Security in the Password-based Identification

Byung-Jun Park and Jong-Min Park, *Member, KIMICS*

Abstract—Almost all network systems provide an authentication mechanism based on user ID and password. In such system, it is easy to obtain the user password using a sniffer program with illegal eavesdropping. The one-time password and challenge-response method are useful authentication schemes that protect the user passwords against eavesdropping. In client/server environments, the one-time password scheme using time is especially useful because it solves the synchronization problem. It is the stability that is based on Square Root Problem, and we would like to suggest PBSI(Password Based Secure Identification), enhancing the stability, for all of the well-known attacks by now including Off-line dictionary attack, password file compromise, Server and so on. The PBSI is also excellent in the aspect of the performance.

Index Term—Password Based Secure Identification, one-time password, Off-line dictionary attack, password file compromise

I. INTRODUCTION

Authentication is the process whereby a verifier is assured of the identify of a prover involved in a protocol, and that the prover has actually participated. Identification is a process whereby a verifier is assured that the identity of a prover is as declared, thereby preventing impersonation.

The password system is most widely used authentication scheme because of its advantages including easy implementation, low price and usability, and the attacks which must be guarded in the password system include password disclosure at the outside of the system and line eavesdropping within the system, both of which allow subsequent replay and password guessing including off-line dictionary attacks [12,13].

Several cryptographic techniques such as one time password [9] and salting technique [10] have been proposed for enhancing the security of the password system, but, in spite of endeavor as yet, no identification technique processing the password has been presented.

To overcome the weakness of the password system, several challenge response protocols [2,6,7] and zero knowledge identification protocols have been presented [4,5,8], but the protocols do not process human memorable password, some of the protocols requires the trusted third party, and the protocols are not completed with one pass.

This paper introduces a new cryptographic identification protocol, called Password Based Secure Identification-(PBSI), processing the human memorable password, and before we present the PBSI, we introduce two identification techniques that are related to the PBSI.

The security of the PBSI depends on the square root problem., that is, the security of the PBSI depends on the fact that if n is the product of two primes, then the ability to calculate square roots mod n is computationally equivalent to the ability to factor n [11]. The PBSI is secure against the well known attacks such as replay attack, pre-play attack, man-in-the-middle attack, eavesdropping, off-line dictionary attack, password file compromise, and furthermore secure against the password file compromise with having performed off-line dictionary attack [12,13].

Comparing the PBSI with challenge response identification protocols and zero knowledge identifi-cation [14] protocols, the PBSI processes the human memorable password, and a number of pass of the PBSI is one.

A. Notation

Our notation is shown in the following :

P : A password of a user.

X_1, Y_1, X_2, Y_2 : Integers such that $(X_1 + X_2) \bmod N \equiv P$ and $(Y_1 + Y_2) \bmod N \equiv P$

N: N = pq where p and q are primes such that N is computationally infeasible to factor.

E_K : Symmetric encryption with key K.

D_K : Symmetric decryption with key K.

R: Random integer.

H: One way hash function.

B. Security of Identification

We present a list of basic attacks that authenticated key exchange protocol needs to guard against.

• Replay :

The attacker records messages which were sent in past communications and re-sends them at a later time.

Manuscript received November 28, 2007.

Byung-Jun Park is with Chosun University, Kwang-Ju, Republic of Korea (e-mail : pbj5446@hanmail.net)

Jong-Min Park is with Chosun University, Kwang-Ju, Republic of Korea (e-mail : pjm5234@hanmail.net)

• Pre-play :
The attacker records messages which were sent in past communications and determines a message from the recorded messages for current communication.

• Eavesdropping :
The attacker listens messages on the line and tries to learn some useful information from the ongoing communication.

• Man-in-the-middle :
The attacker intercepts the messages sent between the parties and replaces them with its own messages.

• Password guessing attacks :
The attacker is assumed to have access to a relatively small dictionary containing common choices of password. there are primarily two ways in which the attacker can use the dictionary that are on-line dictionary attack and off-line dictionary attack.

• Off-line dictionary attack :
The attacker records past communication, and then goes over the dictionary and looks for a password which is consistent with the recorded communication. If such a password is found, the attacker concludes that this is the password of the attack.

• On-line dictionary attack :
The attacker repeatedly picks a password from the dictionary and tries to use it in order to impersonate as the user. If the impersonation fails, the attacker eliminates this password from the dictionary and tries again, using a different password. The standard ways of preventing such on line dictionary attack in practice are to either limit the number of failed runs that a user is allowed to have before the password is expired, or reduce the rate in which the user is allowed to make login attempts. For this reason, in this paper, we consider only off line dictionary attack.

• Password file compromise :
The attacker gets access to sensitive data which is supposed to kept secret at the password file to masquerade as a user.

• Server compromise is possible to identification scheme :
The verifier can impersonate the prover.

II. PRELIMINARIES

A. Challenge response identification

A disadvantage of simple password protocols is that when a claimant *A* gives the verifier *B* her password, *B* can thereafter impersonate *A*. Challenge response protocols improve on this: *A* responds to *B*'s challenge to demonstrate knowledge of *A*'s secret in a time vatiant manner, providing information not directly reusable by *B*.

1)Based on symmetric key block cipher

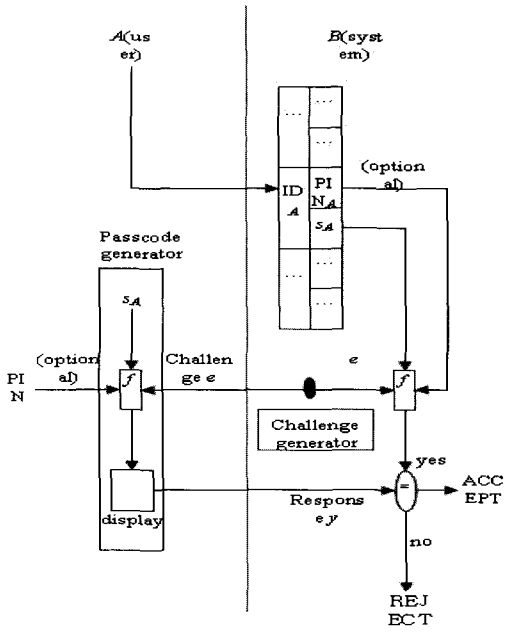

Fig. 1 Challenge response(SKBC)

2)Based on zero-knowledge proof

Zero knowledge protocols are designed to address that allowing a prover to demonstrate knowledge of a secret while revealing no information whatsoever of use to the verifier in conveying tjis demonstration of knowledge to others.

A proves knowledge of s to *B* in t executions of a 3-pass protocol

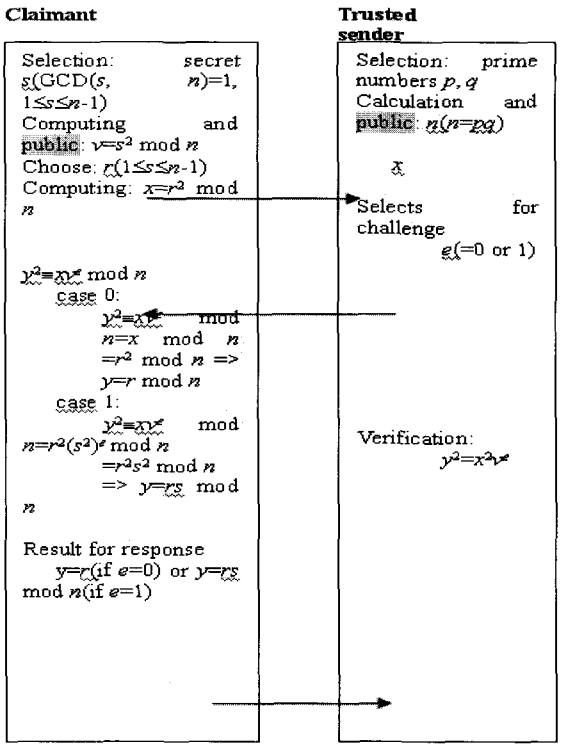

Fig. 2 ZK(Fiat-Shamir identification Protocol)

B. NP-complete

The square root modulo n (SQROOT) problem is to find a square root of a module n for the given composite integer n and quadratic residue a modulo n. If the factors p and q are known, then SQROOT problem can be solved in polynomial time. If the factors p and q are unknown, then the factoring problem of n is reduced to SQROOT problem in polynomial time [11], and the factoring problem of n is NP - complete [1, 3].

Property Let $n=pq$, and two primes p and q be selected such that n is computationally infeasible to factor. Then, the problem finding x in $(x+t)^2 \bmod n$, for a given t, quadratic residue a modulo n and n, is NP-complete.

III. PASSWORD BASED SECURE IDENTIFICATION

We first introduce an identification technique in Fig. 3 which well describes how identification scheme is to be designed so that secure against off line dictionary attack.

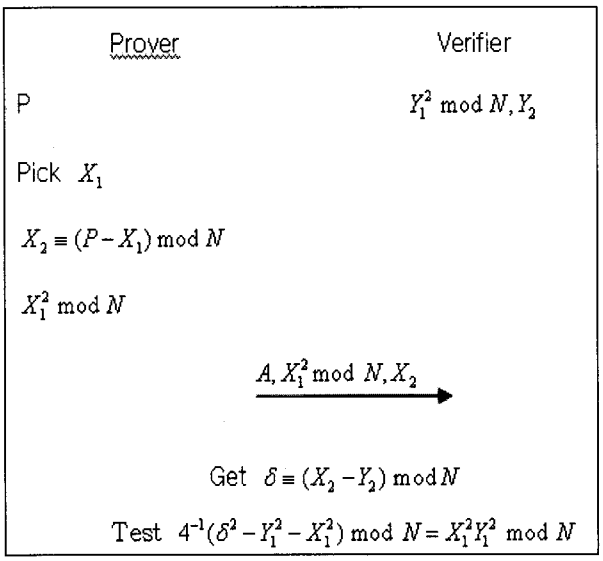

Fig. 3 The first identification technique

In the identification technique described at above, when a user registers P with his identity A, the verifier chooses Y_1 in random ($0 \le Y_1 \le N-1$) and determines Y_2 such that Y_2 such that $Y_2 \equiv (P-Y_1) \bmod N$, and then stores $Y_1^2 \bmod N$ and Y_2 at A of the password file.

When the user inputs P with his identity A, the verifier chooses X_1 in random ($0 \le X_1 \le N-1$) and determines X_2 such that $X_2 \equiv (P-X_1) \bmod N$, and then sends A, $X_1^2 \bmod N$ and X_2 to the verifier. Let $\delta=(X_2-Y_2) \bmod N$, the verifier knows $(Y_1-X_1)^2 \bmod N \equiv \delta^2 \bmod N$ because of $(X_1+X_2) \bmod N \equiv (Y_1+Y_2) \bmod N$. Therefore, the verifier can determine whether $X_1^2 Y_1^2 \bmod N \equiv (4^{-1}(Y_1^2+X_1^2-\delta^2)^2) \bmod N$ is hold or not by using $X_1^2 \bmod N$ and X_2 received from the prover, and $Y_1^2 \bmod N$ and Y_2 stored at the password file.

The technique is secure against replay attack, because X_1 is chosen in random. The technique is also secure against off-line dictionary attack, because it is too huge to store 2^N candidates for $X_1^2 \bmod N$. The technique is also secure against just password file compromise, because it is computationally infeasible to find X_1 from $X_1^2 \bmod N$ even if the password file compromise, because it is computationally infeasible to find X_1 from $X_1^2 \bmod N$ even if the password file is compromised. The technique is also secure against password file compromise with having performed off line dictionary attack, because it is computationally infeasible to find X_1 from $X_1^2 \bmod N$, and X_1 is chosen in random.

Let A, $X_{11}^2 \bmod N$ and X_{21} and A, $X_{12}^2 \bmod N$ and X_{22} be two messages sent to the verifier, then the attacker eavesdropped the messages in past communications can determine X_{11}^2 or X_{12} by using $X_{11}^2 \bmod N \equiv (X_{12}+\varepsilon)^2 \bmod N$ where $\varepsilon = X_{22} - X_{21}$. Therefore, the technique is vulnerable to eavesdropping, pre-play attack and man-in-the-middle attack.

The second identification technique in Fig. 4 describes how identification scheme is to be designed so that secure against eavesdropping.

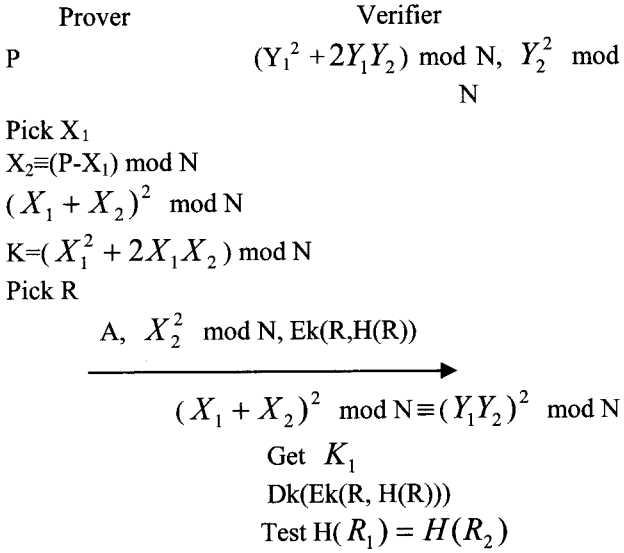

Fig. 4 The second identification technique

In the second identification technique, when a user registers P with his identity A, the verifier chooses Y_1 in random ($0 \le Y_1 \le N-1$) and determines Y_2 such that $Y_2 \equiv (P-Y_1) \bmod N$, and then stores $(Y_1^2+2Y_1Y_2) \bmod N$ and $Y_2^2 \bmod N$ at A of the password file.

When the user inputs P with his identity A, the verifier chooses X_1 and R_1 in random ($0 \le X_1 \le N-1$) and determines X_2 such that $X_2 \equiv (P-X_1) \bmod N$, and then sends A, $X_2^2 \bmod N$, $Ek_1(R, H(R))$ to the verifier where $k_1=(X_1^2+2X_1X_2) \bmod N$. the verifier can determine K_1 by using the fact that $(X_1+X_2)^2 \bmod N \equiv (Y_1+Y_2)^2 \bmod N$. The verifier performs $Dk(Ek(R, H(R)))$ and then tests whether $H(R_1)=H(R_2)$ is hold or not where $(R_1, H(R_2))=Dk(Ek(R, H(R)))$. The verifier accepts the prover only when $H(R_1)=H(R_2)$ is hold.

We can see that above property is true, because the problem finding a square root of a modulo n for the given composite integer n and quadratic residue a modulo n is a special case of the problem finding x in $(x+t)^2$ mod n for a given t, quadratic residue a modulo n and n.

The technique is secure against replay attack, because X_1 is chosen in random. From the property, it is computationally infeasible to find X_1 in $(X_1^2+2X_1X_2)$ mod N under weak environments for the protocol that X_2 in X_2^2 mod N can be found in reasonable time heuristically, and the symmetric cryptosystem is vulnerable to ciphertext only attack. Therefore, the technique is secure against pre-play, eavesdropping and man-in-the-middle.

The technique is secure against off-line dictionary attack, because it is too huge to store all candidates for $E_k(R,H(R))$.

The technique is secure against password file compromise without having performed off-line dictionary attack, because the password was not stored at the password file in clear text.

The technique is vulnerable to password file compromise with having performed off-line dictionary attack because of P^2 mod $N\equiv(Y_1^2+2Y_1Y_2+Y_2^2)$ mod N. The PBSI described in Fig. 5.

Prover — Verifier

P — Y_1^2 mod N, Y_2^2 mod N

Pick X_1

$X_2\equiv(P-X_1)$ mod N

X_1^2 mod N, X_2^2 mod N

A, X_1^2 mod N, X_2^2 mod N

→

Test $X_1^2X_2^2Y_1^2Y_2^2$ mod N = C mod N

Fig. 5 PBSI

In the PBSI, $C=64^{-1}(Y_1^2+Y_2^2 - X_1^2 - X_2^2)^2 - 4X_1^2X_2^2 - 4Y_1^2Y_2^2)^2$. When a user registers P with his identity A, the verifier chooses Y_1 in random $(0\le Y_1\le N-1)$ and determines Y_2 such that $Y_2\equiv(P-Y_1)$ mod N, and then stores Y_1^2 mod N and Y_2^2 mod N at A of the password file.

When the user inputs P with his identity A, the verifier chooses X_1 in random $(0\le X_1\le N-1)$ and determines X_2 such that $X_2\equiv(P-X_1)$ mod N, and then sends A, X_1^2 mod N and X_2^2 mod N to the verifier. The verifier knows $(X_1+X_2)^2$ mod $N\equiv(Y_1+Y_2)^2$ mod N because of (X_1+X_2) mod $N\equiv(Y_1+Y_2)$ mod N, and therefore the verifier can calculate $2(X_1X_2-Y_1Y_2)$ mod $N\equiv(Y_1^2+Y_2^2 - X_1^2 - X_2^2)$ mod N, leaded from $(X_1+X_2)^2$ mod N. Therefore, the verifier can determine whether $(X_1^2X_2^2Y_1^2Y_2^2)$ mod $N\equiv(64^{-1}((Y_1^2+Y_2^2 - X_1^2 - X_2^2)^2 - 4X_1^2X_2^2 - 4Y_1^2Y_2^2)^2)$ mod N is hold or not by using X_1^2 mod N, X_2^2 mod N and received from the prover, and Y_1^2 mod N, Y_2^2 mod N stored at the password file.

The PBSI is secure against replay attack, because X_1 is chosen in random. It is computationally infeasible to determine X_1 in X_1^2 mod N. Therefore, the PBSI is secure against pre-play, eavesdropping, man-in-the-middle. The PBSI is secure against off-line dictionary attack, because it is too huge to store 2^N candidates for X_1^2 mod N. The PBSI is secure against password file compromise without having performed off-line dictionary attack, because the password was not stored at the password file in clear text, and furthermore the PBSI is also secure password compromise with having performed off line dictionary attack, because the security against password file compromise of the PBSI depends on SQROOT problem, and it is too huge to store 2^N candidates for Y_1^2 mod N.

The performance of PBSI is in Table 1. In Table 1, we have assumed that the verifier first calculates $((Y_1^2+Y_2^2 - X_1^2 - X_2^2)^2) - 4X_1^2X_2^2 - 4Y_1^2Y_2^2)^2$ mod N and then calculates $64^{-1}(Y_1^2+Y_2^2 - X_1^2 - X_2^2)^2 - 4X_1^2X_2^2 - 4Y_1^2Y_2^2)^2$ mod N.

Table1 The performance of PBSI

	Prover	Verifier (Off line)	Verifier (On line)
Pass	1	0	0
Random number generation	1	1	0
Modular square multiplication	2	2	1
Modular multiplication	0	0	2

IV. CONCLUSIONS

We have presented an identification protocol the PBSI processing human memorable password. The PBSI is secure against the well known attacks such as replay attack, pre-play attack, man-in-the-middle attack, eavesdropping, off-line dictionary attack, password file compromise, and furthermore secure against the password file compromise with having performed off-line dictionary attack. It is the stability that is based on Square Root Problem, and we would like to suggest PBSI, enhancing the stability, for all of the well-known attacks by now including Off-line dictionary attack, password file compromise, Server and so on. The PBSI is also excellent in the aspect of the performance.

REFERENCES

[1] E. Bach, Algorithmic Number Theory, Volumn 1: Efficient Algorithms, MIT Press, Cambridge Massachusetts, 1996.

[2] M. J. Beller and Y. Yacobi, "Limitations of the kerberos authentication system", computer Communication Review, Vol. 20, pp. 119-132, 1990.

[3] H. Cohen, A Course in Computational Algebraic Number Theory Springer-Verlag, Berlin, 1993.

[4] U. Feige, A. Fiat and A. Shamir, "Zero Knowledge proof of identity", Journal of Cryptology, Vol.1, pp. 77-94, 1983.

[5] A. Fiat and A. Shamir, "How to prove yourself : Practical solutions to identification and signature

problems", Advances in Cryptology-CRYPTO' 86, LNCS 263, pp. 186-194, 1987.
[6] K. Gaarder and E. Snekkenes, "Applying a formal analysis technique to the CCITT X. 509 strong two way authentication protocol", Journal of Cryptology, Vol.3, pp. 81-98, 1991.
[7] L. Gong, "A security risk of depending on synchronized clocks", Operating System Review, Vol.26, pp. 49-53, 1992.
[8] L. C. Guillou and J. -J. Quisquater, "A practical zero-knowledge protocol to security microprocessor minimizing both transmission and memory", Advances in Cryptology-EUROCRYPT '88, LNCS 330, pp. 123-128, 1988.
[9] L. Lamport, "Password authentication with insecure communication", Communications of the ACM, Vol.24, pp. 770-772, 1981.
[10] R. Morris and K. Thompson, "Password security : a case history", Communications of the ACM, Vol.22, pp. 594-597, 1979.
[11] H. Woll, "Reductions among number theoretic problems, Information and Computation, Vol. 72, pp. 167-179, 1987.
[12] Jong-Min Park, Yong-Hun Kim, Beom-Joon Cho, "Password System Enhancing the Security against", The Korean Institute of Maritime Information & Communication Science, Vol. 8, No. 8, pp. 1790-1795, 2004.
[13] Jong-Min Park, "Efficient and Secure Authenticated Key Exchange", The Korean Institute of Maritime Information & Communication Science, Vol. 3, No. 3, pp. 163-166, 2005.
[14] Byung-Jun Park, Jong-Min Park, "One Pass Identification processing Password-based", The Korean Institute of Maritime Information & Communication Science, Vol. 4, No. 4, pp. 166-169, 2006.

Byung-Jun Park
He received the Apply Statistics and Ph.D.
degrees in the Dept. of Computer & Statistics from Chosun University.
His research interests in information security, information statistics, quality control, fourier series.

Jong-Min Park
He received the A.I and Ph.D.
degrees in the Dept. of computer Engineering from Chosun University.
His research interests information security, bio metrics, pattern recognition, artificial intelligence.