Title	Another Countermeasure to Forgeries over Message Recovery Signature							
Author(s)	MIYAJI, Atsuko							
Citation	IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences, E80-A(11): 2192-2200							
Issue Date	1997-11							
Туре	Journal Article							
Text version	publisher							
URL	http://hdl.handle.net/10119/4433							
Rights	Copyright (C)1997 IEICE. Atsuko MIYAJI, IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences, E80-A(11), 1997, 2192-2200. http://www.ieice.org/jpn/trans_online/ (許諾番号:08RB0099)							
Description								



PAPER Special Section on Information Theory and Its Applications

Another Countermeasure to Forgeries over Message Recovery Signature

Atsuko MIYAJI†, Member

Nyberg and Rueppel recently proposed a new SUMMARY ElGamal-type digital signature scheme with message recovery feature and its six variants ([12]). The advantage of small signed message length is effective especially in some applications like public key certifying protocols or the key exchange. But two forgeries that present a real threat over such applications are pointed out ([13],[14]). In certifying public keys or key exchanges, redundancy is not preferable in order to store or transfer small data. Therefore the current systems should be modified in order to integrate the Nyberg-Rueppel's signature into such applications. However, there has not been such a research that prevents the forgeries directly by improving the signature scheme. In this paper, we investigate a condition to avoid the forgeries directly. We also show some new message recovery signatures strong against the forgeries by adding a negligible computation amount to their signatures, while not increasing the signature size. The new scheme can be integrated into the above application without modifying the current systems, while maintaining the security.

key words: message recovery signature, discrete logarithms, forgery

1. Introduction

The RSA signature ([15]), which is based on the difficulty of factoring, has a message recovery feature. On the other hand, the ElGamal signature ([2]) and its six variants ([11], [16]) called EG-signatures in this paper, which are based on the difficulty of the discrete logarithm problem, do not have the message recovery feature. Recently Nyberg and Rueppel proposed a method to add the message recovery feature to all EG-signatures ([12]). Their method has an advantage of smaller signature data size for short messages. Therefore it is effective especially in some applications like public key certifying protocols ([3]) or the key exchange protocol ([1]). But two forgeries are pointed out ([13], [14]) in which an attacker can forge an exponentiation of a basepoint and he knows the exponent. These forgeries present a real threat over such applications since the attacker can forge a certificate or an exchanged key. In general, a redundancy generating function is a countermeasure to forgeries. But in public key certifying protocols or authenticated key exchange protocols redundancy is not preferable from the point of view of small stored or transferred data size. A possible countermeasure to

Manuscript received January 8, 1997. Manuscript revised May 6, 1997. the forgeries in such applications is to let each user use each different basepoint, or to let the authenticated key exchanges use an exponentiation of the receiver's public key instead of that of basepoint ([13],[14]). But such a countermeasure would spoil the benefit of El-Gamal, sharing the system-common parameters, and it would also require modifying the current systems based on DLP in order to integrate Nyberg-Rueppel's signature into them. Thus their signature would not realize an idea of adding only the message recovery feature to EG-signatures.

This paper's motivation is to study another countermeasure that the ElGamal-type message recovery signature is made strong against the two forgeries by improving Nyberg-Rueppel's signatures. In public key certifying protocols or authenticated key exchange protocols such a countermeasure is preferable. There are many variants ([4]) in the ElGamal-signature and the elliptic curve versions ([7],[8]). But little is known about their relative strength against forgeries. Therefore our approach would be also effective in classifying many variants under strength to forgeries, which would give a condition to select more secure variants.

In this paper, we first extend the known forgeries and analyze a general condition for the extended forgeries. By using the condition we show the new scheme which can be integrated into public key certifying protocols and the key exchange protocol without harming security and modifying the current system. Furthermore it can be constructed by adding only a negligible computation amount to the Nyberg-Rueppel's scheme, not increasing the signature size.

This paper is organized as follows. Section 2 summarizes the EG-signatures and Nyberg-Rueppel's signatures. Section 3 discusses some applications to which the forgeries become serious threat and extends the forgeries to the further three forgeries. Section 4 analyzes the general condition to avoid the forgeries. Section 5 shows some new message recovery signatures strengthened against the forgeries.

2. Message Recovery Signature Scheme

This section summarizes EG-signatures and Nyberg-Rueppel's signatures which add the message recovery feature to EG-signatures. The Nyberg-Rueppel's signatures are collectively called MR-signatures in this

[†]The author is with Matsushita Electric Industrial Co., LTD., Kadoma-shi, 571 Japan.

paper. In any signature schemes, the trusted authority chooses system parameters, that are a large prime p, a large prime factor q of p-1 and a basepoint $g \in \mathbb{F}_p = GF(p) = \{0, \cdots, p-1\}$ whose order is q. These system parameters are known to all users. The signer Alice has a secret key x_A and publishes its corresponding public key $y_A = g^{x_A}$.

EG-signatures

To sign a message $m \in \mathbb{F}_p^*$, she chooses a random number $k \in \mathbb{F}_q$, and computes $r_1 \equiv g^k \pmod p$, $r_1' \equiv r_1 \pmod q$ and

$$ak \equiv b + cx_A \pmod{q},\tag{1}$$

where (a,b,c) is a permutation of $(\pm m,\pm r_1',\pm s)$. Then the triplet $(m;(r_1,s))$ constitutes the signed message. The signature verification is done by checking the next equation,

$$r_1^a \equiv g^b y_A^c \pmod{p}. \tag{2}$$

MR-signatures

MR-signatures can be derived from EG-signatures by adding the message-mask equation (4) and replacing m (resp. r'_1) by 1 (resp. r'_2) in Eq. (1). To sign a message $m \in \mathbb{F}_p^*$, she chooses a random number $k \in \mathbb{F}_q$, and computes

$$r_1 \equiv g^k \pmod{p} \tag{3}$$

$$r_2 \equiv mr_1^{-1} \pmod{p} \tag{4}$$

$$r_2' \equiv r_2 \pmod{q}$$

$$ak \equiv b + cx_A \pmod{q},\tag{5}$$

where (a,b,c) is a permutation of $(\pm 1,\pm r_2',\pm s)$. Then the signature is given by (r_2,s) . The message can be recovered by computing a recovery equation

$$m \equiv g^{b/a} y_A^{c/a} r_2 \pmod{p}. \tag{6}$$

The verification of the signature may require further steps that add redundancy to the message before it is signed and that check the redundancy after recovery. The signature equation (5) leads to the following six equations if we neglect the \pm signs.

The ElGamal-type signatures can be constructed in other groups, as long as the discrete logarithm problem (DLP) is hard. So all the six MR-signatures can be also constructed on an elliptic curve, which are called MRE-signatures in this paper. In MRE-signatures, the system parameters are: an elliptic curve E/\mathbb{F}_p , a basepoint $G \in E(\mathbb{F}_p)$ and the prime order q of G. The signer Alice has a secret key x_A and publishes the corresponding public key $Y_A = x_A G$. Alice's procedure to

make a signature on $m \in \mathbb{F}_p^*$ is done in the same way as MR-signatures except for Eqs. (3) and (4), where these are changed to:

$$R_1 = kG, (7)$$

$$r_2 \equiv m \ x(R_1)^{-1} \pmod{p},\tag{8}$$

respectively. Here $x(R_1)$ denotes the x-coordinate of R_1 and Eq. (7) is computed in E. Also in MRE-signatures, the signature is given by (r_2,s) . The message can be recovered by computing $m \equiv x\left(\frac{b}{a}G + \frac{c}{a}Y_A\right)r_2 \pmod{p}$, where $\frac{b}{a}G + \frac{c}{a}Y_A$ is computed in E. Note that the signature equations of MR-signatures and MRE-signatures are the same whereas the message-mask equations are different each other.

3. Forgeries against MR-Signatures

Two forgeries against some schemes of MR-signatures have been presented in [13], [14], which are called the recovery-equation attack (recovery-eq-attack) using the basepoint g and the signature-equation attack (signature-eq-attack) using g in this paper. In order to analyze the general condition of forgeries, we extend the two forgeries to further three forgeries. The extended two forgeries, called the recovery-eq-attack using y_A and the signature-eq-attack using y_A in this paper, can be constructed by changing the function of g in each original attack to y_A . The last forgery, which is the same as the signature-eq-attack using g and y_A in one chosenmessage scenario, can control the forged message. The forgery is called the homomorphism attack in this paper. This section summarizes the extended three forgeries, and discusses their impact on some applications.

3.1 Extended Forgeries

The recovery-eq-attack using y_A against (S3)

This forgery can compute a signature (r_2,s) on a message of a special form $m \equiv My_A^e \pmod{p}$ for any chosen $M \in \mathbb{F}_p^*$ without ever seeing any signature and Alice's secret key. A forger chooses $\forall U, V \in \mathbb{F}_q$ and $\forall M \in \mathbb{F}_p^*$, and sets $r_2 \equiv My_A^Ug^V \pmod{p}$, $s \equiv -V \pmod{q}$ and $e \equiv r_2' + U \pmod{q}$. He sends (r_2,s) as a signature on $m = My_A^e$. We see that (r_2,s) is a valid signature on m since

$$g^{s}y_{A}^{r'_{2}}r_{2} \equiv g^{-V}y_{A}^{e-U}My_{A}^{U}g^{V} \equiv My_{A}^{e}$$
$$\equiv m \pmod{p}.$$

The signature-eq-attack using y_A against (S6)

This forgery can compute a signature $(\tilde{r_2}, \tilde{s})$ for a message of a special form $\tilde{m} \equiv my_A^{-n} \pmod{p}$ $(\forall n \in \mathbb{F}_q)$ if a message m and the signature (r_2, s) are given. A forger sets $\tilde{r_1} \equiv (mr_2^{-1})y_A^{-n} \equiv r_1y_A^{-n} \pmod{p}$ for $\forall n \in \mathbb{F}_q - \{0\}$. (There are q-1 variants.) He also sets a message $\tilde{m} \equiv my_A^{-n} \pmod{p}$, $\tilde{r_2} = r_2$ and $\tilde{s} \equiv s-n$

 \pmod{q} , and sends $(\tilde{r_2}, \tilde{s})$ as a signature on \tilde{m} . We see that $(\tilde{r_2}, \tilde{s})$ is a valid signature on \tilde{m} since

$$g^{\tilde{r_2}'}y_A^{\tilde{s}}\tilde{r_2} \equiv g^{r_2'}y_A^{s-n}r_2 \equiv my_A^{-n} \equiv \tilde{m} \pmod{p}.$$

The homomorphism attack against (S3)

This forgery can compute a signature $(\tilde{r_2}, \tilde{s})$ for any message \tilde{m} by assuming one chosen-message attack scenario: a forger has obtained Alice's signature (r_2, s) on one message $m \in \{\tilde{m}g^{-n}|n \in \mathbb{F}_q\} - \{\tilde{m}\}$. For simplicity we set the chosen-message $m \equiv \tilde{m}g \pmod{p}$. Then a forger sets $\tilde{r_2} = r_2$ and $\tilde{s} \equiv s - 1 \pmod{q}$. He sends $(\tilde{r_2}, \tilde{s})$ as a signature of \tilde{m} . We see that $(\tilde{r_2}, \tilde{s})$ is a valid signature on \tilde{m} since

$$g^{\tilde{s}} y_A^{\tilde{r}_2'} \tilde{r_2} \equiv g^{s-1} y_A^{r_2'} r_2 \equiv m g^{-1} \equiv \tilde{m} \pmod{p}.$$

The homomorphism attack is one chosen-message attack of the signature-eq-attack using g or y_A . But all schemes of MR-signatures that are vulnerable to the signature-eq-attack are not necessarily weak to the homomorphism attack. The reason will be analyzed in Sect. 4.

3.2 Forgery Threat to Some Applications

Both the recovery-eq-attack using g and the signature-eq-attack using g can forge a message of special form like mg^n , where the exponent n is known to the attacker. Such forgeries may often present a real threat on some applications. Actually they have serious impact on the most appropriate applications for MR-signatures like authenticated key exchanges and certifying public keys ([13], [14]).

Authenticated key exchanges

In the original idea of Diffie-Hellman key exchanges ([1]), an exponentiation $g^K \pmod{p}$ is transferred from Alice to Bob and the session key $K_{AB} \equiv y_B^K \pmod{p}$ is computed, where y_B is Bob's public key. The integration of MR-signatures is as follows ([12]): Alice transfers the signature (r_2,s) on $g^K \pmod{p}$ instead of $g^K \pmod{p}$, and Bob first recovers g^K from (r_2,s) and then computes $K_{AB} = (g^K)^{x_B} \pmod{p}$ with Bob's secret key x_B . MR-signatures become useful to transfer an authenticated g^K with smaller additional data size |s|. Comparing with the additional data size $|r_1| + |s|$ of the authenticated key exchanged by EG-signatures, it is reduced to about $\frac{1}{4}$ since s is about $\frac{1}{3}$ of s.

However if we use an MR-scheme vulnerable to the recovery-eq-attack using g or the signature-eq-attack using g, then the authenticated key exchange protocols become serious as follows.

First by recovery-eq-attack using g, an impersonator sets M=1 and forges an exponentiation $g^e \pmod p$ for his known e and establish a phony key $\tilde{K}_{AB} \equiv y_B^e \pmod p$.

Next we assume that an impersonator gets one session key K_{AB} and the corresponding signature for g^K . By

the signature-eq-attack using g he sets a random number n and forges $g^K g^n \equiv g^{K+n} \pmod{p}$. Then he can establish a phony key $\tilde{K}_{AB} \equiv K_{AB} y_B^n \equiv (g^{K+n})^{x_B} \pmod{p}$.

Public key certifying protocol

The basic idea of public key certifying protocols ([3]) is as follows: 1. a user's public key can be obtained from the user's name and the certificate; 2. the authenticity of the certificate is not directly verified, but the correct public key can be obtained only from the authentic certificate. This protocol has an advantage that the size of certificates is small. MR-signatures is suitably applied to this scheme ([12]): a Key Center generates the message recovery signature on an user name times the public key $ID_A \cdot g^{x_A} \pmod{p}$ using Center's secret key and publishes the signature as the user's certificate.

However if we use an MR-scheme vulnerable to the recovery-eq-attack using g or the signature-eq-attack using g, then the public key certifying protocol becomes serious as follows.

First by recovery-eq-attack using g, an attacker sets $M = ID_A$ for any user A and forges $ID_Ag^e \pmod{p}$ for the attacker's known e. Then the attacker can publish the signature as a phony certificate of A.

Next by the signature-eq-attack using g, any user A can change his/her public key as he/she likes and make a phony certificate if he/she once gets an authentic certificate of a public key $g^{x_A} \pmod{p}$ as follows. A can forge $(ID_Ag^{x_A})g^n \equiv ID_Ag^{x_A+n} \pmod{p}$ for his/her chosen n and publish the signature as a phony certificate of his/her new public key $g^{x_A+n} \pmod{p}$.

All MR-signatures are vulnerable to either attack as we will see in Sect. 4. Therefore MR-signatures cannot be integrated naturally into the above applications without harming security. In authenticated key exchanges or certifying public keys no redundancy is desired in order to transfer or store smaller data size. A possible countermeasure to the forgeries is as follows ([13], [14]): each user or Key Center has each different basepoint, or the authenticated key exchanges use an exponentiation y_B^K \pmod{p} instead of $g^K \pmod{p}$. But such a countermeasure would spoil the ElGamal scheme's benefit of sharing the system-common parameters, and it would also require modifying the current used systems based on DLP in order to integrate MR-signatures into them. Therefore another approach of avoiding the forgeries directly and maintaining the efficiency of MR-signatures would be preferable.

We have seen that the forgeries of a message of a special form become serious. Both the recovery-eq-attack using y_A and the signature-eq-attack using y_A forge a message of another special form like my_A^n (mod p), where the exponent n is known to the attacker. Though we don't think of an application on which such forgeries become serious for now, we had better avoid such forgeries also. Next section will show how to strengthen against all the five forgeries by im-

	MR-signatures						MRE-signatures					
	(S1)	(S2)	(S3)	(S4)	(S5)	(S6)	(S1)	(S2)	(S3)	(S4)	(S5)	(S6)
recovery-eq-attack: g	*	-*	††*	_*	††*	_*	tt	††	††	††	††	††
$:y_{A}$		††		_		††	††	††	tt	††	††	††
signature-eq-attack: g	††	††	_*		_*	_	††	††	_	_	_	_
$: y_A$	_	-	-	††	††	_	—	—	—	††	††	_
homomorphism attack	††		_	††	_	_	††	††	††	††	††	tt

 Table 1
 Strongness of MR- and MRE-signatures against forgeries.

proving MR-signatures.

4. Analyze the Condition of Forgery

The previous section extended the known two forgeries to further three forgeries as a first step. This sections analyzes these extended five forgeries. Table I summarizes the results. It shows strongness of each signature against each forgery, where "††" denotes strong, "—" denotes vulnerable, and "*" means the results shown in [13], [14]. We also analyze the necessary and sufficient condition of each attack and make clear why some types are vulnerable but the others are strong.

4.1 Recovery-Eq-Attack

Table 1 says that all MR-signatures are vulnerable to the recovery-eq-attack whereas all MRE-signatures are strong. The following discussion analyzes the condition on the recovery-eq-attack. Without loss of generality, we deal with the attack using y_A against MR-signatures.

From the recovery equation (6), if a forger finds a set of solutions of three variables (r_2, s, e) for a chosen $M \in \mathbb{F}_p^*$ that satisfies

$$r_2 \equiv (My_A^e)g^{-b/a}y_A^{-c/a} \pmod{p},\tag{9}$$

then (r_2, s) is a valid signature on a special form $m \equiv My_A^e$. For this special form, solving (9) can be reduced to solving the next simultaneous equations for two variables (s, e),

$$\begin{cases}
U \equiv -b/a \pmod{q} \\
V \equiv e - c/a \pmod{q}
\end{cases}$$
(10)

by setting $r_2 \equiv Mg^Uy^V_A \pmod{p}$ for any chosen $U, V \in \mathbb{F}_q$. Since (a,b,c) are represented by 1, r_2 and s, the solutions of (10) always exist except for special cases such that the former equation of (10) does not include s. But such special cases, (S2) and (S6), can be easily attacked by using Mg^e instead of My^e_A . Thus all MR-signatures are vulnerable to this attack using y_A or q.

To sum up, the necessary and sufficient condition for the recovery-eq-attack is that solving (9) can be reduced to solving simultaneous equations (10). This condition is that two algebraic relations among the three ex-

ponents e, b/a and c/a can be derived from the recovery-equation (9) for My_A^e , where the recovery-equation is determined by the message-mask equation (4). In order to find the message-mask equation strong against the recovery-eq-attack, let us re-define (4) as

$$r_2 \equiv f(r_1, m) \pmod{p},\tag{11}$$

where $f: \mathbb{F}_p \times \mathbb{F}_p \longrightarrow \mathbb{F}_p$ is a map such that m can be computed by $m = f^{-1}(r_1, r_2)$. The map f is known to all users. The above discussion, including the case of Mg^e , is summarized in the next theorem.

Theorem 1: The recovery-eq-attack is invalid for the DLP-based message recovery signature with a new message-mask equation (11) if and only if two algebraic relations among the three exponents e, b/a and c/a like (10) cannot be derived from either $r_2 \equiv f(g^{b/a}y_A^{c/a}, Mg^e) \pmod{p}$ or $r_2 \equiv f(g^{b/a}y_A^{c/a}, My_A^e) \pmod{p}$.

Let us describe the above map f concretely. We set $f(r_1,m)=mf_1(r_1)^{-1}$, where f_1 is a map from \mathbb{F}_p to \mathbb{F}_p . Namely we change Eq. (4) to

$$r_2 \equiv m f_1(r_1)^{-1} \pmod{p}.$$
 (12)

Then the recovery-eq-attack forges a special-form message $Mf_1(y_A^e)$ by deriving the simultaneous equation (10) from

$$r_2 \equiv M f_1(y_A^e) / f_1(g^{b/a} y_A^{c/a}) \pmod{p}.$$
 (13)

What kind of f_1 leads two algebraic relations among the three exponents e, b/a and c/a? If f_1 is a homomorphism, then Eq. (13) is changed to

$$r_2 \equiv M f_1(g^{-b/a} y_A^{e-c/a}) \pmod{p}.$$

So the three exponents e, b/a and c/a are converted to two algebraic relations (10). The recovery-eq-attack succeeds by first setting $r_2 \equiv M f_1(g^U y_A^V) \pmod{p}$ for any chosen $U, V \in \mathbb{F}_q$, and then solving (10). In MR-signatures, we can regard the map f_1 as an identity map, a kind of a homomorphism map. Therefore solving (9) can be reduced to solving (10). In the case of the attack using g, Eq. (13) is as follows,

$$r_2 \equiv M f_1(g^e) / f_1(g^{b/a} y_A^{c/a}) \pmod{p},$$
 (14)

where $m \equiv M f_1(g^e) \pmod{p}$. The above discussion is summarized as follows.

Corollary 1: If f_1 is a homomorphism map, then the three exponents e, b/a and c/a of both (13) and (14) are converted to two algebraic relations. Therefore the DLP-based message recovery signature with such a message-mask equation (12) is vulnerable to the recovery-eq-attack.

From Corollary 1, we would call the property that two algebraic relations are derived from Eq. (13) or (14) as a homomorphism-like property. So we must choose f_1 that does not have the homomorphism-like property. Here we show each example of f and f_1 .

Example 1: Define a map $f: \mathbb{F}_p \times \mathbb{F}_p \longrightarrow \mathbb{F}_p$ $((x,y) \longrightarrow x+y)$. Then two algebraic relations among e, b/a and c/a cannot be derived from

$$r_2 \equiv g^{b/a} y_A^{c/a} + M y_A^e \pmod{p}.$$

The same also holds in the case of Mg^e . Therefore the recovery-eq-attack is invalid.

Example 2: Define a map $f_1: \mathbb{F}_p \longrightarrow \mathbb{F}_p (x \longrightarrow x+g)$. Then Eq. (13) is

$$r_2 \equiv M(y_A^e + g)/(g^{b/a}y_A^{c/a} + g) \pmod{p}.$$

From the above equation, two algebraic relations among e, b/a and c/a can not be derived. The same also holds in Eq. (14). Therefore the recovery-eq-attack is invalid.

Note that the recovery-eq-attack becomes invalid by changing the message mask-equation slightly.

As for MRE-signatures, we can discuss in the same way as the above. But only the message-mask equation (8) is different from MR-signatures. In fact, Eq. (8) is equal to the case that the map f_1 in Eq. (12) is the x-coordinate function of an elliptic curve. The x-coordinate function on E, whatever an elliptic curve E is chosen, does not have a homomorphism-like property: since Eq. (13) is represented as

$$r_2 \equiv Mx(eY_A)/x\left(\frac{b}{a}G + \frac{c}{a}Y_A\right) \pmod{p},$$

two algebraic relations among e, b/a and c/a can not be derived. The same also holds in Eq. (14). Therefore all MRE-signatures are strong against the recovery-eqattack.

In [4], the message-mask equations different from that of MR-signatures are presented. Theorem 1 can be also applied to their message-mask equations.

4.2 Signature-Eq-Attack

This subsection makes clear the condition of the signature-eq-attack.

The signature-eq-attack using the basepoint

Assume that a forger gets Alice's MR-signature (r_2, s) for a message m. Then the forger can always construct a new commitment $\tilde{r_1} \equiv r_1/g \equiv g^{k-1} \pmod{p}$. He

does not know the correct discrete logarithm of $\tilde{r_1}$ but more importantly he knows it is equal to the value subtracted by 1 from the discrete logarithm of r_1 . First he transforms the signature equation (5) standing for the original $m \ (\equiv r_1r_2 \pmod{p}), r_2, s$ and k to that for the new $\tilde{m} \ (\equiv \tilde{r_1}\tilde{r_2} \pmod{p}), \tilde{r_2}, \tilde{s}$ and k-1, maintaining the congruity of the original signature equation. Namely he tries to find $(\tilde{r_2}, \tilde{s})$ satisfying the following equation,

$$a(k-1) \equiv (b-a) + cx_A \pmod{q},\tag{15}$$

where (a, b, c) is a pre-fixed permutation of $(1, r'_2, s)$. Then he sets $\tilde{m} \equiv \tilde{r_1}\tilde{r_2} \pmod{p}$. There exists a set of $(\tilde{r_2}, \tilde{s})$ satisfying (15) if and only if (a, b - a, c) is equal to the pre-fixed permutation of $(1, \tilde{r_2}', \tilde{s})$. Since the one coefficient b is not fixed, we see that the signature-eqattack succeeds if and only if we use the schemes of $b \neq 1$, namely b = s or $b = r'_2$, in Eq. (5): schemes (S3) and (S5), or (S4) and (S6) respectively. In scheme (S3) (resp. (S5)), a forger can generate the signature $(\tilde{r_2}, \tilde{s})$ by setting $\tilde{r_2} = r_2$ and $\tilde{s} \equiv s - 1 \pmod{q}$ (resp. $\tilde{s} \equiv s - r_2'$ \pmod{q} for $\tilde{m} \equiv \tilde{r_1}\tilde{r_2} \equiv (r_1/g)r_2 \equiv m/g \pmod{p}$. In scheme (S4) (resp. (S6)), he can generate the signature $(\tilde{r_2}, \tilde{s})$ on $\tilde{m} \equiv \tilde{r_1}\tilde{r_2} \pmod{p}$ by setting $\tilde{s} = s$ and $\tilde{r_2} \equiv r'_2 - s \pmod{q} \pmod{q}$ (resp. $\tilde{r_2} \equiv r'_2 - 1 \pmod{q}$). Note that, only in the schemes of setting $\tilde{r_2} = r_2$ (i.e. (S3) and (S5)), \tilde{m} is represented by m and a known parameter, q.

The signature-eq-attack using Alice's public key

The above attack uses a basepoint g in order to modify the original commitment $r_1 = g^k$. Considering the signature equation (5), a forger can construct a different commitment $\tilde{r_1} \equiv r_1/y_A \equiv g^{k-x_A} \pmod{p}$ by using Alice's public key y_A . The following discussion is almost the same as the above attack using g. First he transforms Eq. (5) standing for the original $m = r_1 r_2$, r_2 , $r_3 = r_4$ and $r_4 = r_4$, namely tries to find $r_4 = r_4$ satisfying the following equation,

$$a(k - x_A) \equiv b + (c - a)x_A \pmod{q},\tag{16}$$

where (a,b,c) is a pre-fixed permutation of $(1,r_2',s)$ and (a,b,c-a) is the pre-fixed permutation of $(1,\tilde{r_2'},\tilde{s})$, and next sets $\tilde{m}=\tilde{r_1}\tilde{r_2}$. Therefore the signature-eq-attack using y_A succeeds if and only if we use the schemes of $c \neq 1$, namely c=s or $c=r_2'$, in Eq. (5): schemes (S2) and (S6), or (S1) and (S3) respectively. In scheme (S2) (resp. (S6)), he can generate the signature $(\tilde{r_2},\tilde{s})$ by setting $\tilde{r_2}=r_2$ and $\tilde{s}\equiv s-r_2'\pmod{q}$ (resp. $\tilde{s}\equiv s-1$) for $\tilde{m}\equiv\tilde{r_1}\tilde{r_2}\equiv(r_1/y_A)r_2\equiv m/y_A\pmod{p}$. In scheme (S1) (resp. (S3)), he can generate the signature $(\tilde{r_2},\tilde{s})$ on $\tilde{m}\equiv\tilde{r_1}\tilde{r_2}\pmod{p}$ by setting $\tilde{s}=s$ and $\tilde{r_2}\equiv r_2'-s\pmod{q}$ (resp. $\tilde{r_2}\equiv r_2'-1\pmod{q}$). Note that, in the same way as the attack using g, only in the schemes of setting $\tilde{r_2}=r_2$ (i.e. (S2) and (S6)), \tilde{m} is represented by m and a known parameter, y_A .

As for MRE-signatures, the same discussion as MR-signatures holds since this attack utilizes only the characteristic of the signature equation.

To sum up, the necessary and sufficient condition for the signature-eq-attack is that a forger can construct the signature equation standing for a new commitment $\tilde{r}_1 = r_1/g$ or r_1/y_A by transforming the original signature equation while maintaining its congruity. This means that a set of coefficients (a, b, c) in the signature equation (5) satisfies (15) or (16). Thus the cases that (a, b, c) is a permutation of $(1, r'_2, s)$ (i.e. all MR- and MRE-signatures) are vulnerable to this attack. In order to find the coefficients strong against the signature-eqattack, let us re-define Eq. (5) as

$$h_a(r_2', s, 1)k$$

$$\equiv h_b(r_2', s, 1) + h_c(r_2', s, 1)x_A \pmod{q}, \qquad (17)$$

where h_a , h_b , and h_c are suitable maps from $\mathbb{F}_q \times \mathbb{F}_q \times \mathbb{F}_q$ to \mathbb{F}_q such that s can be computed in (17). Then the above discussion is summarized in the next theorem.

Theorem 2: The signature-eq-attack is invalid for the DLP-based message recovery signature with the new signature equation (17) if and only if for chosen r_2 and s, three maps h_a , h_b , and h_c satisfy the next two conditions for all but some fixed values \tilde{r}_2 and \tilde{s} :

1. if $h_a(r'_2, s, 1) \equiv h_a(\tilde{r_2}', \tilde{s}, 1) \pmod{q}$ and $h_c(r'_2, s, 1) \equiv h_c(\tilde{r_2}', \tilde{s}, 1) \pmod{q}$, then $h_b(r'_2, s, 1) - h_a(r'_2, s, 1) \not\equiv h_b(\tilde{r_2}', \tilde{s}, 1) \pmod{q}$ (avoiding Eq. (15)), 2. if $h_a(r'_2, s, 1) \equiv h_a(\tilde{r_2}', \tilde{s}, 1) \pmod{q}$ and $h_b(r'_2, s, 1) \equiv h_b(\tilde{r_2}', \tilde{s}, 1) \pmod{q}$, then $h_c(r'_2, s, 1) - h_a(r'_2, s, 1) \not\equiv h_c(\tilde{r_2}', \tilde{s}, 1) \pmod{q}$ (avoiding Eq. (16)).

Here "some fixed values $\tilde{r_2}$ and \tilde{s} " mean trivial cases such that the signature-eq-attack succeeds if and only if $\tilde{r_2}' = r_2' = 0$ like the next Example 3.

Example 3: Set $h_a(r'_2, s, 1) \equiv r'_2$, $h_b(r'_2, s, 1) \equiv r'_2 + s + 1 \pmod{q}$ and $h_c(r'_2, s, 1) \equiv s \pmod{q}$. Namely set the signature equation as follows:

$$r_2'k \equiv (r_2' + s + 1) + sx_A \pmod{q}.$$
 (18)

Then the signature-eq-attack does not succeed except for $\tilde{s} = s$ and $\tilde{r_2}' = r_2' = 0$. Such trivial cases can be easily excluded by restricting beforehand $r_2' \in \mathbb{F}_q$ to $\mathbb{F}_q - \{0\}$.

In the same way as Example 3, any permutation of $(h_a, h_b, h_c) = (r'_2, r'_2 + s + 1, s)$ can avoid the signature-eq-attack by excluding each trivial case like $r'_2 = 0$ or s = 0. Example 3 is the optimal case since the signature generation needs only one inversion $\frac{1}{x_A + 1}$ which can be precomputed beforehand. Compared with the optimal MR-signatures (S3), Example 3 requires only one additional inversion in the signature verification. In [4], signature equations different from Eq. (5) are presented. Theorem 2 can be also applied to the signature equations.

4.3 Homomorphism Attack

The homomorphism attack is one chosen-message at-

tack of the signature-eq-attack and forges any message. Table 1 shows that all MR-signatures $(S1)\sim(S6)$ are not necessarily vulnerable to the homomorphism attack though all of them are vulnerable to the signature-eqattack using g or y_A . This subsection deals with MR-signatures and investigates why some cases are extended to the homomorphism attack.

As we have seen in Sect. 4.2, there are two types of \tilde{m} forged by the signature-eq-attack: first \tilde{m} depends only on a message m and known parameters like g or y_A , and second \tilde{m} depends both on m and the signature (r_2, s) . In the former type, a suitable chosen-message m for an intentional message \tilde{m} can be constructed by setting $m \equiv \tilde{m}g \pmod{p}$ or $m \equiv \tilde{m}y_A \pmod{p}$ as shown in Sect. 3.1. But in the latter type, a forger cannot set such m since he cannot guess beforehand what \tilde{m} is forged. This difference is determined by whether the forged \tilde{m} by the signature-eq-attack is independent of the parameters k, r_2 and s which the signer Alice can take arbitrarily for m. Namely the signature-eq-attack is extended to the homomorphism attack if and only if \tilde{m} can be represented only by the original m and a known parameter. Therefore the homomorphism attack is serious only in the cases of (S2), (S3), (S5) and (S6).

The condition for the homomorphism attack is more explicitly written as follows: the forged message \tilde{m} by the signature-eq-attack is represented as

$$\tilde{m} (\equiv \tilde{r_1} \tilde{r_2}) \equiv \exists \varphi(m, g, p, q, y_A) \pmod{p},$$
 (19)

where φ is a suitable function to \mathbb{F}_p such that $m \equiv \varphi^{-1}(\tilde{m}, g, p, q, y_A) \pmod{p}$ exists. Then a chosen-message m for an intentional message \tilde{m} can be defined as $m \equiv \varphi^{-1}(\tilde{m}, g, p, q, y_A) \pmod{p}$. Since Eq. (19) is determined by the message-mask equation (4), we will improve (4) in order to be strong against the homomorphism attack. Using Eq. (11) in Sect. 4.1, the relation equation (19) between m and \tilde{m} forged by the signature-eq-attack is represented as:

if the attack using g, then

$$\tilde{m} \equiv f^{-1}(\tilde{r_1}, r_2)$$

$$\equiv m f^{-1}(q^k q^{-1}, r_2) / f^{-1}(q^k, r_2) \pmod{p} \quad (20)$$

if the attack using y_A , then

$$\tilde{m} \equiv f^{-1}(\tilde{r_1}, r_2)$$

$$\equiv m f^{-1}(g^k y_A^{-1}, r_2) / f^{-1}(g^k, r_2) \pmod{p} \quad (21)$$

where $m \equiv f^{-1}(g^k, r_2) \pmod{p}$. Therefore the homomorphism-attack is invalid if at least one term of k or r_2 , which the signer Alice can take arbitrarily, is not cancelled in Eqs. (20) and (21).

We have seen the condition for the homomorphism attack to be extended from the signature-eq-attack. Of course the homomorphism attack is invalid for a message recovery signature on which the signature-eq-attack is invalid. Thus the condition to avoid the homomorphism attack is summarized as follows.

Theorem 3: The homomorphism-attack is invalid for the DLP-based message recovery signature if the signature-eq-attack does not work for it (i.e. if it satisfies Theorem 2), or if at least one term of k or r_2 in Eqs. (20) and (21) is not cancelled with a new message-mask equation (11).

Let us describe the above map f concretely by using f_1 of Eq. (12). Then the relation between m and \tilde{m} forged by the signature-eq-attack is represented as: if the attack using g, then

$$\tilde{m} \equiv f_1(\tilde{r_1})r_2 \equiv mf_1(g^k g^{-1})/f_1(g^k) \pmod{p}$$
(22)

if the attack using y_A , then

$$\tilde{m} \equiv f_1(\tilde{r_1})r_2 \equiv mf_1(g^k y_A^{-1})/f_1(g^k) \pmod{\dot{p}}$$
(23)

What kind of f_1 cancels k in Eqs. (22) and (23)? If f_1 is a homomorphism, then the term k is cancelled. So Eqs. (22) and (23) lead $\tilde{m} \equiv \varphi(m,g,y_A)$ in both cases. We can regard that MR-signatures use an identity map, a homomorphism map (a vulnerable map) as f_1 . The above discussion is summarized as follows.

Corollary 2: If f_1 is a homomorphism map, the term k in Eqs. (22) and (23) is cancelled in both cases. Therefore in the DLP-based message recovery signature with such a message-mask equation (12), the homomorphism-attack is extended from the signature-eq-attack.

From Corollary 2, we would call the property that the term k in Eqs. (22) and (23) is cancelled as another homomorphism-like property.

Here we show each example of f and f_1 .

Example 4: With the same map f as defined in Example 1, Eq. (20) becomes

$$\tilde{m} \equiv m(g^{k-1} + r_2)/(g^k + r_2) \pmod{p}.$$

So neither the term k nor r_2 are cancelled. The same also holds in (21). Therefore the homomorphism attack is invalid.

Example 5: With the same map f_1 as defined in Example 2, Eq. (22) becomes

$$\tilde{m} \equiv m(g^{k-1} + g)/(g^k + g) \pmod{p}.$$

So the term k is not cancelled. The same also holds in (22). Therefore the homomorphism attack is invalid.

As for MRE-signatures, the map f_1 in Eq. (12) is the x-coordinate function. The x-coordinate function on E, whatever an elliptic curve E is chosen, does not have the homomorphism-like property:

$$x((k-1)G)/x(kG) = x(kG-G)/x(kG)$$
$$= \varphi(k,G),$$

where the term k is not cancelled in φ . Therefore all MRE-signatures are strong against the homomorphism attack.

4.4 Further Discussion

There are many variants in the ElGamal-signature, including elliptic curve versions. However, a general condition of selecting more secure variants or avoiding insecure variants has been scarcely known, though there might exist a general bad property which often causes forgeries. Our approach is effective in studying such a condition. Actually we have seen that two types of forgery are caused by a similar homomorphism-like property (Corollary 1 and 2). Also in RSA-signature, such a property like homomorphism causes some attacks. Importantly, in ElGamal-type message recovery signature the homomorphism-like property can be excluded by only changing a message-mask equation, though in RSA-signature the property is considered not to be excluded. Especially in elliptic curve message recovery signature, the x-coordinate function in messagemask equation already excludes such a bad property, and what is more important, the function does not require any additional computation. The x-coordinate function is also used in the original ElGamal-type signature over elliptic curves, which might present a good feature of avoiding another forgery.

5. Suitable Message Recovery Signatures

The advantage of message recovery signature is small signed message length for a short message. Therefore it is effective especially in some applications like public key certifying protocols or the key exchange. As we have seen in Sect. 3, both the recovery-eq-attack using g and the signature-eq-attack using g have serious impact on the most appropriate applications like authenticated key exchanges and certifying public keys. However All MR-signatures are vulnerable to either attack as we have seen in Sect. 4. In authenticated key exchanges or certifying public keys no redundancy is desired in order to transfer or store smaller data size. Therefore MR-signatures should be directly strong against the critical attack, the recovery-eq-attack using g and the signature-eq-attack using g, at least.

Here we present some schemes suitable for message recovery signature, based on the results of Sect. 4. The efficiency is also discussed from the point of view of the computation amount and the signature size.

Suitable message recovery signature over \mathbb{F}_p

(F1) The signature-equation of (S2) is strong against the signature-eq-attack using g but the message-mask equation is vulnerable to the recovery-eq-attack using g. So we use the signature-equation of (S2) and change the message-mask equation to

$$r_2 \equiv m(r_1 + g)^{-1} \pmod{p},$$

according to Theorem 1. Here we summarize the scheme (F2): to sign a message $m \in \mathbb{F}_p^*$, Alice chooses a random number $k \in \mathbb{F}_q$, computes $r_1 \equiv g^k \pmod p$ and

$$r_2 \equiv m(r_1 + g)^{-1} \pmod{p},$$

and sets $r_2' \equiv r_2 \pmod{q}$. Then she computes s from

$$r_2'k \equiv 1 + sx_A \pmod{q}$$
.

The signature is given by (r_2, s) . The message can be recovered by computing

$$m \equiv (g^{1/r_2'}y_A^{s/r_2'} + g)r_2 \pmod{p}.$$

(F2) If we consider all the five attacks, the message recovery signature can be constructed as follows. To sign a message $m \in \mathbb{F}_p^*$, Alice chooses a random number $k \in \mathbb{F}_q$, computes $r_1 \equiv g^k \pmod{p}$ and

$$r_2 \equiv m(r_1 + g)^{-1} \pmod{p}$$
 (Theorem 1),

and sets $r_2' \equiv r_2 \pmod{q}$. Then she computes s from

$$r'_2 k \equiv (1 + r'_2 + s) + sx_A \pmod{q}$$
(Theorem 2 and 3).

The signature is given by (r_2, s) . The message can be recovered by computing

$$m \equiv (g^{(1+r_2'+s)/r_2'}y_A^{s/r_2'} + g)r_2 \pmod{p}.$$

Comparing with the optimal scheme (S3) of MR-signatures, the above two new \mathbb{F}_p -signatures require an additional inversion only in the signature verification (we can precompute $\frac{1}{x_A}$ or $\frac{1}{x_A+1}$ in the signature generation), and the signature size does not increase. As a result, the efficiency of the two schemes are the same as scheme (S2) in MR-signatures. Furthermore the scheme (F1) is strong against the critical attack, the recovery-eq-attack using g and the signature-eq-attack using g, and the scheme (F2) is strong against all the five attacks. Therefore theses schemes can be integrated naturally into public key certifying protocols and the key exchange protocol without harming security and modifying the current system. As for the relative security between the new message recovery signatures and EG-signatures, almost the same discussion as [9] holds.

Suitable message recovery signature over E/\mathbb{F}_p

In the same way as the case of \mathbb{F}_p , suitable message recovery signatures over E/\mathbb{F}_p can be constructed as follows.

(E1) To sign $m \in \mathbb{F}_p^*$, Alice chooses a random number $k \in \mathbb{F}_q$, computes $R_1 = kG$ and

$$r_2 \equiv m \ x(R_1)^{-1} \pmod{p}$$
 (Theorem 1),

and sets $r_2' \equiv r_2 \pmod{q}$. Then she computes s from

$$r_2'k \equiv 1 + sx_A \pmod{q}$$
 (S2).

The signature is given by (r_2, s) . The message can be recovered by computing

$$m \equiv x \left(\frac{1}{r_2'}G + \frac{s}{r_2'}Y_A\right)r_2 \pmod{p}.$$

(E2) To sign $m \in \mathbb{F}_p^*$, Alice chooses a random number $k \in \mathbb{F}_q$, computes $R_1 = kG$ and

$$r_2 \equiv m \ x(R_1)^{-1} \pmod{p}$$
 (Theorem 1),

and sets $r_2' \equiv r_2 \pmod{q}$. Then she computes s from

$$r_2'k \equiv (1 + r_2' + s) + sx_A \pmod{q}$$
(Theorem 2 and 3).

The signature is given by (r_2, s) . The message can be recovered by computing

$$m \equiv x \left(\frac{1 + r_2' + s}{r_2'} G + \frac{s}{r_2'} Y_A \right) r_2 \pmod{p}.$$

These new schemes are also implemented only by adding one inversion in the signature verification to the optimal scheme (S3) in MRE-signatures. The signature size does not increase also. Furthermore the scheme (E1) is strong against the critical attack, the recovery-eq-attack using g and the signature-eq-attack using g, and the scheme (E2) is strong against all the five attacks. Note that the E/\mathbb{F}_p -signature schemes can be strengthened by improving only the signature-equation. As for the relative security between the new message recovery signatures and EG-signatures over elliptic curves, almost the same discussion as [9] holds.

Remark

A countermeasure to forgeries might be introducing redundancy in message. Here we discuss the differences, especially in public key certifying protocol, between our approach and the redundancy function ([6]). The redundancy function apparently extends size of input message double. Therefore the public key certifying protocol with the redundancy function first divides a public key into two blocks and then generates the signature on each block. Consequently the size of certificates is just twice as large as that with our approach. Therefore our approach can make smaller size of certificates. As for security, in our approach the authenticity of the certificate is not directly verified. In the case of redundancy function, the authenticity is directly verified, but it is not shown how much redundancy is enough. Neither our approach nor the redundancy function are proved to be secure against all attack.

6. Conclusion

In this paper, we have shown the next two results.

1. The relative strength among the DLP-based message recovery signatures against the five forgeries has been clarified:

There are many variants in the DLP-based message recovery signatures ([4], [13]). But little was known about the relative strength against the forgeries and the causes of forgeries. We have shown the condition of the forgeries against all these variants and shown how to choose

the signature-equation and the message-mask equation in order to overcome each forgery.

2. A new countermeasure to the serious forgeries against MR-signatures has been proposed:

We have discussed what kinds of forgery become serious and shown that the forgeries against a message of a special form like mg^n , the recovery-eq- and signature-eq-attack, become serious. These forgeries have serious impact on typical applications with no redundancy, like certifying public keys, or key exchange protocols. We have shown some message recovery signatures strong against the critical forgeries by adding a negligible computation amount to MR-signatures, and not increasing the signature size. The new signatures would realize an idea of adding only the message recovery feature to EG-signatures: they can be integrated naturally into the current applications like certifying public keys and key exchange protocols.

Acknowledgements

The author wishes to thank the anonymous referees for their valuable comments.

References

- [1] W. Diffie and M. Hellman, "New directions in cryptography," IEEE Trans. Inf. Theory, vol.IT-22, pp.644-654, 1976.
- [2] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," IEEE Trans. Inf. Theory, vol.IT-31, pp.469-472, 1985.
- [3] C.G. Günther, "An identity-based key-exchange protocol," Advances in Cryptology-Proceedings of Eurocrypt'89, Lecture Notes in Computer Science, vol.434, pp.29–37, Springer-Verlag, 1990.
- [4] P. Horster, M. Michels, and H. Petersen, "Meta-Message Recovery and Meta-Blind signature schemes based on the discrete logarithm problem and their applications," Advances in Cryptology-Proceedings of Asiacrypt '94, Lecture Notes in Computer Science, vol.917, pp.224-237, Springer-Verlag, 1995.
- [5] L. Harn and Y. Xu, "Design of generalized ElGamal type digital signature schemes based on discrete logarithm," Electronics letters, vol.30, no.24, p.2025, 1994.
- [6] ISO/IEC, 9796, Information technology—Security techniques— "Digital signature scheme giving message recovery."
- [7] N. Koblitz, "Elliptic curve cryptosystems," Mathematics of Computation, vol.48, pp.203-209, 1987.
- [8] V.S. Miller, "Use of elliptic curves in cryptography," Advances in Cryptology-Proceedings of Crypto'85, Lecture Notes in Computer Science, vol.218, pp.417–426, Springer-Verlag, 1986.
- [9] A. Miyaji, "A message recovery signature scheme equivalent to DSA over elliptic curves," Advances in Cryptology-Proceedings of ASIACRYPT'96, Lecture Notes in Computer Science, vol.1163, pp.1-14, Springer-Verlag, 1996.
- [10] A. Menezes, T. Okamoto, and S. Vanstone, "Reducing elliptic curve logarithms to logarithms in a finite field," Proc. 22nd Annual ACM Symposium on the Theory of Computing, pp.80-89, 1991.

- [11] "Proposed federal information processing standard for digital signature standard (DSS)," Federal Register, vol.56, no.169, pp.42980-42982, 30 Aug. 1991.
- [12] K. Nyberg and R.A. Rueppel, "A new signature scheme based on the DSA giving message recovery," Proc. 1st ACM Conference on Computer and Communications Security, 1993.
- [13] K. Nyberg and R.A. Rueppel, "Message recovery for signature schemes based on the discrete logarithm problem," Advances in Cryptology-Proceedings of Eurocrypt'94, Lecture Notes in Computer Science, vol.950, pp.182-193, Springer-Verlag, 1995.
- [14] K. Nyberg and R.A. Rueppel, "Message recovery for signature schemes based on the discrete logarithm problem," Designs Codes and Cryptography, vol.7, pp.61-81, 1996.
- [15] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," Communications of the ACM, vol.21, no.2, pp.120-126, 1978.
- [16] C.P. Schnorr, "Efficient identification and signatures for smart cards," Advances in cryptology-Proceedings of Crypto'89, Lecture Notes in Computer Science, vol.435, pp.239-252, Springer-Verlag, 1989.



Atsuko Miyaji She was born in Osaka, Japan, in 1965. She received the B.Sc. and the M.Sc. degrees in mathematics from Osaka University, Osaka, Japan in 1988 and 1990, respectively. Since 1990, she has been with Multimedia Development Center in Matsushita Electric Industrial Co., LTD. and engaged in research for secure communication. Her research interests include the application of projective varieties theory into cryptography and in-

formation security. She is a member of the International Association for Cryptologic Research.