Cryptanalysis of the Immunized LL Public Key Systems

Yair Frankel[1] * and Moti Yung[2]

[1] Sandia National Laboratories, Albuquerque, NM
[2] IBM T. J. Watson Research Center, Yorktown Heights, NY

Abstract. In CRYPTO '93 Lim and Lee provided a valuable investigation of public key encryption systems secure against *adaptive chosen ciphertext attacks*. In this paper we identify several insecurities of both their RSA and El Gamal based schemes. We first demonstrate that the RSA based scheme is insecure under an adaptive chosen ciphertext attack. We also point weaknesses in the design of both their RSA and El Gamal based schemes regarding the use of pseudorandom-generators, and in particular show that their choice of pseudorandom-generators for the RSA based scheme may be insecure even with respect to a *known ciphertext only attack*.

They further claim that their schemes are particularly useful in the context of group-oriented cryptosystems due to the unique verification method used. (In fact their scheme is the only group-oriented practical encryption claimed to be secure against chosen ciphertext attacks). Group oriented cryptosystems distribute the decryption process amongst a multiple of individuals in order to provide a mechanism in which no single person is trusted. We further demonstrate that both their schemes are completely insecure in this setting.

* Research was performed while the author was at GTE Laboratories Incorporated.

D. Coppersmith (Ed.): Advances in Cryptology - CRYPTO '95, LNCS 963, pp. 287-296, 1995.

1 Introduction

We analyze the Lim and Lee public key cryptosystems (which are built upon the RSA [RSA78] and El-Gamal [ElGamal85] cryptosystems) presented at CRYPTO '93 [LimLee93]. Their systems introduced new insights and were designed (1) to strengthen the basic public-key systems against adaptive chosen ciphertext attacks, (2) to improve problems with previous such schemes, and (3) to be adaptable for group-oriented systems. In this paper we:

- demonstrate that the Lim-Lee RSA based encryption scheme is insecure against an adaptive chosen ciphertext attack;
- show that the Lim-Lee El Gamal based encryption scheme is insecure against an adaptive chosen ciphertext attack when used in a group oriented setting;
- find weaknesses with regards to the usage of pseudorandom generators in the Lim-Lee schemes (in the RSA based scheme, one implementation is even insecure w.r.t. known ciphertext only attack); and
- make heuristic recommendations on potential improvements of the Lim-Lee El Gamal based encryption scheme, strengthening the bindings between ciphertext portions of the scheme.

1.1 Chosen ciphertext security

For many years, no public key system [DiffieHellman76] was shown to be secure under a chosen ciphertext attack. Indeed, Rabin's scheme [Rabin79] which was proven to be secure under chosen message attack and a dense message space was shown, in fact, to be strongly insecure with respect to chosen ciphertext attacks. Thus, the question of how to design or prove security against such attacks was open for a while.

Provably secure chosen-ciphertext secure systems :
The first solution for such systems was given in [NaorYung90] which presented a public key cryptosystem secure against chosen ciphertext attack based on zero-knowledge non-interactive proof systems and probabilistic encryption. Chosen ciphertext attack formalizes the situation where the adversary (say, an **operator** of the decryption equipment in a company) as part of the attack obtains the decryption equipment and is allowed to sequentially query it as a black box (an input-output oracle). The system is said to be secure under a chosen ciphertext attack if the attacker cannot decrypt a new message. This is also known as the "lunch break attack", "midnight-attack" or "querying-only attack". The names visualize the situation where the supervisor is out of the office (for lunch) and the attacker is using the opportunity to play with the equipment over the break but it has no meaningful ciphertext in its possession (as the supervisor has put relevant documents in a safe and suspended communication).

Strengthening this attack is possible: an "adaptive attack" is one where the attacker gets to query the decryption equipment also after receiving the meaningful ciphertext that the attacker wishes to open. However, the attacker is not

allowed to query the machine with the target ciphertext. This can be visualized as an "equipment testing available-ciphertext attack" in which the attacker (a **technician**) who now can have access to a meaningful ciphertext it tries to understand. The technician is allowed to test the equipment in front of the supervisor, yet the attacker should query the machine with other messages (which look seemingly harmless to the supervisor overlooking the testing). That is, in this attack the adversary may ask queries based on the ciphertext it wants to decrypt but not the ciphertext itself. In [DoDwNa91] a system secure against adaptive chosen ciphertext attack was given as part of "non-malleable cryptography". Also [RackoffSimon91] gave a solution to such an attack assuming a stronger setting where the sender (the technician) who queries the machine also has a personal public key known to the decryption machine.

Practical chosen-ciphertext secure systems :
The above theoretical methods employed the impractical non-interactive zero-knowledge proof systems. This motivated the study of practical immunization methods by somewhat simulating what is achieved by the theoretical methods. First [Damgard91] suggested a practical approach (against the operator adversary), then [ZhengSeberry92] claimed a system that is secure even against adaptive chosen ciphertext attacks (the technician adversary). They also showed that the scheme of [Damgard91] does not withstand such attacks. Under a very strong assumption about a cryptographic hash function being as good as a random oracle a method like theirs has been validated to be secure [BellareRogaway93]. Finally, [LimLee93] suggested public key schemes which they claimed to be secure against the (technician) adaptive attack as well. They also demonstrated some problems with the methods of [ZhengSeberry92]. Then they further suggested that their immunization technique works in the context of group-oriented cryptography [DesmedtFrankel89] where the cryptographic power to decrypt is shared amongst a multiple of agents.

Immunization methods against chosen ciphertext attacks attempt to entangle cryptographic operations and introduce dependencies among them so that it may be easy to be convinced that the party which has generated the ciphertext, indeed must have known the cleartext and thus could not have abused the system as it must have known the result of its query to begin with. This is the spirit of the theoretical work.

We would like to remark that the work by [LimLee93] certainly teaches us desirable properties of a practical chosen ciphertext public key. They do point out certain problems with previous schemes and describe how a scheme for the group-oriented case should work. In fact for this purpose they advocate a validity check which is based on the ciphertext and not on the recovered plaintext; which is a true statement. In short, we feel that their analysis is valuable. In their arguments they follow the above direction of thinking by simulating the theoretical approach. It would have seemed that it is plausible to argue that due to the computational relationships within the ciphertext gram, the attacker cannot generate something new based on the available information without actually starting from a known cleartext message. Its information being past answers to queries

and, in the adaptive attack – the available ciphertext challenge. In fact this is what [LimLee93] argue, and they further give additional claims about "semantic security" (i.e., all bits are secure based on a hard problem [GoldwMicali84]) and that certain approaches to breaking the method do not work. Nevertheless, we show that such general arguments should be taken with a grain of salt, and there are ways to abuse the system even when the queries must have portions which look quite dependent on each other and the plaintext message, and even if in addition weaker attacks seem not to be successful.

1.2 Organization of paper

In Section 2 we discuss the necessary background for the paper. The Lim-Lee RSA based scheme is cryptanalyzed in Section 3, and the El Gamal based scheme is broken in Section 4 under the group-oriented cryptography setting. In Section 5 we discuss inherent weaknesses in their use of pseudo-random generators. Potential improvements and conclusion are given in Section 6.

2 Background

Notation:
Throughout this paper we denote $|x|$ as the length of the string x and $x_1||x_2$ as the concatenation of the string x_1 with string x_2.

Let $G(x_1, x_2)$ denotes an x_1 bit output sequence produced by cryptographically strong pseudorandom generator on a seed x_2 and let $h(x)$ denote the output of a cryptographically secure hash function on a string x.

Attacks:
We consider "adaptive chosen ciphertext attacks" when the adversary is able to make queries to a decryption oracle based on but different from a target ciphertext. We say that an attack is successful if, for a large enough fraction of the cases, the attacker is able to recover the ciphertext after the attack. Namely, for a large portion and appropriate choice of system's parameters (as specified by the designer), the attacker can be successful. A strong attack, is one that every ciphertext (not only a fraction of them) can be recovered.

Group-Oriented Cryptography:
Let us give a brief description of group-oriented cryptosystems in our public key decryption setting. Group oriented cryptosystems, also called threshold cryptosystems, were developed as a method to distribute the decryption process amongst a multiple of individuals in order to provide a decryption mechanism in which no single person is trusted [Desmedt87, DesmedtFrankel89]. In such a scheme, the decryption key is initially distributed to a number of agents and each agent is given a partial key. When an encryption is received by the group, a quorum of agents is available to apply its partial keys. The results of the partial keys applied to the ciphertext are called partial results. These partial results are then combined by a polynomial time algorithm into the final decryption result.

No security is placed in the combining function (it can be a public server or any available agent).

As long as a less than a threshold of agents is attacked by an intruder that reads their memory the system remains secure (the intruder does not have access to a threshold of the partial keys). A successful chosen ciphertext attack on a group oriented cryptosystem allows the adversary access to the combining function (as it is a public function with no security requirements put on it) and access to less than a threshold of the partial keys. Then, the attacker is given the allowed information, he is allowed queries, and it attempts to break the system. Lim and Lee provide important insight on why group oriented cryptosystems are particularly prone to chosen ciphertext attacks [LimLee93].

3 Breaking the Lim-Lee RSA based scheme

First let us review the details of their RSA based scheme.

For user A let e_A be the public exponent, N_A be the public modulus and d_A be the private exponent of the RSA encryption scheme [RSA78].

Encryption:

The ciphertext for a message m is $C = (c_0, c_1, c_2)$ where:

- $c_1 = s^{3 \cdot e_A} \bmod N_A$ where $s \in_R \{1, \ldots N_A - 1\}$
- $c_2 = z \oplus m$ where $z = G(|m|, s)$
- $c_0 = s^{3 \cdot H} \bmod N_A$ where $H = h(c_1 || c_2)$

Decryption:

To decrypt $C = (c_0, c_1, c_2)$.

- Verify that $c_0^{3 \cdot e_A} = c_1^{H'}$ where $H' = h(c_1 || c_2)$; return NULL if false
- Compute $z = G(|c_2|, s)$ where $s \equiv c_1^{(1/3) \cdot d_A} \bmod N_A$
- Return $z \oplus c_2$

Note that the ciphertext was generated by encrypting the seed s, employing the seed for encryption (stream cipher) and encrypting a one-way integrity certificate of the two fields. This seems to be a strong binding of the seed and the cleartext.

We now break the Lim-Lee RSA based encryption scheme.

Theorem 1. *The Lim-Lee RSA based encryption scheme can be broken using an adaptive chosen ciphertext attack.*

Proof. We first describe the attack and then show that the adversary is successful with sufficient probability.

Observe using the Extended Euclidean Algorithm that there exist integers u, v such that $e_A \cdot u + H \cdot v = \gcd(e_A, H)$. Since e_A is public and H can be

computed from publicly known values c_1 and c_2, the values of u and v can be generated in polynomial time. Then if $\gcd(e_A, H) = 1$, the attacker is successful and it can compute $s^3 \equiv c_1^u \cdot c_0^v \bmod N_A$. The attacker, when successful, can simply send the oracle the ciphertext $C' = (s^{3 \cdot H''}, c_1, \mathbf{0})$ where $H'' = h(c_1 || \mathbf{0})$ and $\mathbf{0}$ is a string of all zero bits of length $|c_2|$. To complete the attack notice that the oracle will pass the verification step and return $z = G(|c_2|, s)$.

It is now shown that the adversary succeeds with a large enough fraction of the ciphertexts. Using [Apostol76, Thm 3.9] with probability about $\frac{6}{\pi^2}$ the $\gcd(e_A, H) = 1 \equiv 1 \bmod \phi(N_A)$. In this estimate we assumed that the public exponent has been chosen at random and that H is almost random, since h as a strong cryptographic hash function – has an almost-random output distribution. In case when e_A is chosen as a constant (e.g., 3 is a popular choice for a globally chosen exponent) then the probability of a random number in the range being relatively prime can be estimated as the constant $\frac{\phi(e_A)}{e_A}$ (e.g., $\frac{2}{3}$ for exponent 3) as the range of the function h is quite large. □

4 Attacks on the El Gamal based scheme

First let us review the details of the Lim-Lee El Gamal based scheme.

Let α be a generator for GF(p) where p is a large prime. For user A let $y_A \equiv \alpha^{x_A}$ be the public exponent, and x_A be the private key of the El Gamal encryption scheme [ElGamal85].

Encryption:

The ciphertext for a message m is $C = (c_1, c_2, c_3, c_4)$ where:

- $c_0 \equiv \alpha^{r_0}$ where $r_0 \in_R \{1, \ldots, p-1\}$
- $c_1 \equiv \alpha^{r_1}$ where $r_1 \in_R \{1, \ldots, p-1\}$
- $c_2 = z \oplus m$ where $z = G(|m|, y_A^{r_1} c_0)$
- $c_3 = h(c_0 || c_2)$
- $c_4 \equiv r_0 + c_3 r_1 \bmod p - 1$.

Decryption:

To decrypt $C = (c_1, c_2, c_3, c_4)$.

- Verify that $c_3 = h(c_0' || c_2)$ where $c_0' \equiv \alpha^{c_4} c_1^{-c_3}$; return NULL if false
- Compute $z = G(|c_2|, s)$ where $s = c_1^{x_A} c_0' \bmod p$
- Return $z \oplus c_2$

4.1 Cryptanalyzing the group-oriented setting

Lim and Lee discuss the case of using this scheme for group-oriented cryptography. In this case, we secretly share the private exponent and we need a secret sharing scheme to do it. Thus, what is generally needed is to make sure that the generator works in a field (or module), so assume $p - 1 = wq$ where w is

small and q is a prime and choose the generator α above to be in GF(q), similarly the choice of exponents r_0 and r_1 (see [DesmedtFrankel89]). Otherwise, the description stays the same as above.

With the group-oriented setting the decryption agents verify the hash value (the first step in the decryption process) individually, before they jointly compute s. This is an important property for group-oriented decryption as [LimLee93] suggest. To compute s, an agent uses its share of the private key x_A and c_1 to get its partial result. A threshold of partial results can be interpolated to get the result $c_1^{x_A}$. Now s can be computed given c_0'. If the verification process is successful with a different ciphertext, then a random seed unrelated to s should be generated. We demonstrate that this is not the case and in fact one can attack the system.

Theorem 2. *The Lim-Lee El Gamal based encryption scheme in the group-oriented setting can be strongly broken using an adaptive chosen ciphertext attack.*

Proof. We only need to prove that an attacker can generate an s' related to the s of a given ciphertext and recover s itself.

The attacker can generate a ciphertext $C' = (c_1' \equiv c_1 \cdot \alpha \bmod p, c_2, c_3, c_4' \equiv r_0 + c_3 r_1 + c_3 \bmod q)$ and provide it to the oracle. Observe that the verify step will be accepted since c_2 remains the same and $c_0' \equiv c_0 \bmod p$. The seed generated in the decryption step will be $s' = (\alpha^{r_1+1})^{x_A} c_0' \equiv (\alpha^{r_1+1})^{x_A} c_0$. Thus $s \equiv s'(\alpha^{x_A})^{-1} \bmod p$.

To finish the proof we note that one agent (or combiner server) will receive the information needed in order to calculate s' and therefore can generate s. That person (server) can generate C' and be successful in performing the adaptive chosen ciphertext attack. □

5 Cautions on the use of pseudorandom generators

Pseudorandom number generators do not necessarily exhibit the strength to withstand attacks, when correlated seeds can be generated or when the seed is poorly encrypted. Lim and Lee suggest to use generators like [AleChGoSch88] as possible pseudorandom generators. They specifically say that this is a good choice since these generators rely on the same (RSA) assumption as the cryptosystems and there is no need for further tools. They claim their system is semantically secure (that is, all encrypted bits are secure). However we show that the opposite is true and, in fact, this choice may be quite bad due to an algebraic interplay between the generator and the insecurity of the seed encryption.

Claim 1 *The Lim-Lee RSA based scheme when used with the RSA based pseudorandom generator of [AleChGoSch88] with exponent being 3, is insecure with respect to a "known ciphertext" only attack.*

Proof. From Theorem 1 an attacker can determine s^{3g} where $g = \gcd(e_A, H)$. Therefore, as argued, the attacker is able to determine, s^3 with good probability. Now using [AleChGoSch88] with the RSA function with exponent 3, the pad used in the encryption comprised of the concatenation of least significant bits of s, $s^3 \bmod n$, $(s^3)^3 = s^9 \bmod n$, $(s^9)^3 = s^{27} \bmod n$, and so on. Knowing s^3, in turn, gives all the bits but the first one, since exponentiation is easy! □

Note again that the attack above did not employ any access to the device as an oracle, it was a pure known ciphertext attack. Encryption exponent being 3 (as in the generator) is a popular choice, since it provides the fastest modular exponentiation for RSA.

In addition we can also point out that Theorem 2 tells us that an attacker can generate outputs from the decryption functions based on seeds related to the original seed used to encrypt the meaningful ciphertext (i.e., the original C being attacked). Therefore we deduce that:

Corollary 3. *The Lim-Lee El Gamal based encryption scheme has a weakness when attacked by an adaptive chosen ciphertext attack – the attacker can correlate the seeds of the various ciphertexts.*

6 Conclusion and Discussion

The work of [LimLee93] discusses immunized public key encryption systems. It shows problems with previous such systems, and carefully discusses original requirements and designs of the first group-oriented such systems. We found their design goals, criticism of previous systems, discussions, methodologies, and some of their efficient techniques highly valuable.

On the other hand, we presented cryptanalysis of and security problems with the actual schemes of [LimLee93]. We showed that one can generate relations in a meaningful way even if it does not seem so or even if the ciphertext gram is partially authenticated by a hash function. We showed that in the case of group-oriented cryptography, the availability of partial results in the combiner may enable strong attacks (in the original schemes no trust was put in any single point and surely not in the combining stage). We also showed potential problems with the usage of a single key and the same algebraic problem for more than one task in a cryptosystem.

6.1 Potential improvements to the scheme

It seems difficult to modify the Lim-Lee El Gamal based scheme so that it is strong as the Diffie-Hellman problem due to the value c_4 in the ciphertext. That is, due to the way c_4 is generated, we do not see how opening up ciphertext C reduces to breaking the Diffie-Hellman problem.

We make the following preliminary suggestions to improve their El Gamal scheme.

(1) The function G should be a pseudorandom function rather than a pseudorandom number generator where s comprises of two values: a function index and an input value for that function. (In practice it is a block cipher operation). This, heuristically, will reduce the relation between ciphertext grams, but it may make the computations less efficient, though. In fact a random function based on generators built on a totally different algebraic problem than the seed encryption is preferable (e.g., one based on DES or triple-DES) to disentangle potential algebraic dependencies that may ease cryptanalysis.

(2) The value c_3 should be $h(c_0||c_1||c_2)$ instead of $h(c_0||c_2)$; it is preferable that again this hash function should not be based on the same algebraic problem as the other encryption mechanisms. The type of bindings in the resulting "ciphertext gram" above has been recently formalized and a suggestion for immunized RSA based system has been shown as well in [FranklinReiter95].

(3) At some point in their work, Lim and Lee suggest to also include in the preimage of h in c_3 information like date and other related information under the hash function. We remark that, in doing so, one has to be careful not to introduce unstructured redundancy which may enable potential "birthday" attack on this ciphertext validation component by playing with the additional information.

(4) Finally, when claiming security against chosen ciphertext attacks, the claims that the system is semantically secure and that it seems that the sender knows the cleartext from the ciphertext structure are not enough. Also, having a key perform different functions may be dangerous. We feel that the following two things better be done carefully when designing a chosen-ciphertext system based on heuristics. First, it may be a good idea to attempt to prove chosen ciphertext security under certain (strong, if not necessary known or widely assumed) properties of the tools used. Second, it is also useful to extensively characterize how the attacker may have produced the ciphertext or what may be easily computed from it, and given a ciphertext what ciphertexts can be computed adaptively from it (as the quality of the system against adaptive chosen ciphertext attacks relates directly to how the designers capture this fact in their arguments).

A natural open problem is designing a chosen ciphertext system which is practical and proven secure (e.g., as secure as El Gamal or RSA).

References

[AleChGoSch88] W. Alexi, B. Chor, O. Goldreich and C. P. Schnorr, *RSA and Rabin functions: certain parts are as secure as the whole*, SIAM J. Computing vol. 17 (2). 1988.

[Apostol76] T. M. Apostol *Introduction to analytic number theory*, Springer-Verlag, New York, 1976.

[BellareRogaway93] M. Bellare and P. Rogaway, *Random Oracles are Practical: a paradigm for designing efficient protocols*, ACM, 1-st Comp. and Com. Sec. 1993.

[Damgard91] I. Damgård, *Towards practical public key cryptosystems secure against chosen ciphertext attacks*, Advances in Cryptology–Proc. of Crypto '91.

[Desmedt87] Y. Desmedt, *Society and group oriented cryptography: a new concept*, Advances in Cryptology, Proc. of Crypto '87, Springer-Verlag, 1988.

[DesmedtFrankel89] Y. Desmedt and Y. Frankel, *Threshold cryptosystems*, Advances in Cryptology, Proc. of Crypto '89 Springer-Verlag, 1990.

[DiffieHellman76] W. Diffie and M. Hellman, *New Directions in Cryptography* , IEEE Trans. on Information Theory 22 (6), 1976, pp. 644-654.

[DoDwNa91] D. Dolev, C. Dwork and M. Naor, *Non-Malleable Cryptography*, Proc. of the 23rd Annual ACM Symposium on the Theory of Computing, 1991, pp. 542–560.

[ElGamal85] T. El Gamal, *A Public key cryptosystem and a signature scheme based on discrete logarithm*, IEEE Trans. on Information Theory 31, 465-472, 1985.

[FranklinReiter95] M.K. Franklin and M.K. Reiter, *Adaptive Chosen Ciphertext Security for RSA from G-Q Signatures*, Preliminary manuscript.

[GoGoMi86] O. Goldreich S. Goldwasser and S. Micali, *How to Construct Random Functions* , J. of the ACM 33 (1986), pp. 792-807.

[GoldwMicali84] S. Goldwasser and S. Micali, *Probabilistic Encryption*, J. Com. Sys. Sci. 28 (1984), pp 270-299.

[LimLee93] C. H. Lim and P. J. Lee, *Another method for attaining security against adaptive chosen ciphertext attacks*, Advances in Cryptology–Proc. of Crypto '93.

[NaorYung90] M. Naor and M. Yung, *Public-key cryptosystem provably secure against chosen ciphertext attack*, Proc. of the 22nd Annual Symposium on the Theory of Computing, 1990, pp. 427–437.

[Rabin79] M. O. Rabin, *Digital Signatures and Public Key Functions as Intractable as Factoring*, Technical Memo TM-212, Lab. for Computer Science, MIT, 1979.

[RackoffSimon91] C. Rackoff, and D. Simon, *Non-Interactive Zero-Knowledge Proof of Knowledge and Chosen Ciphertext Attacks*, Advances in Cryptology– Proc. of Crypto '91.

[RSA78] R. Rivest, A. Shamir and L. Adleman, *A Method for Obtaining Digital Signature and Public Key Cryptosystems*, Comm. of ACM, 21 (1978), pp 120-126.

[ZhengSeberry92] Y. Zheng and J. Seberry, *Immunizing public key cryptosystems against chosen ciphertext attacks*, IEEE Jour. on Selected Areas in Communications, 11(5), 1993, pp. 715-724. (Also in: Advances in Cryptology–Proc. of Crypto '92).