CS4830: Encryption

Instructor: Elaine Shi, Weikai Lin

March 22, 2017

1 Hybrid Encryption

Definition 1. *(Secure Symmetric-Key Encryption, 91.1). The encryption scheme* $(\mathsf{gen}, \mathsf{enc}, \mathsf{dec})$ *is said to be single-message secure if* $\forall$ *non uniform p.p.t.* D*, there exists a negligible function* $\epsilon(\cdot)$ *such that for all* $n \in \mathbf{N}$*,* $m_0, m_1 \in \{0,1\}^n$*,* D *distinguishes between the the following distributions with probability at most* $\epsilon(n)$*:*

- $\{k \leftarrow \mathsf{gen}(1^n) : \mathsf{enc}_k(m_0)\}_n$
- $\{k \leftarrow \mathsf{gen}(1^n) : \mathsf{enc}_k(m_1)\}_n$

Definition 2. *(Secure Public Key Encryption, 102.2). The public key encryption scheme* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is said to be secure if for all non uniform p.p.t.* D*, there exists a negligible function* $\epsilon(\cdot)$ *such that for all* $n \in \mathbf{N}$*,* $m_0, m_1 \in \{0,1\}^n$*,* D *distinguishes between the the following distributions with probability at most* $\epsilon(n)$*:*

- $\{(pk, sk) \leftarrow \mathsf{Gen}(1^n) : (pk, \mathsf{Enc}_{pk}(m_0))\}_n$
- $\{(pk, sk) \leftarrow \mathsf{Gen}(1^n) : (pk, \mathsf{Enc}_{pk}(m_1))\}_n$

Public-key encryption is typically slower than symmetric-key encryption. Therefore, when we have a long message to encrypt, it is a good idea to use the public key encryption to encrypt a symmetric key, and then use the symmetric key to encrypt the message.

Formally, let $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ denote a (single-message) secure public-key encryption, and let $(\mathsf{gen}, \mathsf{enc}, \mathsf{dec})$ denote a (single-message) secure symmetric-key encryption. Consider the following public-key encryption scheme $(\mathsf{Gen}', \mathsf{Enc}', \mathsf{Dec}')$:

- $\mathsf{Gen}'(1^n)$: call $(pk, sk) \leftarrow \mathsf{Gen}(1^n)$, and output the public key pk and secret key sk.
- $\mathsf{Enc}'(pk, m)$: call $k \leftarrow \mathsf{gen}(1^n)$, and output the following ciphertext: $\mathsf{Enc}_{pk}(k),\ \ \mathsf{enc}_k(m)$
- $\mathsf{Dec}'(sk, \mathsf{ct})$: parse $\mathsf{ct} := (c_0, c_1)$. Call $k := \mathsf{Dec}_{sk}(c_0)$, and then call $m := \mathsf{dec}_k(c_1)$.

Please prove that this is a secure encryption scheme. *Hint: we are sampling* k *randomly, but it is for all* m_0, m_1 *in the definition of secure public key encryption.*

Sol.

Proof. Assume for contradiction, there exists nuPPT D, polynomial p, for infinitely many $n \in \mathbf{N}$, exists $m_0, m_1 \in \{0,1\}^n$ such that D distinguishes between the the following distributions with probability $1/p(n)$:

$$C_0 = \{(pk, sk) \leftarrow \mathsf{Gen}'(1^n) : \mathsf{Enc}'(pk, m_0)\},$$

$$C_1 = \{(pk, sk) \leftarrow \mathsf{Gen}'(1^n) : \mathsf{Enc}'(pk, m_1)\}.$$

To define hybrids, define following encryption algorithm:

$$\mathsf{Enc}''(pk, m)\text{: call } k \leftarrow \mathsf{gen}(1^n), \text{output the following ciphertext: } \mathsf{Enc}_{pk}(0), \mathsf{enc}_k(m).$$

Then, define following hybrid ensembles.

- $H_0 = \{(pk, sk) \leftarrow \mathsf{Gen}'(1^n) : \mathsf{Enc}''(pk, m_0)\}$.
- $H_1 = \{(pk, sk) \leftarrow \mathsf{Gen}'(1^n) : \mathsf{Enc}''(pk, m_1)\}$.

By Hybrid Lemma, D must be able to distinguish between one of three pairs of distributions with probability at least $1/3p(n)$: (C_0, H_0), (H_0, H_1), or (H_1, C_1). We show that all cases are impossible, and then $(\mathsf{Gen}', \mathsf{Enc}', \mathsf{Dec}')$ is a single-message secure public-key encryption.

- C_0, H_0. Rewriting C_0 and H_0 with procedures in $\mathsf{Gen}', \mathsf{Enc}', \mathsf{Enc}''$, and D distinguishes between them with probability $\geq 1/3p(n)$:

$$|\Pr[(pk, sk) \leftarrow \mathsf{Gen}(1^n); k \leftarrow \mathsf{gen}(1^n) : D(1^n, \mathsf{Enc}_{pk}(k), \mathsf{enc}_k(m_0)) = 1] -$$
$$\Pr[(pk, sk) \leftarrow \mathsf{Gen}(1^n); k \leftarrow \mathsf{gen}(1^n) : D(1^n, \mathsf{Enc}_{pk}(0), \mathsf{enc}_k(m_0)) = 1]| \geq 1/3p(n).$$

Rewriting the LHS with summation (and omitting the sampling of pk, k for readability),

$$\begin{aligned} & |\Pr[D(1^n, \mathsf{Enc}_{pk}(k), \mathsf{enc}_k(m_0)) = 1] - \Pr[D(1^n, \mathsf{Enc}_{pk}(0), \mathsf{enc}_k(m_0)) = 1]| \\ =& \Bigg| \sum_a \Pr[D(1^n, \mathsf{Enc}_{pk}(a), \mathsf{enc}_a(m_0)) = 1 | k = a] \Pr[k = a] \\ & - \sum_a \Pr[D(1^n, \mathsf{Enc}_{pk}(0), \mathsf{enc}_a(m_0)) = 1 | k = a] \Pr[k = a] \Bigg| \\ =& \sum_a \Pr[k = a] \left|\Pr[D(1^n, \mathsf{Enc}_{pk}(a), \mathsf{enc}_a(m_0)) = 1 | k = a] - \Pr[D(1^n, \mathsf{Enc}_{pk}(0), \mathsf{enc}_a(m_0)) = 1 | k = a]\right| \\ =& \sum_a \Pr[k = a] |d(a)|, \end{aligned}$$

where $d(a) = \Pr[D(1^n, \mathsf{Enc}_{pk}(a), \mathsf{enc}_a(m_0)) = 1] - \Pr[D(1^n, \mathsf{Enc}_{pk}(0), \mathsf{enc}_a(m_0)) = 1]$. Note that there is no k in $d(a)$. Given $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is a secure public key encryption, there exists a negligible function ϵ such that for all $n \in \mathbf{N}$, for all a, $\Pr[D(1^n, \mathsf{Enc}_{pk}(a)) = 1] - \Pr[D(1^n, \mathsf{Enc}_{pk}(0)) = 1] \leq \epsilon(n)$. By closure under efficient operation, $|d(a)| \leq \epsilon(n)$. Hence, $\sum_a \Pr[k = a] |d(a)| \leq \sum_a \Pr[k = a] \epsilon(n)$ for all $n \in \mathbf{N}$, which contradicts D distinguishes between C_0, H_0 with probability $\geq 1/3p(n)$ for infinitely many n.

- H_0, H_1. Rewriting H_0 and H_1 with procedures in $\mathsf{Gen}', \mathsf{Enc}''$, and D distinguishes between them with probability $\geq 1/3p(n)$:

$$|\Pr[(pk, sk) \leftarrow \mathsf{Gen}(1^n); k \leftarrow \mathsf{gen}(1^n) : D(1^n, \mathsf{Enc}_{pk}(0), \mathsf{enc}_k(m_0)) = 1] -$$
$$\Pr[(pk, sk) \leftarrow \mathsf{Gen}(1^n); k \leftarrow \mathsf{gen}(1^n) : D(1^n, \mathsf{Enc}_{pk}(0), \mathsf{enc}_k(m_1)) = 1]| \geq 1/3p(n).$$

Define nuPPT as $M(x) := (pk, sk) \leftarrow \mathsf{Gen}(1^n)$, output $\mathsf{Enc}_{pk}(0), x$. Rewriting LHS of the above equation,

$$|\Pr[k \leftarrow \mathsf{gen}(1^n) : D(1^n, M(\mathsf{enc}_k(m_0))) = 1] - \Pr[k \leftarrow \mathsf{gen}(1^n) : D(1^n, M(\mathsf{enc}_k(m_1))) = 1]|,$$

we found M is an efficient operation of $\mathsf{enc}_k(m_0)$ or $\mathsf{enc}_k(m_1)$. By $(\mathsf{gen}, \mathsf{enc}, \mathsf{dec})$ is a secure single message encryption, and then by Closure under Efficient Operation, D cannot distinguish H_0, H_1 with probability $1/3p(n)$ for infinitely many $n \in \mathbf{N}$. It is a contradiction as desired.

- H_1, C_1. Following the arguments of C_0, H_0 symmetrically with m_1, we can lead to a contradiction.

□

2 Constructing Secure Symmetric-Key Encryption

Definition 3. *(Pseudo-random Function, 96.2). A family of functions* $\{f_s : \{0,1\}^{|s|} \to \{0,1\}^{|s|}\}_{s\in\{0,1\}^*}$ *is pseudo-random if*

- *(Easy to compute):* $f_s(x)$ *can be computed by a p.p.t. algorithm that is given input* s *and* x
- *(Pseudorandom):* $\{s \leftarrow \{0,1\}^n : f_s\}_n \approx \{F \leftarrow \mathsf{RF}_n : F\}_n$.

Assume $m \in \{0,1\}^n$ and let $\{f_k\}$ be a PRF family. Let U_n be uniform distribution over $\{0,1\}^n$.

- $\mathsf{Gen}(1^n)$: $k' \leftarrow U_n$. Let $k = k'_l||0^{n-l}$.
- $\mathsf{Enc}_k(m)$: Pick $r \leftarrow U_n$. Output $(r, m \oplus f_k(r))$
- $\mathsf{Dec}_k((r,c))$: Output $c \oplus f_k(r)$

Is it a single-message secure encryption if (a) $l = 100$, (b) $l = \log n$, (c) $l = n/2$, (d) $l = n-1$? Is it a multi-message secure encryption if (a) $l = 100$, (b) $l = \log n$, (c) $l = n/2$, (d) $l = n-1$?