rate is 8kHz and the frame size is 160 samples per frame. To keep the continuity of the first and last points in the frames, each frames is multiplied with hamming window. To minimize the distortion in signal, hamming window taper the voice sample to zero at the beginning and end of the each frame. The coefficients of hamming windows are computed by the given equation.

3) *Fast Fourier Transform (FFT):* FFT is used to convert each frame of N Samples from time domain into frequency domain and to obtain the magnitude frequency response of each frame. Instead of the frequency response we need envelope of the response of frequency for that we need triangular bandpaas filters. The triangular bandpass filter is explained in the next step.

4) *Mel-Filter Bank:* The mel filter bank is a set of 20-40 triangular filters. The standard value of number of filters is 26. This filters are equally spaced along the Mel frequency. The filters are calculated by using following equation. This filters are applied to the periodogram estimate of power spectrum. The filter bank is in the form of 26 vectors of length 257. By multiplying each filter bank with the power spectrum and then adding the coefficient, we get 26 filter bank energies. Then the log of each of obtained energies is calculated.

5) *Discrete Cosine Transform (DCT):* We applied DCT on the log energy obtained from previous step to calculate the mel-scale cepstral coefficients. DCT of 26 log filter bank energies gives 26 Cepstral coefficients. Only lower 12-13 of the 26 coefficients are kept. The resulting feature i.e. 12 number for each frame are called Mel Frequency Cepstral Coefficients. DCT transforms signal from the frequency domain to time domain. The features obtained are called as mel-scale cepstral coefficients. The set of mel coefficients are called acoustic vectors. Hence the inputted the voiceprint is transformed into a sequence of acoustic vectors.

6) *Gaussian Mixture Model:* In Speaker Recognition System, GMM is most oftenly used to identify the class of the speaker specific feature vector. The acoustic class represents some general speaker dependent vocal tract properties which are useful for the classification of identity of speaker. The spectral shape of the acoustic class can be represented by the mean (μ) of the component density, and variations of the average spectral shape can be represented by the covariance matrix (Σ).

C. One-Time-Password System

Security of one-time password (OTP) is essential because nowadays most of the e-commerce transactions are performed with the help of this mechanism. OTP is used to overcome the replay attack/eavesdropping. Replay attack or eavesdropping is one type of attacks on network-connected computing environment. On the client side, the appropriate OTP must be displayed. On the server/host side, the server must be able to verify the OTP received from client side.

The generated OTP message is encrypted at the server side by using the DES algorithm. For the encryption we need public key, private key and secrete key. The private key of the user the hash value of the feature vector of that user. This hash value is calculated by using the SHA 256 hash function. The Elliptic Curve Diffie Hellman algorithm is then used for the key generation and key exchange. In communication channel, an encrypted OTP is transferred from the server and gets decrypted at the client side as shown in Fig. 3. After getting the plaintext of OTP code, user inputs the code in application and server verifies that OTP code.

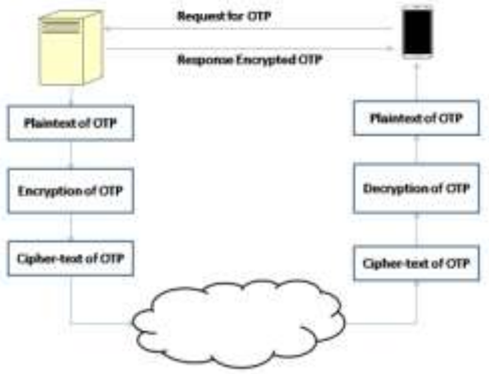

Figure 3. Implementation of OTP System

V. RESULTS AND DISCUSSIONS

This section analyses the performance assessment of proposed system in comparison with traditional system.

Traditional static text-based password technique is the most common method used by the application for authentication. Another method is graphical method or image based method. But both the techniques are nonresistant to the shoulder surfing and also difficult to remember. We have used an alternate authentication technique to mitigate the attacks mentioned above. Since noisy password technique uses different password string, it has been proved resistant to the attacks like eves dropping, key loggers, Trojan attacks etc.

In the proposed system, voices of 20 users are collected. For each user, 5 different utterances are tested, out of which two are valid utterances and three are invalid utterances. The system proposed in [24] gave 85% efficiency when 32 filters are used. When 12 Gaussian filters are used, the existing system gave 65% efficiency. Following table 1 shows the result of speaker recognition performance in the form of False Acceptance (FA) and False Rejection (FR). By analyzing the table entries we can say that the efficiency of the proposed system is 88% when 12 number of Gaussian filters are used.

Using the data values of table 6.1 we can plot the graph of performance of the GMM as shown in figure 4. We can see that, as the number of utterances increases, the performance of GMM also increases depending on the accuracy of registration process.

TABLE I. PERFORMANCE OF SRS

Speaker	No. of Attempts	False Acceptance	False Rejection
S1	5	0	0
S2	5	3	0
S3	5	2	0
S4	5	0	0
S5	5	2	0
S6	5	0	0
S7	5	0	0
S8	5	0	0
S9	5	0	1
S10	5	0	1
S11	5	0	0
S12	5	0	0
S13	5	0	0
S14	5	0	0
S15	5	0	0
S16	5	0	0
S17	5	0	0
S18	5	0	1
S19	5	0	2
S20	5	0	0
Total	100	7	5

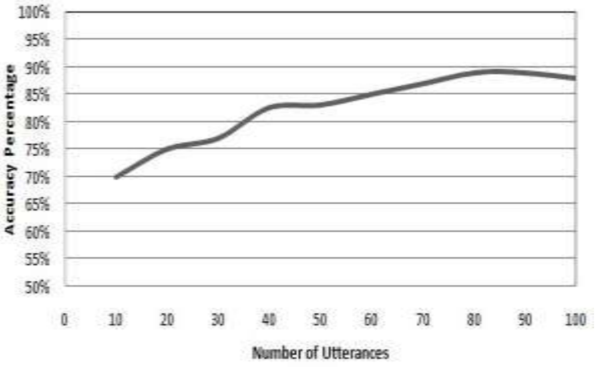

Figure 4. Performance of GMM

A. Comparison between MD5 and SHA 256

This section provides the theoretical and practical comparison between MD5 and SHA 256 hash function. Table II gives the theoretical differences between these two hash functions. The MD5 is less secure as compared to SHA 256 since the length of message digest of MD5 is less i.e. 128 bits than the SHA 256. If we try 2^{64} bit operation sequence, we can find two messages having same message digest in MD5 and 2^{80} operations are required to find two messages having same message digest in SHA 256. Since only 64 iterations are there in MD5, it is faster and less complex than SHA 256. But MD5 is reported as collision-nonresistant.

TABLE II. THEORETICAL COMPARISON BETWEEN MD5 AND SHA 256

	MD5	SHA 256
Security	Less Secure than SHA	High Secure than MD5
Length of Message Digest	128 bits	160 bits
Attacks required to find out original message	2^{128} bit operations	2^{160} bit operations
Speed	Faster	Slower than MD5
Attacks	Collision Attack	No attacks reported yet
Complexity	Less	More

Table III shows the differences between MD5 and SHA 256 occurred during execution. The same input i.e. feature vector of user is given to both of the hash functions. The message digest generated for the same file using the MD5 and SHA 256 is given in Table III. The SHA 256 took 13 cm to generate the hash value whereas MD5 hash algorithm took 9 ms to generate hash value of the same input file. Hence we can say that MD5 is less complex as compared to SHA 256. Since it has been reported that MD5 is less secured than SHA 256. A private key generated by using the SHA 256 hash function is used in cryptography while encrypting the OTP message.

TABLE III. PRACTICAL COMPARISON BETWEEN MD5 AND SHA 256

	MD5	SHA 256
Input File	Feature Vector of user A	Feature Vector of Same user A
Private Key of user (Hash value)	a4c478d05f589cbc220bae1a4ff7616f	4365401846a0ee4b60c5df7016dc00fda982bd7877c2898332b70a18a56c34c5
Processing Time	9 ms	13 ms
Complexity	Less	More

The comparison between the proposed OTP method and the traditional OTP method is given in Table IV. Though the uploading of voice sample of user takes little extra time as compared to traditional method but it ensures the security of the OTP code when it is in communication channel. Again SHA 256 hash function is adopted for the generation of keys. This helps to raise the information security.

TABLE IV. COMPARING PROPOSED OTP SCHEME WITH TRADITIONAL OTP

	Proposed OTP method	**Traditional OTP method**
OTP	Randomly Generated	Randomly Generated
Hash	Yes	No
OTP code transferred in the form	Ciphertext	Plaintext
Biometric	Yes	No

VI. CONCLUSION AND FUTURE SCOPE

Noisy password is an effective technique to overcome the shoulder surfing attack or peeping attack. For user authentication, when we combined Noisy password with Biometric technique, it produces a more secure system. There are more disadvantages than advantages for using biometric authentication systems. This is the reason why this system is not yet widely used. But the advantages are more important that we can benefit from it, and try to reduce the disadvantages. Among all the biometric procedures, voice biometric is easy to use for normal user. It is less complex as compared to other biometric techniques. Also it requires lower implementation cost. To increase the level of security, OTP mechanism is also used. But more attention is given to the security of OTP message in communication channel. The only way to enhance the security of OTP is to encrypt it using cryptography. To ensure the degree of security, a biometric is combined with the cryptography. So the proposed system enhances the drawback of the present e-commerce transaction system which uses OTP mechanism.

The text independent speaker recognition method is implemented. As a part of the future work, the plan is to implement text-dependent speaker recognition method for authentication of users. The LPC gives the best result for speech recognition but does not give high performance for speaker identification. MFCC gives the highest performance for speaker identification but does not give best result for speech recognition. Text-dependent speaker recognition involves both speech and speaker verification. The plan is to combine the functions of both LPC and MFCC to implement Text-dependent speaker recognition system.

ACKNOWLEDGMENT

I represent my sincere gratitude to Prof. K. V. Warkar, Assistant Professor, BDCOE, Sevagram, for her constantly guidance throughout the work and support.

REFERENCES

[1] D. Mahto and D. K. Yadav, “Enhancing security of one-time password using elliptic curve cryptography with biometrics for e-commerce application,” In International Conference on Computing for sustainable Global Development, Mar. 2015, pp. 1737-1742, doi:

[2] K. Alghathbar and H. A. Mahmoud, “Noisy password schme: A new one time password system,” in Candian Conference on Electrical and Computer Engineering. IEEE, May 2009, pp. 841–846.

[3] K. Alghathbar and H. Mahmoud, “Noisy password security technique,” in International Conference for Internet Technology and Secured Transaction. IEEE, Nov 2009, pp. 1–5.

[4] L. Muda, M. Begam, and I. Elamvazuthi, “Voice recognition algorithms using mel frequency cepstral coefficient (mfcc) and dynamic time warping (dtw) techniques,” Journal of Computing, vol. 2, no. 3, pp. 138–143, March 2010.

[5] K. Chakraborty, A. Talel, and S. Upadhya, “Voice recognition using mfcc algorithm,” International Journal of Innovative Research in Advanced Engineering (IJIRAE), vol. 1, no. 20, Nov. 2014, pp. 158–161,

[6] E. Kalaikavitha and J. Gnanaselvi, “Secure login using encrypted one time password (otp) and mobile based login methodology,” International Journal Of Engineering and Science, vol. 2, no. 10, April 2013, pp. 14–17.

[7] L. M. Yee and A. M. Ahmad, “Comparative study of speaker recognition methods,” University of Technology Malaysia.

[8] P. P. Singh and P. Rani, “An approach to extract feature using mfcc,” IOSR Journal of Engineering (IOSRJEN, vol. 04, no. 08, August 2014, pp. 21–25.

[9] S. B. Dhonde and S. M. Jagade, “Feature extraction techniques in speaker recognition: A review,” International Journal on Recent Technologies in Mechanical and Electrical Engineering (IJRMEE), vol. 2, no. 5, May 2015, pp. 104–106.

[10] P. Gupta and S. Kumar, “A comparative analysis of sha and md5 algorithm,” International Journal of Computer Science and Information Technology, vol. 5, no. 3, 2014, pp. 4492–4495.

[11] V. Tiwari, “Mfcc and its application in speaker recognition,” International Journal on Emerging Technologies, vol. 1, no. 1, 2010, pp. 19–22.

[12] J. A. Harris, “Opa: A one-time password system,” in Proceedings of the International Conference on Parallel Processing Workshops, 2002.

[13] E. Sediyono, K. I. Santoso, and Suhartono, “Secure login by using onetime password authentication based on md5 hash encrypted sms,” in International Conference on Advances in Computing, Communication and Informatics (ICACCI). IEEE, 2013, pp. 1604–1608.

[14] K. V. Warkar and N. J. Janwe, “A review on two level graphical authenticaton against key-logger spyware,” in National Conference on Emerging Trends in Computer Science and Information Technology (ETCSIT). Proceedings published in International Journal of Computer Applications, 2011, pp. 1–4.