6.875J/18.425J Cryptography and Cryptanalysis	October 3, 2001
Lecture 8	
Instructor: Anna Lysyanskaya	*Scribed by: Eamon Walsh*

Today's lecture will be dedicated to showing that GM-security is equivalent to semantic security. The theorem below is originally due to Goldwasser and Micali. It was later refined by Micali, Rackoff and Sloan. This specific version of the definitions given and the outline of the proof below were given to me by Yevgeniy Dodis. This version of the proof of equivalence is a follow-up on the proof given by Dodis and Ruhl in 1998, when Dodis was the TA for 6.875, and Ruhl was taking the class.

1 Two definitions of security

Definition 1 *(One-pass GM-security) A cryptosystem* (G, E, D) *is* one-pass GM-secure *if*

$$\begin{aligned}
&\forall PPTF\{Eve_k\}\, \exists f(k) \in negl, \exists K \in \mathbb{N} \text{ such that } \forall k > K:\\
&\quad \Pr[\ (PK, SK) \leftarrow G(1^k);\\
&\qquad (\alpha, m_0, m_1) \leftarrow Eve_k(PK) \text{ where } |m_0| = |m_1|;\\
&\qquad b \leftarrow \{0,1\};\\
&\qquad c \leftarrow E(PK, m_b);\\
&\qquad b' \leftarrow Eve_k(\alpha, c) \quad : \quad b = b'] \leq 1/2 + f(k)
\end{aligned}$$

Definition 2 *(One-pass semantic security) A cryptosystem* (G, E, D) *is* one-pass semantically secure *if*

$$\begin{aligned}
&\forall PPTF\{Eve_k\}\, \exists PPTF\{S_k\}, \exists f(k) \in negl, \exists K \in \mathbb{N}\\
&\text{such that } \forall k > K,\ p_0(k) - p_1(S_k, k) \leq f(k), \text{ where}\\
&\quad p_0(k) = \Pr[\ (PK, SK) \leftarrow G(1^k);\\
&\qquad\qquad (\alpha, M, R) \leftarrow Eve_k(PK) \text{ where } \forall m_0, m_1 \in M, |m_0| = |m_1|\\
&\qquad\qquad m \leftarrow M;\\
&\qquad\qquad c \leftarrow E(PK, m);\\
&\qquad\qquad z \leftarrow Eve(\alpha, c) \quad : \quad R(m, z)].\\
&\quad p_1(S_k, k) = \Pr[\ (PK, SK) \leftarrow G(1^k);\\
&\qquad\qquad (\alpha, M, R) \leftarrow Eve_k(PK)\\
&\qquad\qquad m \leftarrow M;\\
&\qquad\qquad z \leftarrow S_k(\alpha, M, R, PK) \quad : \quad R(m, z)].
\end{aligned}$$

Remember our intuition for semantic security: whatever can be inferred from the ciphertext can be inferred negligibly less well without it.

We see that one-pass security gives a great deal of control to the adversary Eve_k, both in terms of the message space M and the relation R. By proving one-pass GM-security to be equivalent to one-pass semantic security, we will gain the powerful implications of semantic security while being able to work with the simpler GM-security definition.

2 Equivalence

Theorem 3 *A cryptosystem (G, E, D) is (one-pass) semantically secure if and only if it is (one-pass) GM-secure.*

2.1 GM security implies semantic security

Suppose a cryptosystem (G, E, D) is not semantically secure. We wish to show that it is not GM-secure either.

By assumption, there exists an adversary $\{\text{Eve}_k\}$ such that for all simulators $\{S_k\}$, $p_0(k) - p_1(S_k, k) \geq \epsilon(S_k, k)$ where $\epsilon(S_k, k)$ is non-negligible, i.e., it is inverse-polynomial for infinitely many k.

So, in particular, let S_k be as follows: On input (α, M, R, PK), where α, M, and R were obtained from $Eve_k(PK)$, sample a message $m' \leftarrow M$, let $c \leftarrow E(PK, m')$, and output $z \leftarrow Eve_k(\alpha, c)$. By assumption, $p_0(k) - p_1(S_k, k) \geq \epsilon(S_k, k)$ where $\epsilon(S_k, k)$ is non-negligible, i.e., it is inverse-polynomial for infinitely many k.

Let us now consider the following adversary $\tilde{\text{Eve}}_k$, who makes subroutine calls to $\{\text{Eve}_k\}$, and whose aim is to break GM-security of (G, E, D):

- On input PK,
  - Obtain $(\alpha, M, R) \leftarrow \text{Eve}_k(PK)$.
  - Sample $m_0 \leftarrow M$, $m_1 \leftarrow M$ and output $\tilde{\alpha} = (\alpha, M, R, m_0, m_1), m_0$, and m_1.
- On input $(\tilde{\alpha}, c)$,
  - Obtain $z \leftarrow \text{Eve}_k(\alpha, c)$.
  - If $R(m_b, z) = TRUE \neq R(m_{\bar{b}}, z)$, output b.
  - Otherwise, output 1 or 0 with equal probability.

The following diagram helps in viewing the reduction:

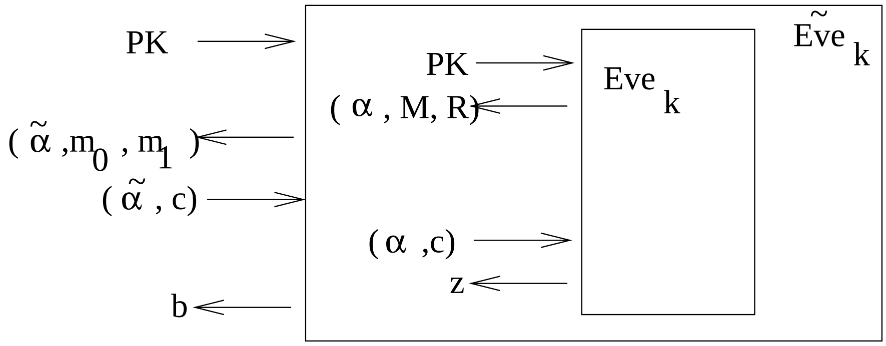

Figure 1: Visual idea behind reduction $\tilde{\text{Eve}}_k$

Let's estimate the success probability of $\tilde{Eve}_k$: Let A be the event that $R(m_b, z) = TRUE$. More formally,

$$\begin{aligned} A = (\ & (PK, SK) \leftarrow G(1^k); \\ & (\tilde{\alpha}, m_0, m_1) \leftarrow \tilde{\text{Eve}}_k(PK); \\ & b \leftarrow \{0,1\}; c \leftarrow E(PK, m_b); \\ & z \leftarrow \text{Eve}_k(\alpha, c) \quad : \quad R(m_b, z)\ ) \end{aligned}$$

Let B be the event that $R(m_{\bar{b}}, z) = TRUE$. More formally:

$$\begin{aligned} B = (\ & (PK, SK) \leftarrow G(1^k); \\ & (\tilde{\alpha}, m_0, m_1) \leftarrow \tilde{\text{Eve}}_k(PK); \\ & b \leftarrow \{0,1\}; c \leftarrow E(PK, m_b); \\ & z \leftarrow \text{Eve}_k(\alpha, c) \quad : \quad R(m_{\bar{b}}, z)\ ) \end{aligned}$$

Note that, for our adversary $\tilde{\text{Eve}}_k$,

$$\Pr[success] = \Pr[A \wedge \bar{B}] + 1/2(\Pr[A \wedge B] + \Pr[\bar{A} \wedge \bar{B}])$$

which is the probability either that R correctly accepts z for m_b and rejects z for $m_{\bar{b}}$, or that the coin flip, by chance, is b. On the other hand,

$$\Pr[fail] = \Pr[\bar{A} \wedge B] + 1/2(\Pr[A \wedge B] + \Pr[\bar{A} \wedge \bar{B}])$$

which is the probability either that R incorrectly accepts z for $m_{\bar{b}}$ and rejects z for m_b, or that the coin flip, by chance, is $\bar{b}$ Therefore,

$$\begin{aligned} \Pr[success] - \Pr[fail] &= \Pr[A \wedge \bar{B}] - \Pr[\bar{A} \wedge B] \\ &= \Pr[A \wedge \bar{B}] + \Pr[A \wedge B] - \Pr[A \wedge B] - \Pr[\bar{A} \wedge B] \\ &= \Pr[A] - \Pr[B]. \end{aligned}$$

So, let us estimate $\Pr[A]$ and $\Pr[B]$.

$$\begin{aligned} \Pr[A] = \Pr[\ & (PK, SK) \leftarrow G(1^k); \\ & (\tilde{\alpha}, m_0, m_1) \leftarrow \tilde{\text{Eve}}_k(PK); \\ & b \leftarrow \{0,1\}; c \leftarrow E(PK, m_b); \\ & z \leftarrow \text{Eve}_k(\alpha, c) \quad : \quad R(m_b, z)] = \end{aligned}$$

$$\begin{aligned} \Pr[\ & (PK, SK) \leftarrow G(1^k); \\ & (\alpha, M, R) \leftarrow \text{Eve}_k(PK); \\ & m_0 \leftarrow M; m_1 \leftarrow M; \\ & b \leftarrow \{0,1\}; c \leftarrow E(PK, m_b); \\ & z \leftarrow \text{Eve}_k(\alpha, c) \quad : \quad R(m_b, z)] = \end{aligned}$$

$$\begin{aligned} \Pr[\ & (PK, SK) \leftarrow G(1^k); \\ & (\alpha, M, R) \leftarrow \text{Eve}_k(PK); \\ & m \leftarrow M; c \leftarrow E(PK, m); \\ & z \leftarrow \text{Eve}_k(\alpha, c) \quad : \quad R(m, z)] = \end{aligned}$$

$p_0(k)$.

$$
\begin{aligned}
\Pr[B] = \Pr[\ & (PK, SK) \leftarrow G(1^k); \\
& (\tilde{\alpha}, m_0, m_1) \leftarrow \tilde{\mathrm{Eve}}_k(PK); \\
& b \leftarrow \{0,1\}; c \leftarrow E(PK, m_b); \\
& z \leftarrow \mathrm{Eve}_k(\alpha, c) \quad : \quad R(m_{\bar{b}}, z)] = \\
\Pr[\ & (PK, SK) \leftarrow G(1^k); \\
& (\alpha, M, R) \leftarrow \mathrm{Eve}_k(PK); \\
& m_0 \leftarrow M; m_1 \leftarrow M; \\
& b \leftarrow \{0,1\}; c \leftarrow E(PK, m_b); \\
& z \leftarrow \mathrm{Eve}_k(\alpha, c) \quad : \quad R(m_{\bar{b}}, z)] = \\
\Pr[\ & (PK, SK) \leftarrow G(1^k); \\
& (\alpha, M, R) \leftarrow \mathrm{Eve}_k(PK); \\
& m \leftarrow M; m' \leftarrow M; c \leftarrow E(PK, m'); \\
& z \leftarrow \mathrm{Eve}_k(\alpha, c) \quad : \quad R(m, z)] = \\
\Pr[\ & (PK, SK) \leftarrow G(1^k); \\
& (\alpha, M, R) \leftarrow \mathrm{Eve}_k(PK) \\
& m \leftarrow M; \\
& z \leftarrow S_k(\alpha, M, R, PK) \quad : \quad R(m, z)] =. \\
& p_1(S_k, k).
\end{aligned}
$$

Note that the last equality uses the earlier definition of S_k.

Therefore, $\Pr[success] - \Pr[fail] = \Pr[A] - \Pr[B] = p_0(k) - p_1(S_k, k) \geq \epsilon(S_k, k)$, where ϵ is a non-negligible function.

This contradicts GM-security of (G, E, D). (Exercise: why?)

2.2 Semantic security implies GM security

Suppose that our cryptosystem (G, E, D) is not GM-secure. We wish to show that it is not semantically secure either.

By our assumption, there exists an adversary $\{\mathrm{Eve}_k\}$ such that there exists a non-negligible function $\epsilon(k)$, such that

$$
\begin{aligned}
\Pr[\ & (PK, SK) \leftarrow G(1^k); \\
& (\alpha, m_0, m_1) \leftarrow \mathrm{Eve}_k(PK) \text{ where } |m_0| = |m_1|, \\
& b \leftarrow \{0,1\}; \\
& c \leftarrow E(PK, m_b); \\
& b' \leftarrow \mathrm{Eve}(\alpha, c) \quad : \quad b = b'] \geq 1/2 + \epsilon(k)
\end{aligned}
$$

Let us construct an adversary $\{\tilde{\mathrm{Eve}}_k\}$ that breaks semantic security of (G, E, D) as follows:

- On input (PK), $\tilde{\mathrm{Eve}}_k$ does:
  - Obtain $(\alpha, m_0, m_1) \leftarrow \mathrm{Eve}_k(PK)$.
  - Let $\tilde{\alpha} = (\alpha, m_0, m_1)$. Let $M = \{m_0, m_1\}$. Let R be the equality relation. Output $(\tilde{\alpha}, M, R)$.
- On input $(\tilde{\alpha}, c)$, where $\tilde{\alpha} = (\alpha, m_0, m_1)$, $\tilde{\mathrm{Eve}}$ does:

– Obtain $b' \leftarrow \mathrm{Eve}_k(\alpha, c)$.

– Output $m_{b'}$.

It is clear that the the value $p_0(k)$ for this adversary is:

$$\begin{aligned} p_0(k) = \Pr[\ & (PK, SK) \leftarrow G(1^k); \\ & (\alpha, m_0, m_1) \leftarrow \mathrm{Eve}_k(PK), \\ & b \leftarrow \{0,1\}; \\ & c \leftarrow E(PK, m_b); \\ & b' \leftarrow \mathrm{Eve}(\alpha, c) \quad : \quad b = b'] \geq 1/2 + \epsilon(k) \end{aligned}$$

On the other hand, what is the best chances for an algorithm that does not take c as input? It cannot be more than $1/2$, since we may think of it as though first the simulator outputs a guess, and then bit b is picked. Therefore, for all simulators S_k, $p_0(k) - p_1(S_k, k) \geq \epsilon(k)$, which contradicts the semantic security of (G, E, D).

3 Other definitions

Twenty years ago, there were no established definitions of security. But people had intuition that "seemed right." The intuition behind the RSA cryptosystem is that, without any a-priori information on the message, it is hard to infer it from seeing its encrypted form. As we have stressed before, this intuition is much weaker than the intuition behind GM-security. Indeed, if we know a candidate message sent using RSA encryption, then we can check whether or not this was the message sent.

Yao had another intuition, which had to do with effectiveness of data compression. His intuition is that, if a cryptosystem is secure, then it takes just as many bits to transmit a message, whether I happen to have this message in encrypted form already or not. From the point of view of an information-theorist, this intuition is very natural, and quite akin to Shannon's ideas about security. It turns out that Yao's definition is equivalent to semantic security. This was shown by Micali, Rackoff, and Sloan.

It is possible that other notions of security may prove useful, even if they are not comparable to existing ones. The study of useful notions of security is on-going, and if you have any interesting ideas related to it, please be sure to follow up on them!

4 Why worry about definitions so much?

We focus on *provable* security, i.e., if we claim that a cryptosystem is secure, then we also exhibit a proof of this claim, under some reasonable computational assumption (where "reasonable" is in the eye of the beholder, of course). Thus the need for a definition of security is natural. Whether a given formal definition satisfies one's need for security, or comes short of it, is, needless to say, in the eye of the beholder as well. A discussion of possible definitions and their strength is therefore central.

Why do we focus so much on these specific definitions? The concepts of semantic security and indistinguishability are central in cryptography. In the end, it's all about the bad guys not being any better off after seeing something than they were without seeing it.

Proving that this very powerful definition is equivalent to another, much simpler one, besides making our job as cryptographers much simpler, also tells us that there is something fundamental about the notion of indistinguishability of two probability distributions, and inspires us to study this notion further. Looking at it from a more general perspective will be the subject of the next few lectures.