Public-key Encryption and the El Gamal and RSA Encryption Schemes

Michele Ciampi

Introduction to Modern Cryptography, Lecture 14
Introduction

- The introduction of Public-Key Encryption (PKE) marked a revolution in cryptography.
- Parties can communicate securely without having agreed on any secret information in advance!
- One party (receiver) generates a pair of keys \((pk, sk)\) where \(pk\) is the public key and \(sk\) is the private key.
- \(pk\) is used by a sender to encrypt a message. The receiver uses \(sk\) to decrypt the ciphertext.
Introduction

Public key distribution

- Alice can send $pk$ to Bob over an authenticated channel, when she learns that Bob wants to communicate with her.
- Alice generates $(pk, sk)$ in advance and disseminates $pk$ by publishing it on her webpage or placing it in a public directory.
- **Public-key Infrastructure**: a trusted certificate authority issues certificates (signatures) for everyone’s public key.
- In this lecture, we assume that senders are able to obtain a legitimate copy of the receiver’s public key.
Comparison to Private-Key Encryption

- In public-key encryption, only the secrecy of the private key $sk$ is required.
- In public-key encryption different keys are used for encryption and decryption (asymmetry). The roles of the sender and the receiver are not interchangeable.
- Multiple senders can communicate privately with a single receiver.
- Public-key encryption is significantly slower than private-key encryption and implementing it for resource-constrained devices like smartcards can be a challenge.
Definition
A public-key encryption scheme is a triple of polynomial-time algorithms (Gen, Enc, Dec) such that:

1. The key-generation algorithm Gen on input $1^n$ outputs a pair of keys $(pk, sk)$, where $pk$ is the public key and $sk$ is the private key.

2. The encryption algorithm Enc on input $pk$ and a message $M$ from some message space (that may depend on $pk$) outputs a ciphertext $c$. We write this as $c \leftarrow \text{Enc}_{pk}(M)$.

3. The decryption algorithm Dec on input $sk$ and a ciphertext $c$ outputs a message $M$ or a special symbol $\bot$ denoting failure. We write this as $M := \text{Dec}_{sk}(c)$. 


Correctness

For any message $M$, we have that $\text{Dec}_{sk}(\text{Enc}_{pk}(M)) = M$ except with negligible probability over $(pk, sk)$ output by $\text{Gen}(1^n)$. 

Security against Chosen-Plaintext Attacks

Given a public-key encryption scheme $\Pi = (\text{Gen}, \text{Enc}, \text{Dec})$ and an adversary $A$ consider the following experiment:

**The eavesdropping indistinguishability experiment**
$\text{PubK}_{A,\Pi}(n)$:

1. $\text{Gen}(1^n)$ is run to obtain $(pk, sk)$.

2. The adversary $A$ is given $pk$, and outputs a pair of equal-length messages $M_0, M_1$ in the message space.

3. A uniform bit $b \in \{0, 1\}$ is chosen and then a ciphertext $c \leftarrow \text{Enc}_{pk}(M_b)$ is computed and given to $A$. We call $c$ the challenge ciphertext.

4. $A$ outputs a bit $b'$. The output of the experiment is 1 if $b = b'$ and 0 otherwise. If $b = b'$, we say that $A$ succeeds.
Security against Chosen-Plaintext Attacks

Definition

A public-key encryption scheme $\Pi = (\text{Gen}, \text{Enc}, \text{Dec})$ has *indistinguishable encryptions under a chosen plaintext attack*, or it is *CPA-secure*, if for every PPT adversary $A$ it holds that

$$\Pr \left[ \text{PubK}^{eav}_{A,\Pi}(n) = 1 \right] \leq \frac{1}{2} + \text{negl}(n).$$
In 1985, Taher El Gamal observed that the Diffie-Hellman (DH) protocol could be adapted to give a public-key encryption scheme.

In DH protocol, Alice and Bob derive a shared key $k$ which is indistinguishable from a uniform element of a group $\mathbb{G}$.

Bob may use this shared key to encrypt a message $M \in \mathbb{G}$ by sending $k \cdot M$ to Alice.

Alice can recover $m$ (she knows $k$), while an eavesdropper learns nothing about $M$. 
El Gamal encryption

Let $G$ be a group generation algorithm that on input $1^n$ outputs a description of a cyclic group $G$, its order $q$, and a generator $g$.

- **Gen**: on input $1^n$ run $G(1^n)$ to obtain $(G, q, g)$. Then choose a uniform $x \in \mathbb{Z}_q$ and compute $h = g^x$. The public key is $pk = (G, q, g, h)$ and the private key is $sk = (G, q, g, x)$. The message space is $G$.

- **Enc**: on input a public key $pk$ and a message $M \in G$, choose a uniform $y \in \mathbb{Z}_q$ and output the ciphertext

  $$\langle g^y, h^y \cdot M \rangle.$$

- **Dec**: on input a private key $sk$ and a ciphertext $\langle c_1, c_2 \rangle$, output

  $$\hat{M} := c_2/c_1^x.$$

**Figure**: The El Gamal encryption scheme.
Correctness of El Gamal

Let $\langle c_1, c_2 \rangle = \langle g^y, h^y \cdot M \rangle$ with $h = g^x$. Then

$$\hat{M} := \frac{c_2}{c_1^x} = \frac{h^y \cdot M}{(g^y)^x} = \frac{(g^x)^y \cdot M}{g^{xy}} = \frac{g^{xy} \cdot M}{g^{xy}} = M.$$
Lemma

Let $G$ be a finite group and an arbitrary $M \in G$. Then choosing uniform $k \in G$ and setting $k' := k \cdot M$ gives the same distribution for $k'$ as choosing uniform $k' \in G$. Put differently, for any $\hat{g} \in G$

$$\Pr[k \leftarrow G : k \cdot M = \hat{g}] = \frac{1}{|G|}.$$  

Proof.

Let $\hat{g} \in G$ be arbitrary. Then

$$\Pr[k \leftarrow G : k \cdot M = \hat{g}] = \Pr[k \leftarrow G : k = \hat{g} \cdot M^{-1}] .$$

Since $k$ is uniform, the probability that $k$ is equal to the fixed element $\hat{g} \cdot M^{-1}$ is $\frac{1}{|G|}$. \hfill \qed
Security of El Gamal

**Theorem**

If the DDH problem is hard relative to $G$, then the El Gamal encryption scheme is CPA-secure.

**Proof.** Let $\Pi$ denote the El Gamal encryption scheme. Let $A$ be a PPT adversary. We want to show that

$$\Pr \left[ \text{PubK}_{A,\Pi}(n) = 1 \right] \leq \frac{1}{2} + \text{negl}(n).$$

Consider the modified “encryption scheme” $\tilde{\Pi}$, where Gen is as $\Pi$ but the encryption of $M$ with respect to the public key $\langle G, q, g, h \rangle$ is done by choosing uniform $y, z \in \mathbb{Z}_q$ and outputting the ciphertext

$$\langle g^y, g^z \cdot M \rangle.$$
By the Lemma, we have that the second component $g^z \cdot M$ is uniformly distributed and independent of $M$. The first component $g^y$ is also independent of $M$. Thus, the entire ciphertext of $\widetilde{\Pi}$ contains no information about $M$. It follows that

$$\Pr \left[ \text{PubK}^{\text{eav}}_{\mathcal{A},\widetilde{\Pi}}(n) = 1 \right] = \frac{1}{2}.$$
Security of El Gamal

Now consider the following PPT algorithm $\mathcal{D}$ that attempts to solve the DDH problem relative to $\mathcal{G}$. $\mathcal{D}$ receives $(\mathcal{G}, q, g, h_1, h_2, h_3)$, where $h_1 = g^x$, $h_2 = g^y$ and $h_3$ is either $g^{xy}$ or $g^z$ for uniform $x, y, z$.

**Algorithm $\mathcal{D}$:**
The algorithm is given $(\mathcal{G}, q, g, h_1, h_2, h_3)$ as input.

1. Set $pk = \langle \mathcal{G}, q, g, h_1 \rangle$ and run $A(pk)$ to obtain two messages $M_0, M_1 \in \mathcal{G}$.

2. Choose a uniform bit $b$ and set $c_1 := h_2$ and $c_2 := h_3 \cdot M_b$.

3. Give the ciphertext $\langle c_1, c_2 \rangle$ to $A$ and obtain an output bit $b'$. If $b' = b$, then output 1; otherwise, output 0.
Security of El Gamal

Case 1: Say the input of \( \mathcal{D} \) is generated by running \( \mathcal{G}(1^n) \) to obtain \((\mathbb{G}, q, g)\), then choosing uniform \(x, y, z \in \mathbb{G}\) and finally setting \( h_1 := g^x, h_2 := g^y, \) and \( h_3 := g^z \). Then, \( \mathcal{D} \) runs \( \mathcal{A} \) on public key \( \text{pk} = \langle \mathbb{G}, q, g, g^x \rangle \) and ciphertext

\[
\langle c_1, c_2 \rangle = \langle g^y, g^z \cdot M_b \rangle .
\]

Thus, the view of \( \mathcal{A} \) as a subroutine of \( \mathcal{D} \) is identical to the one in experiment \( \text{PubK}_{\mathcal{A}, \tilde{\Pi}}^{\text{eav}}(n) \). Since \( \mathcal{D} \) outputs 1 exactly when the output \( b' \) of \( \mathcal{A} \) is equal to \( b \), we have that

\[
\Pr \left[ \mathcal{D}(\mathbb{G}, q, g, g^x, g^y, g^z) = 1 \right] = \Pr \left[ \text{PubK}_{\mathcal{A}, \tilde{\Pi}}^{\text{eav}}(n) = 1 \right] = \frac{1}{2} .
\]
Case 2: Say the input of $\mathcal{D}$ is generated by running $\mathcal{G}(1^n)$ to obtain $(\mathbb{G}, q, g)$, then choosing uniform $x, y, z \in \mathbb{G}$ and finally setting $h_1 := g^x$, $h_2 := g^y$, and $h_3 := g^{xy}$. Then, $\mathcal{D}$ runs $\mathcal{A}$ on public key $pk = \langle \mathbb{G}, q, g, g^x \rangle$ and ciphertext

$$\langle c_1, c_2 \rangle = \langle g^y, g^{xy} \cdot M_b \rangle = \langle g^y, (g^x)^y \cdot M_b \rangle.$$ 

Thus, the view of $\mathcal{A}$ as a subroutine of $\mathcal{D}$ is identical to the one in experiment $\text{PubKE}_{\mathcal{A},\Pi}(n)$. Since $\mathcal{D}$ outputs 1 exactly when the output $b'$ of $\mathcal{A}$ is equal to $b$, we have that

$$\Pr [\mathcal{D}(\mathbb{G}, q, g, g^x, g^y, g^{xy}) = 1] = \Pr [\text{PubKE}_{\mathcal{A},\Pi}(n) = 1].$$
Security of El Gamal

By the DDH hardness assumption, we have that

\[ \negl(n) \geq \Pr[\mathcal{D}(G, q, g, g^x, g^y, g^{xy}) = 1] - \Pr[\mathcal{D}(G, q, g, g^x, g^y, g^z) = 1] = \]

\[ = \Pr[\text{PubK}_{A,\Pi}^\text{eav}(n) = 1] - \frac{1}{2}, \]

from where we get that \( \Pr[\text{PubK}_{A,\Pi}^\text{eav}(n) = 1] \leq \frac{1}{2} + \negl(n). \)

\[ \square \]
Security against Chosen-Ciphertext Attacks

Given a public-key encryption scheme $\Pi = (\text{Gen}, \text{Enc}, \text{Dec})$ and an adversary $\mathcal{A}$ consider the following experiment:

**The CCA indistinguishability experiment** $\text{PubK}_{\mathcal{A},\Pi}^{\text{cca}}(n)$:

1. $\text{Gen}(1^n)$ is run to obtain $(pk, sk)$.
2. The adversary $\mathcal{A}$ is given $pk$ and access to a decryption oracle $\text{Dec}_{sk}(\cdot)$. It outputs a pair of equal-length messages $M_0, M_1$ in the message space.
3. A uniform bit $b \in \{0, 1\}$ is chosen and then a ciphertext $c \leftarrow \text{Enc}_{pk}(M_b)$ is computed and given to $\mathcal{A}$. We call $c$ the challenge ciphertext.
4. $\mathcal{A}$ continues to interact with the decryption oracle, but may not request the decryption of $c$ itself. Finally, $\mathcal{A}$ outputs a bit $b'$.
5. The output of the experiment is 1 if $b = b'$ and 0 otherwise. If $b = b'$, we say that $\mathcal{A}$ succeeds.
Security against Chosen-Ciphertext Attacks

Definition
A public-key encryption scheme $\Pi = (\Gen, \Enc, \Dec)$ has indistinguishable encryptions under a chosen ciphertext attack, or it is CCA-secure, if for every PPT adversary $A$ it holds that

$$\Pr \left[ \text{PubK}^{\text{cca}}_{A, \Pi}(n) = 1 \right] \leq \frac{1}{2} + \text{negl}(\cdot).$$
Malleability of El Gamal

An encryption scheme is *malleable* if given a ciphertext $c$ that is an encryption of an unknown message $M$, it is possible to generate a modified ciphertext $c'$ that is an encryption of a message $M'$ having some known relation to $M$. 
In El Gamal, an adversary that intercepts a ciphertext $c = \langle c_1, c_2 \rangle$ can construct a ciphertext $c' = \langle c_1, c'_2 \rangle$, where $c'_2 = \alpha \cdot c_2$.

It is easy to check that if $c$ is an encryption of a message $M$, then $c'$ is a valid encryption of the message $\alpha \cdot M$!

El Gamal is malleable, so it is vulnerable against chosen-ciphertext attacks (Exercise!).
The RSA encryption scheme

Theorem
Let $p, q$ be primes. Let $N := pq$ and \( \phi(N) = (p - 1)(q - 1) \). For integer $e > 0$ define $f_e : \mathbb{Z}_N^* \rightarrow \mathbb{Z}_N^*$ by

\[ f_e(x) = x^e \mod N. \]

If $e$ is relatively prime to $\phi(N)$, then $f_e$ is a permutation. Moreover, if $d = e^{-1} \mod \phi(N)$, then $f_d$ is the inverse of $f_e$. 
The RSA encryption scheme

- **Gen:** On input $1^n$ choose two $n$-bit random primes $p$ and $q$. Compute $N = pq$ and $\phi(N) = (p - 1)(q - 1)$. Choose $e > 1$ such that $\gcd(e, \phi(N)) = 1$. Compute $d := e^{-1} \mod \phi(N)$. Return $(N, e)$ as the public key and $(N, d)$ as the private key.

- **Enc:** on input a public key $pk = (N, e)$ and a message $M \in \mathbb{Z}_N^*$, compute the ciphertext

  $$c = M^e \mod N.$$

- **Dec:** on input a private key $sk = (N, d)$ and a ciphertext $c \in \mathbb{Z}_N^*$, compute the message

  $$M = c^d \mod N.$$

**Figure:** The RSA encryption scheme.
Correctness of RSA

Let $c = M^e \mod N$. Then,

$$\hat{M} := c^d \mod N = (M^e)^d \mod N = M.$$ 

This because $f_d(x) = x^d \mod N$ is the inverse of $f_e(x) = x^e \mod N$. 
Security of RSA

- Factoring is at least as hard (but not known to be equivalent) as breaking RSA.
- RSA is deterministic, therefore it is not CPA-secure.
End

References: Sec 11.1, 11.2.1, 11.4.1, Sec 11.4.1, 11.3.2 (only Definition 11.13), 11.5.1 (up to Construction 11.26)