The Analysis of Zheng-Seberry Scheme

David Soldera and Jennifer Seberry
Centre for Computer Security Research
University of Wollongong, NSW 2522, Australia
E-mail: jennie@uow.edu.au

Chengxin Qu
School of Mathematical and Computer Science
University of New England, Armidale, NWS 2531, Australia
E-mail: cxqu@turing.une.edu.au

Abstract

The Zheng-Seberry (ZS) encryption scheme was published in 1993 and was one of the first practical schemes that was considered secure against a chosen ciphertext adversary. This paper shows some problems that the semantic security of the one-way hash variant of the ZS scheme is insecure on some special circumstances. Attempts to modify the ZS scheme resulted on an El-Gamal variant that is provably secure in the random oracle model.

1 Introduction

In 1993 Zheng-Seberry [9] presented a paper introducing three new public-key encryption schemes that were efficient and considered secure against a chosen ciphertext adversary, under some assumptions. Since then much progress has made in the area of provable security for public-key cryptosystems, from those that use the Random Oracle (RO) model [2] to the scheme by Cramer-Shoup (CS) [4] that is provably secure using standard public key cryptography assumptions.

Using the RO model or standard assumptions represent opposite ends of the provable security spectrum. The RO model yields extremely efficient schemes yet practical implementations using hash functions fall short of actual RO's. Using standard assumptions gives us tremendous confidence in security yet schemes are still too inefficient for the majority of practical implementations.

The hardness of the Diffie-Hellman decision problem is essentially equivalent to the semantic security of basic El-Gamal encryption scheme [5]. The basic El-Gamal scheme is completely insecure against adaptive chosen ciphertext attack. This new Secure El-Gamal scheme was born out of the OWH variant of the original ZS scheme, since as shall be seen in section 2.2 it is insecure against a chosen ciphertext adversary (CCA). Securing the ZS scheme meant providing a proof for security and the best proof (in that it requires the least assumptions) was that by Cramer-Shoup. Unfortunately the CS proof cannot easily be used to prove the security of Secure El-gamal. So section 3 presents the new Secure El-Gamal scheme, and a proof of security which borrows many parts of the CS proof. Unfortunately, the proof still needs to rely on the random oracle model, but encouragingly, it only relies on it in a minimal way.

It has become standard practice that the level of security required for a public-key cryptosystem is indistinguishability of encryptions, IND, (equivalently semantic security or non-malleability) against an adaptive chosen ciphertext adversary (CCA2), for formal definitions see [1]. The basic idea behind an IND-CCA2 adversary is that they are given access to an encryption and decryption oracle, they then choose two messages, one of which gets encrypted (they do not know which). They are then presented with the ciphertext of the encrypted message and asked to determine which of the two messages was encrypted. They must succeed with probability non-negligibly better than 0.5. The only restriction is the adversary may not query the decryption oracle with the challenge ciphertext.

2 Original ZS

The ZS paper presented three variants of an El-Gamal like cryptosystem. The three variants were described as 'immunising' the cryptosystem against a CCA2 adversary. The variants incorporated a one-way hash function (OWH), a universal hash function and a digital signature.

2.1 ZS-OWH

The ZS-OWH variant is presented below.

ZS-OWH
Preliminaries: Consider message of length n, a one-way hash function H with output length k_0 and a PRNG G with output length $n + k_0$. Operations are modulo p and there is a generator g.
Key Generation: Private key is $x_R \in GF(p)$ and public key is $y_R = g^{x_R} \bmod p$.
Encryption; 1, $x \in_R [1, p-1]$ 2, $z = G(y_R^x)_{[1K\,(n+k_0)]}$ 3, $t = H(m)$ 4, $c_1 = g^x$ 5, $c_2 = z \oplus (m\|\|t)$ Ciphertext is (c_1, c_2)
Decryption: 1, $z' = G(c_1^{x_R})_{[1K\,(n+k_0)]}$ 2, $w = z' \oplus c_2$ 3, $m = w_{[1\ldots n]}$ 4, $t' = w_{[(n+1)K(n+k_0)]}$ If $H(m) = t'$, then output m else output $\emptyset$.

The security of ZS-OWH depends on the hardness of Diffie-Hellman one way problem.

2.2 Breaking ZS-OWH in IND-CCA2 Sense

> "Due to the involvement of $t = H(m)$, the creation of the ciphertext is apparently impossible without the knowledge of x and m. .... This motivates us to introduce a notion called $sole - samplable$ space." ([9], pg. 721)

If the authors had to pick an assumption in the ZS paper that ultimately turned out to be incorrect, the above assumption would be an appropriate choice. As it turns out an adversary can create a new ciphertext from an existing ciphertext, if the message in the existing ciphertext is known.

To see how this is achieved consider the last part of the ciphertext,

$$c_2 = z \oplus (m||t) = z \oplus (m||H(m)),$$

which just depends on the message. So if the message is known, this part of the ciphertext can be recreated. If the adversary wishes to replace the message m with another message m', this can be achieved via:

$$c_2' \quad = \quad c_2 \oplus (m||H(m)) \oplus (m'||H(m'))$$

$$
\begin{aligned}
&= z \oplus (m||H(m)) \oplus (m||H(m)) \oplus (m'||H(m')) \\
&= z \oplus [(m||H(m)) \oplus (m||H(m))] \oplus (m'||H(m')) \\
&= z \oplus (m'||H(m')),
\end{aligned}
$$

in which $[(m||H(m)) \oplus (m||H(m))] = 0$ due to Boolean addition.

The new ciphertext is (c_1, c_2') and the adversary is successful in manipulating the cryptosystem.

This attack can be used by a CCA2 adversary to defeat IND and the adversary succeeds 100% of the time. In this situation the adversary does not know which of two messages, m_0 or m_1, has been encrypted, but they know one of them has been. Let the encrypted message be m_b where $b \in [0, 1]$. The adversary uses the above attack by setting $m = m_0$ and $m' = m_1$ and creates a new cryptogram via:

$$
\begin{aligned}
c_2' &= c_2 \oplus [m_0||H(m_0)] \oplus [m_1||H(m_1)] \\
&= z \oplus [m_b||H(m_b)] \oplus [m_0||H(m_0)] \oplus [m_1||H(m_1)] \\
&= z \oplus [m_{\neg b}||H(m_{\neg b})]
\end{aligned}
$$

Hence the adversary creates a new ciphertext (c_1, c_2'), which is a valid ciphertext for the message that was not encrypted in the challenge ciphertext. Since the adversary is a CCA2 adversary, and the new ciphertext is not the challenge ciphertext, they may query the decryption oracle with it. The decryption oracle will dutifully return the message that was not encrypted, m_b, and the adversary makes their choice for b as corresponding to the message not returned by the decryption oracle.

The ZS-OWH scheme is largely of theoretical value to the cryptographic community, so while breaking the scheme does not have many practical implications, it is still of theoretical use. This break highlights the importance of adding random information to the check on the message, as shall be shown. Also, as recently as EUROCRYPT 2000, a paper [6] made reference to the ZS paper with the implication being it was secure, under some assumptions. So this attack means ZS-OWH now needs to be added to the list of schemes that were considered secure but turned out to be insecure.

This attack on ZS-OWH is a very trivial one and as could be expected a trivial change to the scheme thwarts this attack. By simply creating a new variable $r = y_R^x$ and changing $t = H(m||r)$, then the attack no longer works. The change incorporates some randomness into the hash calculation and thus defeats the above attack as the adversary can no longer create the concatenation of message and hash because the adversary does not know the

random information. This change defeats the above attack, but of course does not prove the security of the scheme.

This change was borrowed from an authenticated-encryption version of ZS-OWH by Zheng [8], however Zheng stresses that the changes made are only needed for the new scheme proposed and that the original scheme is secure.

3 Secure El-Gamal

The attack and the repair of the original ZS-OWH leaves a rather large question mark over its security. Securing the original ZS-OWH scheme led to a new El-Gamal variant. Great efforts were made to prove the security of this new variant using the CS proof and thus derive a scheme that was secure under some reasonable assumptions, but without using the RO model. Unfortunately, this goal was not realised, but encouragingly the proof does not heavily rely on the RO model.

Secure El-Gamal
Preliminaries: Consider messages of length $n - k_0$, a random oracle H with output length k_0. Operations are modulo p where $p = 2q + 1$ (q is a prime) and a generator g of order q.
Key Generation: Private key is $x_R \in GF(p)$ and public key is $y_R = g^{x_R} \bmod p$.
Encryption: Encrypt message m as follows; 1, $x \in_R [1, p-1]$ 2, $r = y_R^x$ 3, $t = H(m\|\|r)$ 4, $c_1 = g^x$ 5, $c_3 = r \cdot (m\|\|t)^2$ Ciphertext is (c_1, c_3).
Decryption: 1, $r' = c_1^{x_R}$ 2, $w = \sqrt{\frac{c_3}{r'}}$ (choose square root that yields the correct hash) 3, $m = w_{[1\cdots(n-k_0)]}$ 4, $t' = w_{[(n-k_0+1)Kn]}$ If $H(m\|\|r') = t'$, then output m, else output $\emptyset$.

The differences between this and the original El-Gamal scheme is the addition of the hash appended to the message, and the squaring of the message and hash to convert them into a quadratic residue (this makes it an element of the quadratic residues of $GF(p)$, the group of order q). Note

that in step 2 of the decryption, if neither square root yields a correct hash then the output is $\emptyset$.

The proof relies on the difficulty of the Decision Diffie-Hellman Problem (DDHP), the definition of which, from Cramer-Shoup, is given below.

Definition 1 - *([4], pg. 16) Let G be a group of large prime order q. Consider the following two distributions:*

- *the distribution R of random quadruples $(g_1, g_2, u_1, u_2) \in G^4$;*
- *the distribution D of quadruples $(g_1, g_2, u_1, u_2) \in G^4$, where g_1, g_2 are random, and $u_1 = g_1^r$ and $u_2 = g_2^r$ for random $r \in Z_q$.*

An algorithm that solves the DDHP is a statistical test that can effectively distinguish between these two distributions. For a given quadruple coming from one of the two distributions, it should output 0 or 1 and there should be a non-negligible difference between the probability that it outputs 1 given an input from R and the probability that it outputs 1 given an input from D. The decision Diffie-Hellman problem is hard if there is no such polynomial-time statistics test.

The construction of the proof is as follows. It is assumed an adversary that can break the cryptosystem in the IND-CCA2 sense exists, and then it is shown how this adversary can unwittingly be used to help solve what is considered a computationally unfeasible problem, in this case the DDHP. The construction of the proof can be seen in Figure 1.

The input to the proof are quadruples coming from either D or R (but not both). These go to a constructed simulator, which is responsible for, the creation of keys, simulation of an encryption oracle and simulation of a decryption oracle. The IND-CCA2 adversary receives all its information, including oracle queries, from the simulator.

The proof runs as follows. A quadruple is input. The simulator creates a valid secret key (once only) and the public key, which is passed to the IND-CCA2 adversary. The adversary runs its first stage A_1 and produces two messages m_0 and m_1. Then it passes the two messages to the simulated encryption oracle. The simulated encryption oracle chooses a random bit $b \in [0, 1]$, encrypts m_b and passes the challenge ciphertext back to the adversary. The adversary cannot see the simulator's choice for b.

The adversary then runs its second stage, A_2, on the challenge ciphertext and outputs its guess, b', for the random bit. Both the simulator and the adversary pass b and b' respectively to a distinguisher that outputs 1 if $b = b'$ otherwise 0.

Consider the case when the input comes from R, the simulator is unable to create a valid ciphertext (as the relation that quadruples from D have, are not present in quadruples from R). This fact will be crucial in showing the adversary cannot succeed in guessing b with any advantage. Alternatively, when the input comes from D, then the simulator creates a perfectly valid ciphertext and the adversary can guess the bit b with an advantage.

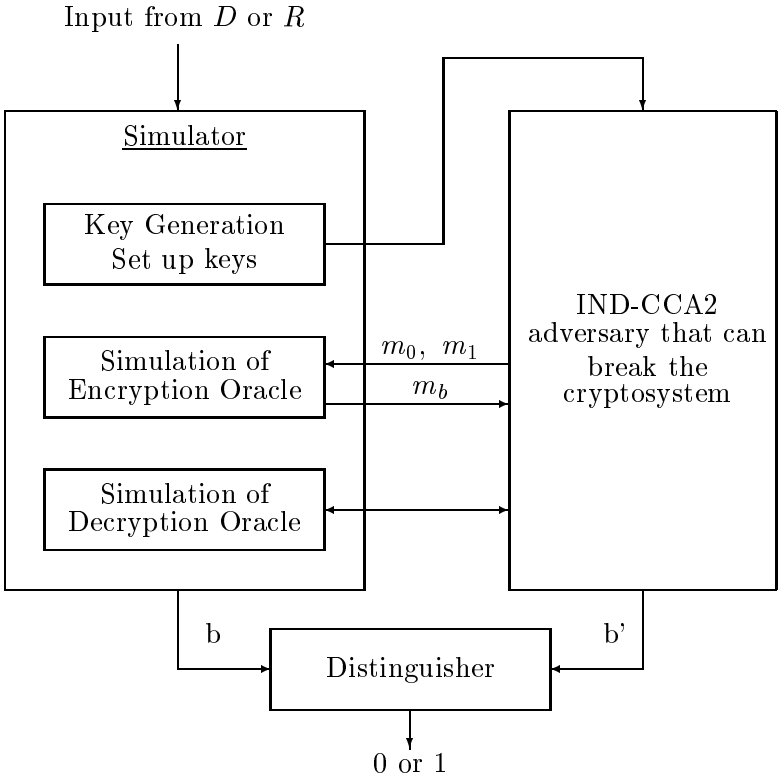

Figure-1: Graphical representation for the construction of the Secure El-Gamal proof.

Hence by observing the distribution of 0's and 1's that are output by the distinguisher, it can be determined which distribution the quadruples are coming from. If the quadruples are coming from R then 1's will occur with probability 0.5 and 0's with probability 0.5. The adversary will only be correct half the time, as it has no advantage. If the quadruples come from D then the adversary has an advantage and 1's will occur with probability $0.5 + \varepsilon$ (where ε is the adversary's non-negligible advantage) and 0's with

probability $0.5 - \varepsilon$.

Hence, by observation of the output distribution, one has a statistical test for the DDHP.

The construction of the proof is relatively simple, however there are several properties that must hold for the proof to be valid.

- The simulator must create a valid ciphertext if the quadruple comes from D and an invalid ciphertext if the quadruple comes from R.
- When the quadruple comes from D the joint distribution of the adversary's view and the random bit b must be statistically indistinguishable from that in an actual attack.
- When the quadruple comes from R the distribution of the random bit b must be (essentially) independent from the adversary's view.

Theorem 1 *Secure El-Gamal is secure against adaptive chosen ciphertext attack in the Random Oracle model assuming that the Diffie-Hellman decision problem is hard in the group $GF(p)$.*

The proof of security is for a scheme that is slight variant of the El-Gamal scheme described above, but the two schemes are interchangeable. The scheme used in the proof has an extra part to the ciphertext, c_2. A ciphertext from the El-Gamal scheme (above) can be transformed into one for this scheme (in the proof) by $(c_1, c_3) \rightarrow (c_1, c_2, c_3 \times c_2)$. The transformation back is obvious.

First the simulator is described. On input the quadruple (g_1, g_2, c_1, c_2) the simulator generates a random private key $x_R \in_R GF(p)$ and outputs the public key as $y_R = g^{x_R}$.

The simulator simulates the encryption oracle as follows. On input two messages m_0 and m_1 it selects a random bit $b \in [0, 1]$ and computes:

$$r = c_1^{x_R}$$
$$c_3 = (r \times c_2) \times (m_b || H(m_b || r))^2$$

The simulated encryption oracle outputs (c_1, c_2, c_3), where c_1 and c_2 come from the input quadruple to the simulator.

The simulator simulates the decryption oracle as follows. On input (c_1, c_2, c_3) it computes:

$r = c_1^{x_R}$
$w = \sqrt{\frac{c_3}{(rc_2)}}$ (choose the square root that yields the correct hash)
$m = w_{[1 \cdots (n-k_0)]}$

If the simulated decryption oracle outputs m, else it outputs $\emptyset$.

The aim now is to show that when the input comes from D the simulator simulates the encryption and decryption oracles perfectly (probabilistically) and the advantage of the adversary is apparent at the distinguisher. Alternatively, if the input comes from R then the output of the simulated encryption oracle will not be a valid ciphertext in the sense that .

It is also important to note that since the DDHP is hard for the adversary they cannot even find out any partial information about the secret key that could be used to determine b.

The theorem follows from the following two lemmas.

Lemma 1 - *When the simulator's input comes from D, the joint distribution of the adversary's view and the hidden bit b is statistically indistinguishable from that in the actual attack.*

In this case it is clear the output of the simulated encryption oracle has the right form as

$$c_1^{x_R} c_2 = (g_1^x)^{x_R} g_2^x = (g_1^{x_R})^x g_2^x = y_R^x g_2^x$$

which is equivalent to the output of the actual encryption oracle. Similarly, the simulated decryption oracle will accept all valid ciphertexts.

It remains to be shown that all invalid ciphertexts are rejected with overwhelming probability. If an invalid ciphertext (in the sense that $\log_{g_1} c_1 \neq \log_{g_2} c_2$) is presented as a query to the decryption oracle it will be rejected as the resulting r will not be correct for recovering m from c_3. More importantly the invalid ciphertext will not pass the check involving the random oracle (H). By using a random oracle it is ensured that the hash is completely non-malleable and no partial information is leaked.

Lemma 2 - *When the simulator's input comes from R, the distribution of the hidden bit is (essentially) independent from the adversary's view.*

First it will be shown that no partial information about b is leaked from just the challenge ciphertext, this essentially is showing IND-CPA security. Then it will be shown that there is only a negligible chance that the simulated decryption oracle gives the adversary any information about b. Since an IND-CCA2 adversary that cannot gain any information from a decryption oracle is equivalent to an IND-CPA adversary, the lemma is proven.

It has been shown that assuming DDHP the El-Gamal cryptosystem is secure in the sense of IND-CPA [3, 7]. To show the IND-CPA security of this

scheme it will be shown how to convert an El-Gamal challenge ciphertext into one for this scheme. First a second generator needs to be created, if p is of the form $p = 2q + 1$, then there are $q - 1$ generators. Hence by considering powers of g_1 a second generator of the form $g_2 = g_1^w$ can be found in polynomial time, with w known. So g_2^x can be calculated as $(g_1^x)^w$. So an El-Gamal challenge ciphertext can be transformed into a Secure El-Gamal challenge ciphertext as

$$(g_1^x, y_R^x \times m_b) \rightarrow (g_1^x, g_2^x, (y_R^x g_2^x) \times m_b).$$

It should be noted that the message is a different size to a message in an actual Secure El-Gamal challenge ciphertext. However this is not an issue, if p is an n bit prime, and the hash function outputs 128 bits, then the chances that two messages chosen at random do not differ in the first $n - 128$ bits is $\frac{n}{2} - 128$, which is negligible for suitable large n. The absence of the appended hash is irrelevant since the use of a random oracle ensures no information about m is leaked to an IND-CPA adversary. Also, without access to a decryption oracle there is no need for a correct hash value to be present in the ciphertext.

The simulated decryption oracle still needs to reject all invalid ciphertexts, otherwise relevant information will be leaked. A valid ciphertext is (c_1, c_2, c_3), an invalid one is (c_1', c_2', c_3'). There are two cases to consider.

1) $(c_3) = (c_3')$. If this happens with non-negligible probability then the random oracle must not be one way since c_1' and c_2' will create a different r (as they are different from c_1 and c_2) and this will cause decryption to a different message and hash. If the hash check passes, then the hash has been created without knowledge of the message.

2) $(c_1, c_2) = (c_1', c_2')$. With $c_3 \neq c_3'$, then the adversary has to replace the message and hash in c_3 to create c_3'. They can't just replace the message as if the hash check passes then a collision has been found. They can't replace the hash, or the message and hash, as without complete knowledge of r the correct hash cannot be calculated, and if it could then a collision could be found.

Using a random oracle means that one-wayness and collision-freeness cannot be defeated, in fact no partial information is leaked about the pre-image of the hash. Thus, the simulated decryption oracle will reject all invalid ciphertexts, except with negligible probability.

Hence if the DDHP is a computationally unfeasible problem then an IND-CCA2 attacker for Secure El-Gamal cannot exist.

4 Conclusion

This paper has shown that the one-way hash variant of the scheme by Zheng-Seberry [9] is insecure in the sense of IND against a chosen ciphertext adversary.

A new scheme was created, called Secure El-Gamal, that was shown to be provably secure in the random oracle model.

Acknowledgments

Breaking the Zheng-Seberry scheme was discovered during discussions with Associate Professor Josef Piepzyck. Also, the authors wish to thanks Dr David Pointcheval for his help in verifying the proof of Secure El-Gamal.

References

[1] M. Bellare, A. Desai, D. Pointcheval and P. Rogaway, *Relations among notions of security for public-key encryption schemes* CRYPTO'98. LNCS 1462, pg 26-45. Springer-Verlag, California, 1998.

[2] M. Bellare and P. Rogaway, *Optimal asymmetric encryption - how to encrypt with RSA* EUROCRYTP'94. LNCS 950, pg 92-111. Springer-Verlag, 1994.

[3] D. Boneh, *The decision Diffie-Hellman problem*, Third Algorithmic Number Theory Symposium (ANTS)LNCS 1423, Springer-Verlag, 1998..

[4] R. Cramer and V. Shoup, *A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack*, CRYPTO'98. LNCS 1462, pg 13-25. Springer-Verlag, California, 1998.

[5] T. El Gamal, *A public key cryptosystem and signature schme based on discrete logarithms*, IEEE Trans. Inform. Theory, 31:469-472, 1985.

[6] V. Shoup, *Using hash functions as a hedge against chosen ciphertext attack* EUROCRYPT'00. LNCS 1807, pg 275-288. Springer-Verlag, 2000.

[7] Y. Tsiounis and M. Yung, *"On the security of El-Gamal based encryption*, PKC'98. LNCS 1431, Spinger-Verlag, Japan, 1998.

[8] Y. Zheng, *Improved public key cryptosystems secure against chosen ciphertext attacks*, Technical Report 94-1, University of Wollongong, 1994.

[9] Y. Zheng and J. Seberry, *Immunizing public key cryptosystems against chosen ciphertext attacks*, IEEE Journal on Selected Areas in Communications, 1993. 11(5): p. 715-724.