User Authentication Scheme based on Voice One Time Password over Smartphone

Jinhyuck Park[a,1], Hyunsung Kim[a,2], Sung Woon Lee[b,3,*]

[a] *Dept. of Cyber Security, Kyungil University, Korea.*
[b] *Dept. of Information Security, Tongmyong University, Korea.*

Abstract

Today's smartphones are capable of doing so much more than the mobile phones of a few years ago. However, this extended range of capabilities has introduced some new security risks. This paper proposes user authentication scheme based on voice one time password, denoted as $\mathrm{AUTH}_{\mathrm{VOTP}}$, on smartphone. It is a multi-factor user authentication that uses password, smartcard, and voice. $\mathrm{AUTH}_{\mathrm{VOTP}}$ uses a voice based one time password (VOTP), which could provide more secure authentication between user and server.

Keywords: Information Security, User Authentication, Voice based One Time Password, Smartphone Security.

1. Introduction

Smartphones are becoming more and more popular due to the increase in their processing power, mobility aspect and personal nature. Smartphone can be connected various subjects, Internet, personal computer, and other mobile devices over wireless network. These features make smartphone useful and most popular mobile device. Contrast with that, these features make that attacker can invade smartphone in various means [1].

There are many research efforts on combining smartphone and voice information [2-6]. Smartphone platforms are adopted to implement a pair of mobile payment devices, both a counter reader and a paying client. So the mobile payment reader is suitably deployed in any temporary outdoor business activities, such as night markets or street fairs, because these overcrowding outdoor scenes are usually out of electric power and WiFi access point [2]. However, current research results show that there are security vulnerabilities in the mobile payment services, especially attacks against user authentication [3]. Thereby, some recent researches provide one time password based user authentication mechanisms to secure Internet banking and the mobile payment services [7-10]. User authentication mechanism based on smartphones requires additional security enhancement compared to the other authentication. Furthermore, the previous one time password (OTP) based on smartphone does not provide proper user authentication to check the ownership of that OTP [11]. Thereby, it needs further research because the mechanism is weak against the man in the

*Corresponding author
URL: http://blog.tu.ac.kr/staroun (Sung Woon Lee)
[1] Student, Dept. Cyber Security, Kyungil University, Korea.
[2] Professor, Dept. Cyber Security, Kyungil University, Korea.
[3] Professor, Dept. Information Security, Tongmyong University, Korea.

middle attack.

To cope with the problem, Cho et al. proposed voice OTP based user authentication scheme [10]. The purpose of Cho's scheme is to use voice as an additional user authentication factor but it does not provide a detailed step of user authentication by using the voice. Furthermore, it additionally needs to provide system set-up especially including registrations.

This paper proposes user authentication scheme based on voice one time password, denoted as $\text{AUTH}_{\text{VOTP}}$, on smartphone. It is a multi-factor user authentication that uses password, smartcard, and voice. $\text{AUTH}_{\text{VOTP}}$ uses a voice based OTP (VOTP), which could provide more secure authentication between user and server.

2. Background

This section reviews the network configuration of smartphone based payments and researches focused on the voice recognition [10, 12-19].

2.1. Network Configuration

Mobile financial services are among the most promising applications in the world. Smartphones allow users to access Internet using a wireless connection, to store contacts in databases and to perform payments over Internet with the network configuration of Fig. 1.

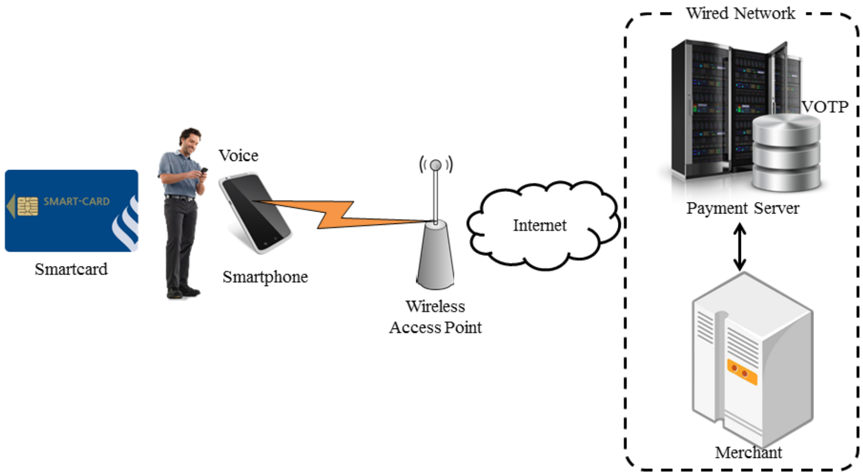

Figure 1: Network Configuration.

There are four participants in the mobile payment environment, user, smartphone, the payment server, and the service server. Smartphone could be considered the client device of a user, which needs to communicate with the payment server or the service server. Furthermore, the configuration assumes that smartphone could embed a smartcard, which could work as the universal subscriber identity module (USIM).

For the convenience, our network configuration omits the service server. Thereby, only three participants are considered in this paper, which are user, smartphone, and the payment server. Note that the server from now on will be considered as the payment server.

2.2. Voice Recognition Mechanisms

A voice biometric is a numerical representation of the sound, pattern, and rhythm of an individual's voice. A voice biometric or voice print, is as unique to an individual as a palm or fingerprint. Any application that employs a voice channel during the out-of-band session is able to add voice biometric authentication to the process for even higher levels of authentication and security [12].

Speaker dependent feature extraction from speech waves, automatic speaker identification and verification, speaker adaptation in speech recognition, and voice conversion techniques. Speaker dependent information exists both in the spectral envelope and in the supra-segmental features of speech. This individual information can be further classified into temporal and dynamic features. Speaker identification/verification methods can be divided into text-dependent and text-independent methods. Although text-dependent speaker verification techniques have almost reached the level suitable for practical implementation, text-independent techniques are still in the fundamental research stage [13].

This research will use the previous research results as voice recognition mechanism to feature user's unique biometric among one of researches in [14-19].

3. User Authentication Scheme based on VOTP

This section proposes a user authentication scheme based on voice one time password, denoted as $AUTH_{VOTP}$, on smartphone.

3.1. Registration Phase

Let x and $PU_S=g^x$ denote the server S_i's private key and its corresponding public key, where g is group generator in Z_p. Note that x is kept secret by S_i and PU_S is stored inside each user's smart card. When a user U_i wants to subscribe services from S_i, he (or she) needs to perform the following steps

Step 1 : U_i generates a random number R_A, inputs his identity ID_i, password PW_i and voice information V_i by reading numbers from 0 to 9, computes $DPW=h(PW_i \| R_A)$ and submits $\{ID_i, DPW, V_i\}$ to the registration server, R_i through a secure channel.

Step 2 : R_i computes $e_i=h(ID_i\|x)\oplus DPW_i$, where x is the server's long term secret key and stores ID_i and V_i into the database of the server. R_i stores $\{h(), PU_S, e_i\}$ on a smartcard, where $PU_S=g^x$ is the server's public key, and issues the smartcard to U_i.

Step 3 : U_i computes $R=ID_i\oplus PW_i\oplus R_A$ and stores R into the smartcard.

Note that the smartcard could be dealt with USIM of the smartphone. For the convenience, R_i considers as the server S_i.

3.2. Authentication and VOTP Generation Phase

When U_i wants to login the remote server S_i, U_i has to perform the following steps

Step 1 : U_i inputs ID_i and PW_i to smartphone. Smartphone generates a random number a, computes $DPW'=h(PW_i\|R_A)$, $Y=e_i\oplus DPW_i$, $R_U=g^a$, $R_{UA}=PU_S^a$, $CID_i=ID_i\oplus R_{UA}$ and $MAC_1=h(Y\|ID_i\|R_{UA})$ and sends the login request $\{CID_i, R_U, MAC_1\}$ to S_i.

Step 2 : S_i computes $R_{UA}'=R_U^x$, $ID_i'=CID\oplus R_{UA}'$ and $Y'=h(ID_i'\|x)$. After that, S_i checks the validity of MAC_1 by comparing it with the computed $h(Y'\|ID_i'\|R_{UA}')$. If the verification is successful, it chooses two random numbers b and c, computes $R_S=g^b$, $Chal=Y'\oplus c$, $SK=R_u^b$ and $MAC_2=h(SK\|R_U\|R_S\|c)$ and sends the reply message $\{R_S, Chal, MAC_2\}$ to smartphone.

Step 3 : After receiving the message, smartphone computes $c'=Chal\oplus Y$ and $SK'=$

R_S^a and checks the validity of MAC_2 by comparing it with the computed $h(SK'||R_U||R_S||c')$. Smartphone asks U_i to read c to generate voice information V_c, computes $VOTP_U=h(Y||SK'||R_U||R_S||c')$, $AV_c=Y\oplus V_c$ and $MAC_3=h(VOTP_U||V_c)$, and sends { AV_c, MAC_3} to S_i.

Step 4 : After receiving the message from smartphone, S_i computes $VOTP_U'=h(Y'||SK||R_U||R_S||c)$ and $V_c'=Y'\oplus AV_c$, and validates MAC_3 and V_c' by comparing MAC_3 with the computed $h(VOTP_U'||V_c')$ and by using the voice authentication check function from one of mechanisms in [14-19], respectively. Only if the verifications are successful, S_i believes the authenticity of U_i .

The generated VOTP could provide freshness, confidentiality and integrity of the session message.

4. Security Analysis

This section provides the security analysis of AUTH$_{\text{VOTP}}$. To analyze the security, we assume that an attacker can access a user's smartcard, extract the secret values stored in the smartcard by some means and intercept the messages communicating between the user and the server.

4.1. Freshness of VOTP

OTP should provide freshness to provide strong security, which could be provided either by using the session dependent random number or timestamp. AUTH$_{\text{VOTP}}$ uses user's voice biometric to generate VOTP based on a session dependent random number c and the other session dependent values R_U, R_S and SK. Thereby, AUTH$_{\text{VOTP}}$ provides freshness of $VOTP_U$. Furthermore, there is no way that an attacker guesses or knows $VOTP_U$, which is based on the difficulty of onewayness on the hash function for MAC_i in each message, the discrete logarithm problems of R_U and R_S, and the Diffie-Hellman problem of SK [20].

4.2. User Impersonation Attack

To impersonate as the legitimate user, an attacker attempts to make a forged login request message which can be authenticated to the server. However the attacker cannot impersonate as the legitimate user by forging the login request message even if the attacker can extract the secret information $\{h(),PU_S, e_i, R\}$ stored in the user's smartcard, because the attacker cannot form the login request $\{CID_i, R_U, MAC_1\}$ without knowing ID_i and computing the parameters for MAC_1 due to the lack of knowledge on Y in AUTH$_{\text{VOTP}}$. Hence, the attacker has no chance to login to AUTH$_{\text{VOTP}}$ by launching the user impersonation attack.

4.3. Server Masquerading Attack

To masquerade as the legitimate server, an attacker attempts to make the forged reply message which can be masqueraded to the user when receiving the user's login request. However the attacker cannot masquerade as the server by forging the reply message, because the attacker cannot compute $\{R_S, Chal, MAC_2\}$ without knowing the secret SK for MAC_2 due to the discrete logarithm problem and the secret Y for $Chal$. Hence, the attacker cannot masquerade as the legitimate server to the user by launching the server masquerading attack in AUTH$_{\text{VOTP}}$.

4.4. Password Guessing Attack

After the attacker extracts the secret information $\{h(),PU_S, e_i, R\}$ stored in the user's smartcard under the described assumption, the attacker attempts to derive the user's password PW_i using $e_i=h(ID_i||x)\oplus DPW_i$. However, the attacker cannot guess the user's password PW_i using the secret values extracted from the legitimate user's smartcard due to the lack of knowledge for $h(ID_i||x)$, PW_i and R_A.

4.5. Mutual Authentication

As described in the user impersonation attack and the server masquerading attack, AUTH$_{VOTP}$ can withstand the attacks, consequently AUTH$_{VOTP}$ provides mutual authentication between the user and the remote server. Namely, even if the attacker can extract the secret information $\{h(),PU_S, e_i, R\}$ stored in the user's smartcard, the user can be authenticated to the server and the server can be authenticated to the server. Because the attacker cannot form a legal login request $\{CID_i, R_U, MAC_1\}$ and a legal reply message $\{R_S, Chal, MAC_2\}$ without knowing ID_i, PW_i and $h(ID_i||x)$ and the parameters to compute MAC_1 and MAC_2.

5. Conclusion

This paper has been proposed a user authentication scheme based on VOTP, denoted as AUTH$_{VOTP}$, on smartphone. It is a multi-factor user authentication that uses password, smartcard, and voice. AUTH$_{VOTP}$ uses VOTP to provide more secure authentication between user and server. The generated VOTP in AUTH$_{VOTP}$ could provide freshness, confidentiality and integrity of the session message. The overall security analyses showed that AUTH$_{VOTP}$ achieves the desired security goals of smartphone. Thereby, AUTH$_{VOTP}$ could be used as the basic security building block for mobile payment applications over smartphone platforms.

Acknowledgments

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (MEST) (NRF-2010-0021575) and was also supported by the National Research Foundation of Korea Grant funded by the Korean Government (MEST) (NRF-2011-0008890).

References

[1] W. Jeon, J. Kim, Y. Lee, and D. Won, "A Practical Analysis of Smartphone Security," *in Lecture Notes in Computer Science*, vol. 6771, pp. 311-320, 2011.

[2] J. Hua, C. Sueng, W. Liao, and C. Ho, "Android-Based Mobile Payment Service Protected by 3-Factor Authentication and Virtual Private Ad Hoc Networking," *in Proceedings of Computing, Communications and Applications Conference*, 2012, pp. 111-116.

[3] M. L. Polla, F. Martinelli, and D. Sgandurra, "A Survey on Security for Mobile Devices," *in IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 446-471, 2013.

[4] Voice Authentication: Making Access a Figure of Speech, http://www.computerworld.com/s/article/86897/Making_access_a_figure_of_speech.

[5] Voice Verification – for Mobile Banking Security?, http://www.finextra.com/community/fullblog.aspx?id=3949.

[6] Voice PIN 2.0, http://www.voiceverified.com/products.htm.

[7] D. Choi, S. Kim, and D. Won, "Technical Analysis and Standardization Efforts on OTP: One Time Password," *in KIISC Review*, vol. 17, no. 3, pp. 12-17, 2007.

[8] K. Kim, "Remarks on Authentication System based on One Time Password," *in KIISC Review*, vol. 17, no. 3, pp. 26-31, 2007.

[9] W. Jang and H. Lee, "Biometric One-Time Password Generation Mechanism and Its Application on SIP Authentication," *in Journal of the Korea Convergence Society*, vol. 1, no. 1, pp. 93-100, 2010.

[10] S. Cho and H. Lee, "Design and Implementation of Voice One-Time Password(V-OTP) based User Authentication Mechanism on Smart Phone," *in The KIPS Transactions Part C*, vol. 18-C, pp. 79-88, 2011.

[11] C. Nickel, T. Wirtl, and C. Busch, "Authentication of Smartphone Users Based on the Way They Walk Using k-NN Algorithm," *in Proceedings of Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2012, pp. 16-20.

[12] http://www.authentify.com/solutions/voice_biometrics.html.

[13] S. Furui, "Speaker-dependent-feature extraction, recognition and processing techniques," *in Speaker Characterization in Speech Technology*, vol. 10, no. 5-6, pp. 505-520, 1991.

[14] R. L. Klevans and R. D. Rodman, *Voice Recognition*, Artech House, 1997.

[15] C. Kim and R. M. Stern, "Feature Extraction for Robust Speech Recognition based on Maximizing the Sharpness of the Power Distribution and on Power Flooring," *in Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing*, 2010, pp. 4571-4577.

[16] B. Singh, R. Kaur, N. Devgun, and R. Kaur, "Review of Feature Extraction Techniques in Automatic Speech Recognition,*" in International Journal of Scientific Engineering and Technology*, vol. 2, no. 6, pp. 479-484, 2013.

[17] A. G. Ghitu, L. J. M. Rothkrantz, P. Wiggers, and J. C. Wojdel, "Comparison between different feature extraction techniques for audio-visual speech recognition," *in Journal of Multimodal User Interfaces*, vol. 1, no. 1, pp. 7-20, 2007.

[18] Z. Tuske, P. Golik, R. Schluter, and F. R. Drepper, "Non-Stationary Feature Extraction for Automatic Speech Recognition," *in Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing*, 2011, pp. 5204-5207.

[19] S. Demircan and H. Kahramanli, "Feature Extraction from Speech Data for Emotion Recognition," *in Journal of Advances in Computer Networks*, vol. 2, no. 1, pp. 28-30, 2014.

[20] B. Schneier, *Applied Cryptography*, Wiley, 1994.