The Cramer-Shoup Public-Key Cryptosystem

Tuesday 25.4.2006

Aleksi Hänninen

Based on a book:

Wenbo Mao: Modern cryptography : theory and practice

Upper Saddle River, NJ : Prentice Hall PTR, cop. 2004
ISBN: 0-13-066943-1

f-OAEP vs. Cramer-Shoup

- Cramer-Shoup has efficient ”reduction to contradiction”
  - vs. square reduction of f-OAEP
- The intractability assumptions are minimal – namely: DDH
  - vs. ROM (there exists none) + RSA Assumption 8.3
- Efficient reduction and weak intractability assumptions are desirable properties

DDH assumption

- In group G, given (g, g^a, g^b, g^c).
  - There is no polynomially bounded algorithm to answer question ”Is ab = c (mod #G)?” with nonneglible Adv.
  - Means that if you have polynomially bounded time, your answers are about 50% right.
- In here (later):
  - #G=q, $g = g_1, g^a = g_2 = g_1^w, g^b = u_1 = g_1^{r_1}, g^c = u_2 = g_2^{r_2} = g_1^{w r_2}$
  - $(g_1, g_2, u_1, u_2) = (g_1, g_1^w, g_1^{r_1}, g_1^{wr_2})$
    - Q: is $r_1 = r_2$ (mod q)?
      - (iff $w*r_1=w*r_2$ and gcd(w,q)=1)
- DDH implies DL -problem: ”find i such that g^i=x (mod q)” is hard

Algorithm – Key Parameters

- G abelian group of large prime order q
  - Every $g \in G \neq 1$ is generator of G (Corollary 5.3)
- Two random elements $g_{1,} g_2 \in_U G$
- Five random integers $x_{1,} x_{2,} y_{1,} y_{2,} z \in [0, q)$
- Three elements $c \leftarrow g_1^{x_1} g_2^{x_2}, d \leftarrow g_1^{y_1} g_2^{y_2}, h \leftarrow g_1^z$
- A cryptographic hash function $H : G^3 \rightarrow [0, q)$
- $(g_{1,} g_{2,} c, d, h, H)$ is public key
- $(x_{1,} x_{2,} y_{1,} y_{2,} z)$ is private key
  - Because public key is made from private by exponentiating known g_1, g_2, private key is secure due to DL assumption, which is weaker than DDH.

Algorithm – Key Setup

- Pick two random $g_{1,} g_2 \in_U G$
- Pick five random integers $x_{1,} x_{2,} y_{1,} y_{2,} z \in [0, q)$
- Compute $c \leftarrow g_1^{x_1} g_2^{x_2}, d \leftarrow g_1^{y_1} g_2^{y_2}, h \leftarrow g_1^z$
- Choose a cryptographic hash function $H : G^3 \rightarrow [0, q)$
- $(g_{1,} g_{2,} c, d, h, H)$ is public key
- $(x_{1,} x_{2,} y_{1,} y_{2,} z)$ is private key

Algorithm – Encryption & Decryption

- Bob encrypts message m by
  - Pick random $r \in [0, q)$
  - $u_1 \leftarrow g_1^r, u_2 \leftarrow g_2^r, e \leftarrow h^r m, \alpha \leftarrow H(u_{1,} u_{2,} e), v \leftarrow c^r d^{r\alpha}$
  - (u_1, u_2, e, v) is the encrypted message
- Alice performs decryption of (u_1, u_2, e, v) by:
  - $\alpha \leftarrow H(u_{1,} u_{2,} e)$
  - Output:
    - $m \leftarrow e / u_1^z$, if $u_1^{x_1 + y_1 \alpha} u_2^{x_2 + y_2 \alpha} = v$
    - REJECT otherwise

Algorithm – Encryption & Decryption

- Bob: $u_1 \leftarrow g_1^r, u_2 \leftarrow g_2^r, e \leftarrow h^r m, \alpha \leftarrow H(u_{1,} u_{2,} e), v \leftarrow c^r d^{r\alpha}$
- Alice: $m \leftarrow e / u_1^z$, if $u_1^{x_1 + y_1 \alpha} u_2^{x_2 + y_2 \alpha} = v$
- If message is not altered en route to Alice, message is not rejected
  - $u_1^{x_1 + y_1 \alpha} u_2^{x_2 + y_2 \alpha} = u_1^{x_1} u_2^{x_1} u_1^{y_1 \alpha} u_2^{y_2 \alpha} = g_1^{r x_1} g_2^{r x_1} g_1^{r y_1 \alpha} g_2^{r y_2 \alpha} = (g_1^{x_1} g_2^{x_1})^r (g_1^{y_1} g_2^{y_2})^{r\alpha} = c^r d^{r\alpha} = v$
  - $e / u_1^z = \frac{h^r m}{u_1^z} = g_1^{rz} \frac{m}{g_1^{rz}} = m$
  - Process is ok

Algorithm – Notions

- Part (u_1, e) is the very same of semantically secure ElGamal cryptosystem
- Therefore IND-CPA secure if the DDH assumption holds by Theorem 14.2
- Hash function helps to provide IND-CCA2 by offering data-integrity validating step

Algorithm - Performance

- Public key consists of five elements in G
  - vs. two of ElGamal
- The size of ciphertext is quadruple
  - Twice that of ElGamal
- Encryption requires 4 and decryption 2 exponentiations
  - Increased from two of encryption and one of decryption of ElGamal

Proof of security

- Proof is (linear) reduction to contradiction
  - Reducing a hard problem supported by the underlying intractability assumption to an alleged IND-CCA2 attack
- Hard problem is the DDH problem
- If Cramer-Shoup is not secure in IND-CCA2 mode, then DDH -problem can be solved
- **D** is the set of Diffie-Hellman quadrubles
  - All quadrubles $(g_1, g_2, u_1, u_2) = (g_1, g_1^{w}, g_1^{r_1}, g_1^{wr_2})$ for which $r_1 = r_2 \pmod{q}$

Proof of security

- Suppose an attacker $\mathcal{A}$ can break Cramer-Shoup
- Then Simon, given (g_1, g_2, u_1, u_2), can construct challenge ciphertext C*, which encrypts one of messages m_0, m_1 given by $\mathcal{A}$ and asks $\mathcal{A}$ to release its attacking advantage
  - If $(g_1, g_2, u_1, u_2) \in \mathbf{D}$, C* is valid Cramer-Shoup ciphertext
    - In this case, $\mathcal{A}$ can use its attacking advantage
  - If not, then message m_b is encrypted in Shannon's information-theoretically secure sense and thus can not be deciphered
    - $\mathcal{A}$ can not have any advantage whatsoever!
- If $\mathcal{A}$ has about 50% right, quadruble is probably not in $\mathbf{D}$

Proof of security – setup

- First, (g_1, g_2, u_1, u_2) is given to Simon
- He picks $x_1, x_2, y_1, y_2, z_1, z_2$ from [0,q)
- And computes $c \leftarrow g_1^{x_1} g_2^{x_2}, d \leftarrow g_1^{y_1} g_2^{y_2}, h \leftarrow g_1^{z_1} g_2^{z_2}$
- Implicit private key is $(x_1, x_2, y_1, y_2, z_1, z_2)$
  - z is not explicitly expressed, but is uniquely determined since $g_2 = g_1^w, g_1^{z_1} g_2^{z_2} = g_1^{z_1} g_1^{w z_2} = g_1^{z_1 + w z_2} = g_1^z$
  - It is possible to cipher and decipher with this impicit information (z_1, z_2)

Proof of security – the challenge ciphertext

- Simon gets m_0 and m_1 from $\mathcal{A}$ and tosses a fair coin and gets b.
- He computes $e = u_1^{z_1} u_2^{z_2} m_b, \alpha = H(u_1, u_2, e), v = u_1^{x_1 + y_1 \alpha} u_2^{x_2 + y_2 \alpha}$
- The challenge ciphertext is C*=(u_1, u_2, e, v)
  - ” But usually e=$h^r m_b$!?? ”
  - This is the trick!
- **If** (g_1, g_2, u_1, u_2) ∈**D**, there exist r such that $u_1 = g_1^r$, $u_2 = g_2^r$
  - $u_1^{z_1} u_2^{z_2} = (g_1^r)^{z_1} (g_2^r)^{z_2} = (g_1^{z_1} g_2^{z_2})^r = h^r$
  - Simulated encryption of (g_1, g_2, u_1, u_2) is valid
  - So $\mathcal{A}$ should know b with positive Adv

Proof of security – the challenge ciphertext

- **Else** as far as $\mathcal{A}$ is considered, C* could be from either one.
- Let's analyze what $\mathcal{A}$ can calculate and form equations

$$g_1^{z_1} g_2^{z_2} = h$$
$$g_1^{z_1 r_1} g_2^{z_2 r_2} = e/m_i \quad \rightarrow$$
for each m_i

$$\begin{pmatrix} 1 & \log_{g_1} g_2 \\ r_1 & r_2 \log_{g_1} g_2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \log_{g_1} h \\ \log_{g_1}(e/m_0) \end{pmatrix} (mod\ q)$$

$$\begin{pmatrix} 1 & \log_{g_1} g_2 \\ r_1 & r_2 \log_{g_1} g_2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \log_{g_1} h \\ \log_{g_1}(e/m_1) \end{pmatrix} (mod\ q)$$

- Matrix on the left hand side is invertible
  - $Det\ M = (r_2 - r_1) \log_{g_1} g_{2,}\ r_1 \neq r_{2,}\ g_2 \neq g_1 \rightarrow \log_{g_1} g_2 \neq 0$
  - So two different implicit private key information (z_1, z_2) can be found, one for m_0 and one for m_1, but both are equally likely!

Proof of security – the challenge ciphertext

- C* encrypts m_b in Shannon's information-theoretical security sense
  - 2 cipher texts, 2 plain texts, equal probability both
- $\mathcal{A}$ does not have any advantage so m_b is absolutely secured
- Q: $(g_1, g_2, u_1, u_2) \in \mathbf{D}$ **?**
- Simon answers: YES if $\mathcal{A}$ was right, NO if $\mathcal{A}$ was not.
  - This is how he gets same Adv as $\mathcal{A}$ when Q is true
  - Then Simon's total Advantage is a half of $\mathcal{A}$'s Advantage (see lecture 6, page 24)

Theorem 15.1

- Let $(g_{1,} g_{2,} c , d , h , H)$ be a public key for the Cramer-Shoup encryption scheme in a group G of a prime order q, where $g_1 \neq 1$ and $g_2 \neq 1$. If $(g_1, g_2, U_1, U_2) \notin \mathbf{D}$ then the probability of successfully solving the following problem is bounded by $\frac{1}{q}$.
  - Input: public key $(g_{1,}\ g_{2,}\ c ,\ d ,\ h ,\ H),\ (U_{1,}\ U_{2,}\ E) \in G^3$
  - Output: V st. $(U_{1,}\ U_{2,}\ E ,\ V)$ is a valid ciphertext deemed by the key owner

- *Note: in here, the problem of finding correct ciphertext is simplified as to give V from the three other. As all other are inputs of the hash function H forming α and V is not, the easiest way is to deduce V from the other three.*

Theorem 15.1

- What can be known from the input?
  - V must satisfy $U_1^{x_1+y_1\alpha} U_2^{x_2+y_2\alpha} = V$
  - From the construction of public key components c and d
    $g_1^{x_1} g_1^{wx_2} = c, \quad g_1^{y_1} g_1^{wy_2} = d$
  - Other information of the (x_1, x_2, y_1, y_2) is not available.

$$\rightarrow \begin{pmatrix} 1 & 0 & w & 0 \\ 0 & 1 & 0 & w \\ r_1 & r_1\alpha & wr_2 & wr_2\alpha \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} \log_{g_1} c \\ \log_{g_1} d \\ \log_{g_1} V \end{pmatrix} (mod\ q) \quad (15.3.9)$$

Theorem 15.1 - continued

- After Gaussian elimination matrix has the following form:

$$\begin{pmatrix} 1 & 0 & w & 0 \\ 0 & 1 & 0 & w \\ 0 & 0 & w(r_2 - r_1) & w(r_2 - r_1)\alpha \end{pmatrix}$$

- $Det\, M \neq 0$, because $r_1 - r_2 \neq 0, w \neq 0$
- Thus (15.3.9) has (non-unique) solutions for each of V.
- So $\mathcal{A}$ cannot set the V unambiguously!
  - Every element of G (q elements) can be V fulfilling everything which A knows of the secret key!
  - Only one is correct, thus $\frac{1}{q}$ probability of correct V

Proof of security – cryptanalysis training courses

- We have not considered the cryptanalysis training course!
- When Simon gets C = (U_1, U_2, E, V) from $\mathcal{A}$, Simon will conduct the data-integrity validating procedure, checking if $u_1^{x_1+y_1\alpha} u_2^{x_2+y_2\alpha} = v$
- If message is not rejected, Simon computes $m = E/(U_1^{z_1} U_2^{z_2})$
- 3 different cases:
  - C for which (g_1, g_2, U_1, U_2) $\in \mathbf{D}$
  - C such that it is rejected
  - C for which (g_1, g_2, U_1, U_2) $\notin \mathbf{D}$ and which is not rejected

Proof of security – cryptanalysis training courses

- What if $\mathcal{A}$ send ciphertext C for which $(g_1, g_2, U_1, U_2) \in \mathbf{D}$?
- So there exist R st.

$$g_1^R = U_1, g_2^R = U_2 \quad \rightarrow \quad U_1^{z_1} U_2^{z_2} = g_1^{R z_1} g_2^{R z_2} = (g_1^{z_1} g_2^{z_2})^R = h^R$$

- Simulated decryption is correct!
- **And** no new information is revealed from z_1 and z_2
  - Because triplet (U_1, U_2, h^R) connects them similarly to (g_1, g_2, h), only the R exponent is more.
- No use of sending this kind of messages

Proof of security – cryptanalysis training courses

- What if $\mathcal{A}$ sends C such that it is rejected?
  - If C is rejected, A knows that $u_1^{x_1+y_1\alpha} u_2^{x_2+y_2\alpha} \neq v$
  - If three of x_1, y_1, x_2, y_2 are known, still the last one can't be easily determined due to DL assumption.
- What if $\mathcal{A}$ sends C for which $(g_1, g_2, U_1, U_2) \notin \mathbf{D}$ and which is not rejected?
  - Due to Theorem 15.1, this is with probability $\frac{1}{q}$!
  - $\mathcal{A}$ could as well guess correctly anything since G *is* of size q
- **All in all**, no profit from the cryptanalysis training courses!