explicit-consent and defining requiring-explicit-consent as the ability of confirming explicit intentions under denial or man-in-the-middle attacks, as mentioned above.

3. Architecture

In this section, we describe our two proposed methods, Two-Factor User Authentication with SMS and Voiceprint Challenge Response (SV-2FA) and SV-2FA/Certificate Provisioning (SV-2FA/CP), by using a netbanking model. These methods are based on two-factor user authentication with SMS and voiceprint challenge response. The former is a basic method, and the later is a revised method from the viewpoint of usability and communication cost by using client certificates.

3.1 SV-2FA method

3.1.1 Setup

- **PC**

  A general terminal, such as a PC/tablet, of a legitimate user (in the following example Alice), can be equivalent to a cellphone. In this case, this authentication method can be used for cellphones.

- **Cellphone**

  A general cellphone with caller ID satisfies the assumption given in Section 2 and has voice communication and SMS functionality. Installing special applications is not needed. However, softphones, such as Skype, are excluded because they can be logged into with passwords. Such software cannot maintain the possession authentication’s security level because it can be spoofed when a password is stolen.

- **Server**

  A server is connected to the PC through the Internet and the cellphone through the telephone network (Fig. 3). The server receives a service request from the PC and sends a user authentication request. If the server succeeds in user authentication, it provides a service to the PC. The server has functions such as WEB server, interactive voice response (IVR), voice recognition, voiceprint determination, SMS transmission, and database. The server has enough telephone numbers, and the IVR can recognize which number is called. The more phone numbers there are, the more secure it becomes. If a user does not have many phone numbers due to the cost, he/she can regularly delete and add phone numbers. To prevent the line from becoming busy, it is preferable to have about ten lines.

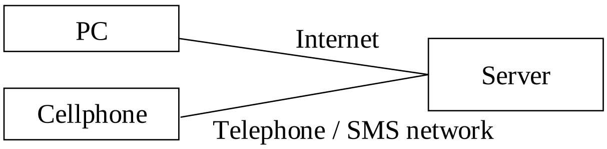
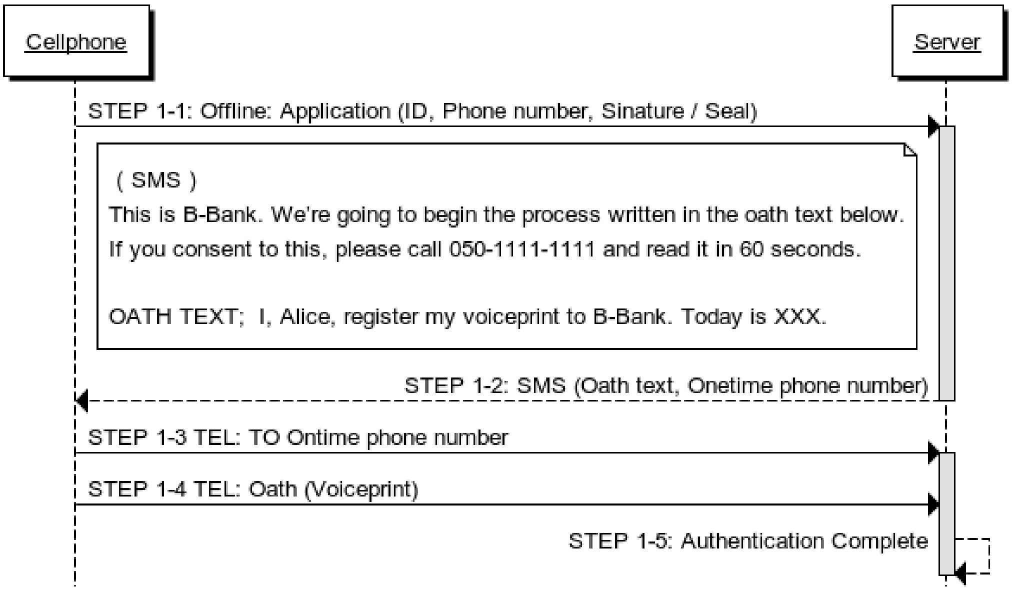


Figure 3. System Structure

3.1.2 Registraton/Changing phase

This section describes how to register an ID, voiceprint, and phone number (Fig. 2). However, in



STEP 1-1: Offline: Application (ID, Phone number, Sinature / Seal)

(SMS)
This is B-Bank. We're going to begin the process written in the oath text below.
If you consent to this, please call 050-1111-1111 and read it in 60 seconds.

OATH TEXT; I, Alice, register my voiceprint to B-Bank. Today is XXX.

STEP 1-2: SMS (Oath text, Onetime phone number)

STEP 1-3 TEL: TO Ontime phone number

STEP 1-4 TEL: Oath (Voiceprint)

STEP 1-5: Authentication Complete

Figure 2. Registration/Changing Phase.

principle, this method does not require passwords. Changing the phone number must not be allowed online because the former phone number has been unused, so the user cannot receive the SMS.

- **STEP 1-1 Authentication Request**

  A user writes his/her ID, cellphone number, and signature or seal on a form and mails the information to the server's operator (in this example, B-Bank). This phase is secure due to the assumption discussed in Section 2.

- **STEP 1-2 SMS**

  A server sends an SMS that includes the oath text and onetime phone number to the cellphone. The onetime phone number changes according to the session.

  An example of the body of such an SMS is as follows. "This is B-Bank. We're going to begin the process written in the oath text below. If you consent to this, please call 050-1111-1111 and read it in 60 seconds. OATH TEXT; I, Alice, register my voiceprint to B-Bank. Today is XXX."

  The oath should include text that changes every time, such as the date, to increase the accuracy of the voiceprint registration. By changing the oath text every time, replay attacks that record and use legitimate users voices illegally can be prevented.

- **STEP 1-3 Voice call**

  The user calls the onetime phone number by clicking on it in the SMS, and the server responds to the incoming call. It is also allowed to make the server not answer any call from a phone number that is not sent in STEP1-2. Therefore, it can prevent dictionary attacks in which an attacker calls the phone number received in the past, or prevent DOS attacks in which an attacker calls many numbers at a time to make the lines busy. However, to prevent DOS attacks, just not answering the call is not enough because ringing status makes the line busy. Therefore, the server should send a call disconnecting message before answering.

- **STEP 1-4 Oath and Voiceprint Recording**

  If voice communication starts, the user reads the oath, and the server checks if the user read the oath correctly by using voice recognition.

- **STEP 1-5 Authentication Complete**

  The server records the oath, voiceprint date, ID, and phone number.

3.1.3 Login Phase

This section describes the process for each login (Fig. 4).

- **STEP 2-1 Authentication Request**

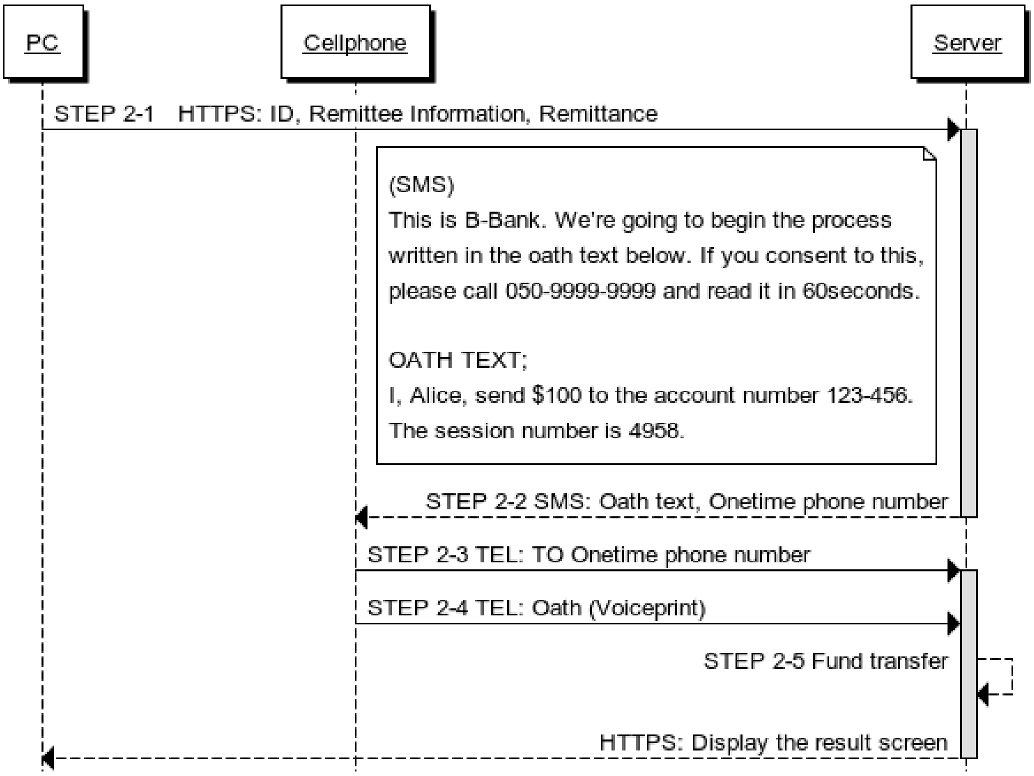

PC
Cellphone
Server
STEP 2-1 HTTPS: ID, Remittee Information, Remittance
(SMS)
This is B-Bank. We're going to begin the process written in the oath text below. If you consent to this, please call 050-9999-9999 and read it in 60seconds.
OATH TEXT;
I, Alice, send $100 to the account number 123-456.
The session number is 4958.
STEP 2-2 SMS: Oath text, Onetime phone number
STEP 2-3 TEL: TO Onetime phone number
STEP 2-4 TEL: Oath (Voiceprint)
STEP 2-5 Fund transfer
HTTPS: Display the result screen

Figure 4. Login phase

The PC sends the ID, remittee information, and remittance to the server (it does not need to be one transaction). At this time, a session number is displayed on the PC (for example 4985).

- **STEP 2-2 SMS**

  The server sends basically the same SMS as in the registration phase to a cellphone. The oath text is below. Since this phase is for voiceprint authentication not for registration, the oath text can be relatively short.

  Oath text example: “I, Alice, send $Y to the account number X. The session number is 4958.”

- **STEP 2-3 Voice Call**

  The user clicks on and calls the onetime phone number in the SMS’s body; the same as in the registration phase.

- **STEP 2-4 Voiceprint authentication and Oath Recording**

  This is almost the same as in the registration phase. However, the difference is that this phase is for voiceprint authentication not registration.

- **STEP 2-5 Authentication Complete:** Fund transfer process, saving the record, and displaying the result screen are obeyed.

3.2 SV-2FA/CP method

The SV-2FA method has a problem from the viewpoint of usability and communication cost because it uses SMS, voice call, and oath recording every time. In this section, we propose SV-2FA/CP, which solves these problems by using a client certificate. This method adds a certificate provisioning phase, which is almost the same as SV-2FA’s login phase, but STEP2-5 was changed to issue a certificate. After that, the server authenticates the user by the certificate. This method also improves upon the conventional commercial certificate authority’s (CA’s) operation, which involves e-mail and passwords. The registration/changing phase is the same as with SV-2FA. However, we omit the same words and phrases if the STEP is almost the same as that with SV-2FA.

3.2.1 Setup

The following is an addition to SV-2FA’s Setup.

- **Certificate Authority (CA)**

This generates a public key pair for the client and issues the corresponding client certificate.

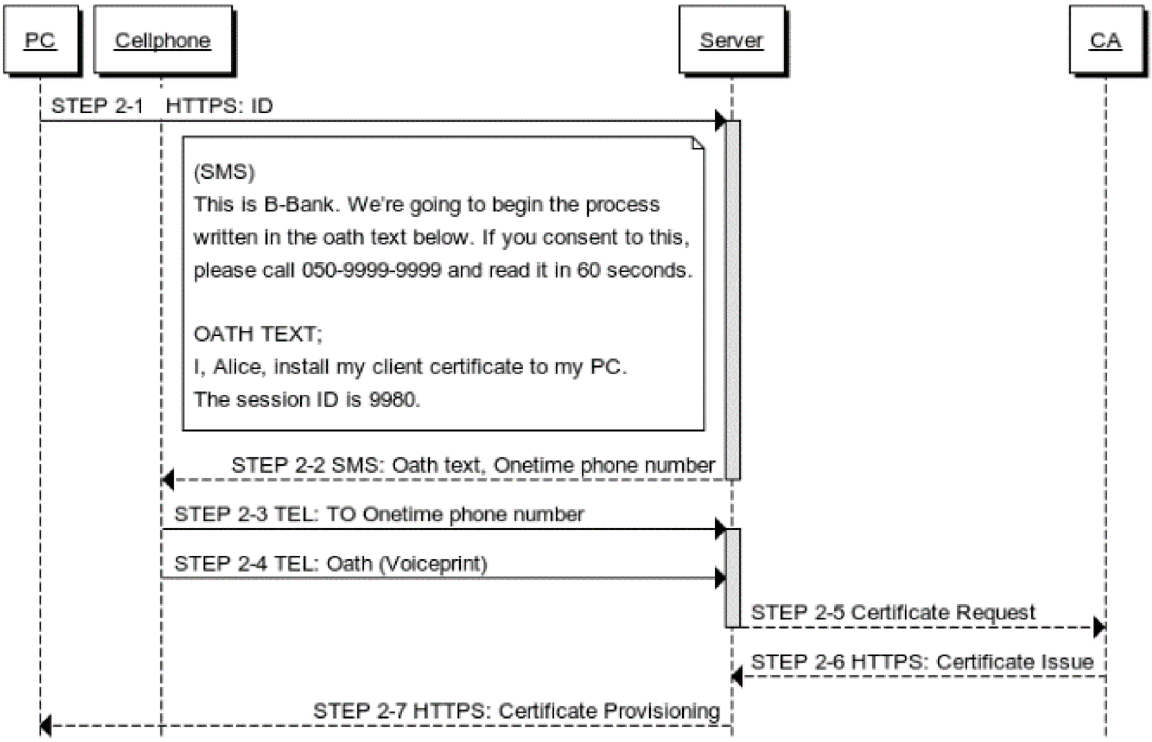
3.2.2 Registraton/Changing phase

All steps are same as SV-2FA.

3.2.3 Certificate provisioning phase

This section explains the certificate provisioning phase (Fig. 5). We omit the same words and phrases if the STEP is the same as that in SV-2FA’s Login phase.

- **STEP 2-2 SMS**



PC
Cellphone
Server
CA
STEP 2-1 HTTPS: ID
STEP 2-2 SMS: Oath text, Onetime phone number
STEP 2-3 TEL: TO Onetime phone number
STEP 2-4 TEL: Oath (Voiceprint)
STEP 2-5 Certificate Request
STEP 2-6 HTTPS: Certificate Issue
STEP 2-7 HTTPS: Certificate Provisioning

(SMS)
This is B-Bank. We're going to begin the process written in the oath text below. If you consent to this, please call 050-9999-9999 and read it in 60 seconds.

OATH TEXT;
I, Alice, install my client certificate to my PC.
The session ID is 9980.

Figure 5. Client Certificate Provisioning Phase

The oath text is as follows.

Oath text example: "I, Alice, install my client certificate to the PC that has session number 9980".

- **STEP 2-5 Certificate Request**

  The server requests the CA to generate the public key pair and issues the client certificate.

- **STEP 2-6 Certificate Issue**

  The CA generates the public key pair for the client, issues the corresponding client certificate, and sends them to the server.

- **STEP 2-7 Certificate Provisioning**

  The server sends the public key pair and client certificate to the PC through HTTPS.

With this method, the CA generates a public key pair. This kind of operation is often seen with some commercial CAs. However, it is also allowed to generate a key pair through a PC. If the PC is a smartphone, certificate provisioning is done in STEP 1-5.

3.2.4 Login Phase

Although there are some differences to the principle method, usability and communication cost problems are solved by the client certificate.

- **STEP 3-1 First Authentication**

  The server authenticates by the client certificate. At this stage, balance checking, small amount of remitting a small amount, or remitting to a registered account are executable.

- **STEP 3-2 Second Authentication**

  This step is required only when the remittance is above a certain amount or sent to an unregistered account.

  Following Steps are same as SV-2FA's Login Phase.

4. Evaluation

In this section, we evaluate our methods against Bonneau et al.'s UDS Evaluation and our added or revised benefits defined in Section 2 (Table 1). We do not describe the original definition of these benefits due to space limitations. Please read the original paper [15]. However, we compare our methods with "Google 2-Step," and "Voice", which are described in Bonneau et al.'s representative examples, which are similar to our methods.

4.1 SV-2FA

4.1.1 Usability

Since our method does not require any information to memorize, it is memory-wise-effortless and scalable-for-users (Bonneau et al. argued that Google 2-Step does not have this benefit since it requires password authentication). It is quasi-nothing-to-carry since one cellphone can be used for any site (same as Google 2-Step). It is quasi-physically-effortless and easy-to-learn since it requires receiving only SMS and reading of the oath (equal to Voice). It is quasi-efficient-to-use since there is a delay in receiving an SMS and in reading the oath. It is not infrequent-errors since the false rejection rate of voiceprint authentication is several percentage points [20]. However, it takes time to replace a cellphone, so it is quasi-easy-recovery-from-loss since storing secret information in the cellphone is not required.

4.1.2 Deployability

The SV-2FA method is quasi-accessible since a blind person cannot read the oath text (there is another way to communicate the oath text through voice guidance). It is not negligible-cost-per-user by taking into account the communication charge of SMSs and voice calls. It is not server-compatible since alternating the server is required, and it is browser-compatible since installations, such as plug-ins, in the browser are not required. Since we have commercialized products that have voice recognition, voiceprint authentication, and telephone authentication functionality [8], there remains only to add SMS-sending and record functionality, making it quasi-mature.

4.1.3 Security

The SV-2FA method is resilient-to-physical-observation, resilient-to-internal-observation, and resilient-to-theft because an attacker cannot send a correct voiceprint if he/she eavesdrops on the oath text or SMS or hijacks the voice call and SMS communication by infecting the PC or cellphone with malware. It is resilient-to-targeted-impersonation since an attacker cannot reuse the voiceprint, which is recorded illegally from the legitimate user, because the oath text changes each time. If the representative voiceprint authentication method's entropy is converted into key space, corresponding to no more than 11.7 bits, the voiceprint cannot be resilient-to-unthrottled-guessing alone [15, page 24]. However, recent research suggests that the cross over error rate is 6.4% [20], furthermore, our method includes possession authentication by SMS; therefore, it can be resilient-to-unthrottled-guessing, as with Google 2-Step. Bonneau et al. stated that biometrics is not resilient-to-leaks-from-other-verifier but our method changes the oath text each time and requires possession authentication by SMS. Therefore, it is resilient-to-leaks-from-other-verifier. It is resilient-to-phishing because of possession authentication. It is not no-trusted-third-parties because it requires a trusted mobile operator. It is requiring-explicit-consent (Revised), which we defined in Section 2, because it can confirm explicit intention with the oath. It is not unlinkable due to voiceprint authentication. It is resilient-to-man-in-the-middle-attack (New), as defined in Section 2. If an attacker conducts a man-in-the-middle attack through the client certificate and changes the remittee information, remittance, and result screen, the legitimate user can be notified of the attack by SMS and stop the remittance.

4.2 SV-2FA/CP

Most benefits are the same as with SV-2FA. This method's "Usability and Deployability" and "Safety" are in a trade-off relationship, and the valance can be changed by changing the setting of STEP 3-2. We describe only the difference in SV-2FA below.

4.2.1 Usability

Although this method requires SMS and voice call at registration, the remittance to be above a certain amount, and so on, usual operation is done by the client certificate seamlessly, and this improves the principle method. Therefore, it is physically-effortless, efficient-to-us, and quasi-infrequent-errors.

4.2.2 Deployability

Because the numbers of SMSs and voice calls decrease, it is quasi-negligible-cost-per-user.

4.2.3 Security

This method cannot prevent illegal access when the PC that has the client certificate is stolen. However, an unregistered account's remittance cannot be obeyed, so it is quasi-resilient-to-theft. It is resilient-to-phishing because even a small amount of remittance uses the client certificate. It is quasi-requiring-explicit-consent (Revised) because this method does not require oath recording when a small amount of remittance is obeyed. It is resilient-to-man-in-the-middle-attack (New) because the server holds the legitimate user's public key and can check if the PC is of the legitimate user from the client certificate.

5. Conclusion

We proposed two strict user authentication methods that use SMS and recording oath and voiceprint authentication through voice calls. These methods enable possession factor authentication by calling back to the one-time phone number sent by SMS and prevent man-in-the-middle attacks and legitimate user denial by recording a user's oath reading. At the same time, they prevent spoofing by voiceprint authentication, even when a cellphone is stolen, and prevent replay attacks perpetrated with illegal voice recordings by changing the oath text every time. By calling back, the methods can prevent social-engineering attacks perpetrated by changing the remote call transfer setting. Our methods are easy to use since they can be used with any cellphone. However, the former version of these methods, which uses caller ID authentication and voiceprint authentication, has been commercialized [8], so we also plan to improve this method.

TABLE 1. COMPARISON OF SV-2FA AGAINST PASSWORDS AND GOOGLE 2-STEP VERIFICATION AND (SIMPLE) VOICEPRINT USING BONNEAU ET AL.'S EVALUATION FRAMEWORK

	Usability								**Deployability**						**Security**											
Scheme	***Memorywise-Effortless***	***Scalable-for-Users***	***Nothing-to-Carry***	***Physically-Effortless***	***Easy-to-Learn***	***Efficient-to-Use***	***Infrequent-Errors***	***Easy-Recovery-from-Loss***	***Accessible***	***Negligible-Cost-per-User***	***Server-Compatible***	***Browser-Compatible***	***Mature***	***Non-Proprietary***	***Resilient-to-Physical-Observation***	***Resilient-to-Targeted-Impersonation***	***Resilient-to-Throttled-Guessing***	***Resilient-to-Unthrotled-Guessing***	***Resilient-to-Internal-Observation***	***Resilient-to-Leaks-from-Other-Verifiers***	***Resilient-to-Phishing***	***Resilient-to-Theft***	***No-Trusted-Third-Party***	***Requiring-Explicit-Consent (Revised)***	***Unlinkable***	***Resilient-to-Man-in-the-Middle-Attack (New)***
SV-2FA	●	●	○	○	●	○		○	○			●	○		●	●	●	●	●	●	●	●		●		●
SV-2FA/CP	●	●	○	●	●	●	○	○	○	○		●	○		●	●	●	●	●	●	●	○		○		●
Passwords			●		●	●	○	●	●	●	●	●	●	●		○							●	○	●	
Google 2-Step			○		●	○	○	○	○			●	●		○	○	●	●		●	●	●	●	○	●	

Voice	●	●	●	○	●	○			○	○		○	○		●		○							○		

●=offers benefit; ○= offers some benefit; no circle= does not offer benefit

6. References

[1] Frequently Asked Questions on FFIEC Guidance on Authentication in an Internet Banking Environment, August 15, 2006.

[2] J. Bonneau, C. Herley, P. C. van Oorschot, and F. Stajano, "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes," IEEE Symposium on Security and Privacy (SP), 2012.

[3] B. Schneier, “Two-Factor Authentication: Too Little, Too Late,” COMMUNICATIONS OF THE ACM, Vol. 48, No. 4, pp. 27–27 , 2005.

[4] Google Inc., “2-step verification: how it works,” 2012, www.google.com/accounts (accessed 2013-8-1).

[5] H. Fujii, Y. Tsuruoka, and Y. Tada, “A User Authentication Scheme with Calling Number Notification of Telephone Network, ” IPSJ Journal, Vol.54, No.2, pp. 992-1001, 2013, (in Japanese).

[6] H. Fujii, and H. Takei, N. Miyake, and E. Kuwana, Japan Patent 3497799 (2001-12-21).

[7] H. Fujii, N. Shigematsu, H. Kurokawa, and T. Nakagawa, “Telelogin: a two-factor two-path authentication Technique Using Caller ID,” NTT Technical review, Vol. 6, No. 8, pp. 1-6, 2008.

[8] NTT Software Corporation: CallPassport[online]. Available: http://www.ntts.co.jp/products/callpassport (accessed 2013-8-1).

[9] H. Sun, Y. Chen, and Y. Lin, "oPass: A User Authentication Protocol Resistant to Password Stealing and Password Reuse Attack," IEEE Transactions on Information Forensics and Security, vol. 7.2, pp. 651-663, 2012.

[10] B. Parno, C. kuo, and A. Perrinng, "Phoolproof phishing prevention," Financial Cryptography and Data Security, pp. 1-19, 2006.

[11] S. Suoranta, A. Andrade, and T. Aura, "Strong Authentication with Mobile Phone," Information Security, pp. 70-85, 2012.

[12] L. O’Gorman, “Comparing passwords, tokens, and biometrics for user authentication,” Proc. IEEE, vol. 91, no. 12, pp. 2021–2040, 2003.

[13] X. Huang, Y Xiang, A. Chonka, J. Zhou, and R. H. Deng, "A generic framework for three-factor authentication: preserving security and privacy in distributed systems," IEEE Transactions on Parallel and Distributed Systems, vol. 22.8, pp. 1390-1397, 2011.

[14] H. Fujii, and Y. Tsuruoka, “Three-Factor User Authentication Method Using Biometrics Challenge Response,“ FC 2013. LNCS, vol. 7859, pp 395–396, Springer, Heidelberg, 2013.

[15] J. Bonneau, C. Herley, P. C. van Oorshot, and F. Stajano, "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes," Technical reports published by the University of Cambridge Computer Laboratory ISSN 1476-2986, 2012.

[16] TS 23.040: Technical Realization Short Message Service (SMS) 3GPP [Online]. Available: http://www.3gpp.org/ (accessed 2013-8-1).

[17] European Telecommunications Standards Institute: ETSI TS 133 102 V10.0.0 Universal Mobile Telecommunications System (UMTS); LTE; 3G security; Security architecture (3GPP TS 33.102 version 10.0.0 Release 10), 2011.

[18] A. Bais, W. T. Penzhorn, and P. Palensky, “Evaluation of UMTS security architecture and services,” IEEE International Conference on Industrial Informatics, pp. 570-575, 2006.

[19] L. O’Gorman, “Comparing passwords, tokens, and biometrics for user authentication,” Proceedings of the IEEE, vol. 91, no. 12, pp.2019–2040, 2003.

[20] N. Tsuchiya, Y. Yamaguchi, Y. Fukumoto, H. Iwano, and S. Furui, “Speech Identification in VoIP (Voice over IP) System,” Symposium on Mobile Interactions and Navigation, 2004/3/17-18. 2004, (in Japanese).