CMSC 858K — Advanced Topics in Cryptography February 19, 2004

Lecture 8

Lecturer: Jonathan Katz *Scribe(s):* *Alvaro A. Cardenas*
Nicholas Sze
Yinian Mao
Kavitha Swaminathan

1 Introduction

Last time we introduced the Naor-Yung construction of a CCA1-secure cryptosystem, and gave a proof of security. We also gave the construction of a CCA2-secure cryptosystem by Dolev-Dwork-Naor. Here, we complete the proof that this cryptosystem is indeed CCA2 secure.

2 The Dolev-Dwork-Naor (DDN) Scheme [1]

Given a public-key encryption scheme $(\mathsf{Gen}', \mathcal{E}', \mathcal{D}')$, an adaptively-secure NIZK proof system $(\mathcal{P}, \mathcal{V})$, and a (strong, one-time) signature scheme $(\mathsf{SigGen}, \mathsf{Sign}, \mathsf{Vrfy})$, the DDN encryption scheme is constructed as follows (in the following, $poly(k)$ represents some unspecified polynomial which is not necessarily always the same):

- $\mathsf{Gen}(1^k)$:
  for $i = 1$ to k do $(pk_{i,0}, sk_{i,0}) \leftarrow \mathsf{Gen}'(1^k)$, $(pk_{i,1}, sk_{i,1}) \leftarrow \mathsf{Gen}'(1^k)$
  Select a random r: $r \leftarrow \{0,1\}^{poly(k)}$
  Output $pk^* = \begin{bmatrix} pk_{1,0} & pk_{2,0} & \cdots & pk_{k,0} \\ pk_{1,1} & pk_{2,1} & \cdots & pk_{k,1} \end{bmatrix}, r$
  and $sk^* = \begin{bmatrix} sk_{1,0} & sk_{2,0} & \cdots & sk_{k,0} \\ sk_{1,1} & sk_{2,1} & \cdots & sk_{k,1} \end{bmatrix}$
  (in fact, we may simplify things and let $sk^* = (sk_{1,0}, sk_{1,1})$; see below).

- $\mathcal{E}_{pk^*}(m)$:
  $(vk, sk) \leftarrow \mathsf{SigGen}(1^k)$
  view vk as a sequence of k bits[1]; i.e., $vk = vk_1|vk_2|\cdots|vk_k$
  for $i = 1$ to k: $w_i \leftarrow \{0,1\}^{poly(k)}$; $c_i \leftarrow \mathcal{E}'_{pk_{i,vk_i}}(m; w_i)$
  $\pi \leftarrow \mathcal{P}(r, \vec{c}, \vec{w})$
  $\sigma \leftarrow \mathsf{Sign}_{sk}(\vec{c}|\pi)$
  Output $(vk, \vec{c}, \pi, \sigma)$

- $\mathcal{D}_{sk^*}(vk, \vec{c}, \pi, \sigma)$:
  If $\mathsf{Vrfy}_{vk}(\vec{c}|\pi, \sigma) = 0$ then output $\perp$
  If $\mathcal{V}(r, \vec{c}, \pi) = 0$ then output $\perp$
  Else output $\mathcal{D}'_{sk_{1,vk_1}}(c_1)$

[1] The scheme can be modified in the obvious way for vk of arbitrary (polynomial) length.

Theorem 1 *The encryption scheme presented above is CCA2 secure if* $(\mathsf{Gen}', \mathcal{E}', \mathcal{D}')$ *is semantically secure,* $(\mathcal{P}, \mathcal{V})$ *is an adaptively-secure NIZK proof system, and* $(\mathsf{SigGen}, \mathsf{Sign}, \mathsf{Vrfy})$ *is a strong, one-time signature scheme.*

Proof Consider an arbitrary PPT adversary A with adaptive access to a decryption oracle. We will use a sequence of games in which the first game will correspond to a real encryption of m_0, the final game will correspond to a real encryption of m_1 (here, m_0, m_1 are the messages output by A), and in each stage along the way we show that the difference in the adversary's probability of outputting "1" is negligible. This then implies that the difference between the probability that it outputs 1 when it gets an encryption of m_0 and the probability it outputs 1 when it gets an encryption of m_1 is also negligible, and that is exactly the definition of CCA2 security.

Game 0 is the encryption of m_0 using the real cryptosystem:

$$\begin{array}{llrcl}
\textbf{Game } 0\textbf{:} & \underline{\text{Stage 1}} & \{(pk_{i,b}, sk_{i,b})\} & \leftarrow & \mathsf{Gen}'(1^k),\text{ for } i = 1, 2, \ldots, k \text{ and } b = 0, 1 \\
 & & r & \leftarrow & \{0,1\}^{\mathsf{poly}(k)} \\
 & & (pk^*, sk^*) & = & ((\{pk_{i,b}\}, r), \{sk_{i,b}\}) \\
 & & (m_0, m_1) & \leftarrow & A^{\mathcal{D}_{sk^*}(\cdot)}(pk^*) \\
 & & & & \\
 & \underline{\text{Stage 2}} & (vk, sk) & \leftarrow & \mathsf{SigGen}(1^k) \\
 & & w_i & \leftarrow & \{0,1\}^{poly(k)},\text{ for } i = 1, 2, \cdots, k \text{ (from now on} \\
 & & & & \text{we let this step be implicit)} \\
 & & c_i & \leftarrow & {\mathcal{E}'}_{pk_{i,vk_i}}(m_0; w_i),\text{ for } i = 1, 2, \cdots, k \\
 & & \pi & \leftarrow & \mathcal{P}(r, \vec{c}, \vec{w}) \\
 & & \sigma & \leftarrow & \mathsf{Sign}_{sk}(\vec{c}|\pi) \\
 & & b^* & \leftarrow & A^{\mathcal{D}_{sk^*}(\cdot)}(pk^*, vk, \vec{c}, \pi, \sigma)
\end{array}$$

Then, we modify Game 0 by simulating r and π to obtain Game 1. Simulator Sim_1 generates r and simulator Sim_2 outputs π without any witness.

$$\begin{array}{llrcl}
\textbf{Game } 1\textbf{:} & \underline{\text{Stage 1}} & \{(pk_{i,b}, sk_{i,b})\} & \leftarrow & \mathsf{Gen}'(1^k),\text{ for } i = 1, 2, \cdots, k \text{ and } b = 0, 1 \\
 & & r & \leftarrow & \boxed{\mathsf{Sim}_1(1^k)} \\
 & & (pk^*, sk^*) & = & ((\{pk_{i,b}\}, r), \{sk_{i,b}\}) \\
 & & (m_0, m_1) & \leftarrow & A^{\mathcal{D}_{sk^*}(\cdot)}(pk^*) \\
 & & & & \\
 & \underline{\text{Stage 2}} & (vk, sk) & \leftarrow & \mathsf{SigGen}(1^k) \\
 & & c_i & \leftarrow & {\mathcal{E}'}_{pk_{i,vk_i}}(m_0; w_i),\text{ for } i = 1, 2, \cdots, k \\
 & & \pi & \leftarrow & \boxed{\mathsf{Sim}_2(\vec{c})} \\
 & & \sigma & \leftarrow & \mathsf{Sign}_{sk}(\vec{c}|\pi) \\
 & & b^* & \leftarrow & A^{\mathcal{D}_{sk^*}(\cdot)}(pk^*, vk, \vec{c}, \pi, \sigma)
\end{array}$$

Claim 2 *Let* $\Pr_i[\cdot]$ *denote the probability of a given event in game* i. *Then for any* PPT A *the following is negligible:* $|\Pr_0[b^* = 1] - \Pr_1[b^* = 1].|$

Sketch of Proof (Informal) The validity of this claim is intuitively clear as if the probabilities are substantially different then A can be used as a distinguisher between a real NIZK proof and a simulated proof. distinguish a simulated proof from a real proof. We provide more details now.

Given a PPT adversary A, construct the following PPT adversary A' (adversary A' will attempt to distinguish between real/simulated proofs):

```
$A'(r)$:  // $r$ is either a truly random string or a string output by $\mathsf{Sim}_1$
          $\{(pk_{i,b}, sk_{i,b})\} \leftarrow \mathsf{Gen}'(1^k)$, for $i = 1, 2, \cdots, k$ and $b = 0, 1$
          $pk^* = (\{pk_{i,b}\}, r)$
          $(m_0, m_1) \leftarrow A(pk^*)$
          $(vk, sk) \leftarrow \mathsf{SigGen}(1^k)$
          $\forall i\ c_i \leftarrow \mathcal{E}'_{pk_{i,,vk_i}}(m_0; w_i)$
          Output $(\vec{c}, \vec{w})$
          get $\pi$     //$\pi$ is either a real proof or a simulated proof
          $\sigma \leftarrow \mathsf{Sign}_{sk}(\vec{c}|\pi)$
          $b^* \leftarrow A^{\mathcal{D}_{sk^*}(\cdot)}(pk^*, vk, \vec{c}, \pi, \sigma)$
          Output $b^*$
```

Note that A' has no problems simulating the decryption oracle for A, since it has all necessary secret keys. If (r, π) are a real string/proof, then A is interacting in Game 0 and so the probability that A' outputs 1 is the probability that A outputs 1 in Game 0. On the other hand, if (r, π) are a simulated string/proof, then A is interacting in Game 1 and so the probability that A' outputs 1 is the probability that A outputs 1 in Game 1. Since the NIZK proof system is adaptively-secure, we must have $|\Pr_0[b^* = 1] - \Pr_1[b^* = 1].|$. □

We construct Game $1'$ as Game 1 except that if A ever makes valid decryption oracle query using vk (where vk is the verification key used to construct the challenge ciphertext), then we simply return $\perp$ in response to this query. We claim that $|\Pr_{1'}[b^* = 1] - \Pr_1[b^* = 1]|$ is negligible. Note that the only difference between the games occurs if A is able to forge a new, valid signature with respect to vk (since ciphertexts submitted to the decryption oracle must be *different* from the challenge ciphertext, and since ciphertexts are only valid if the signature verifies correctly); furthermore, the security of the signature scheme ensures that this event occurs with only negligible probability. Details omitted.

We construct a new game, Game $1''$, which is the same as Game $1'$ except that instead of using sk_{1,vk'_1} to decrypt a ciphertext $(vk', \vec{c'}, \pi', \sigma')$ (i.e., to answer decryption oracle queries for this ciphertext), we look for the first bit position i where vk and vk' differ[2] (i.e., $vk_i \neq vk'_i$) and use key sk_{i,vk'_i} to decrypt. I.e., the decryption oracle now works as follows:

$$\mathcal{D}''_{sk^*}(vk', \vec{c'}, \pi', \sigma') = \begin{cases} \perp & \text{if } vk' = vk; \\ \perp & \text{if } \mathsf{Vrfy}_{vk'}(\vec{c'}|\pi', \sigma') = 0 \text{ or } \mathcal{V}(r, \vec{c'}, \pi') = 0; \\ \mathcal{D}'_{sk_{i,vk'_i}}(c'_i) & \text{otherwise (where } i \text{ is as discussed above)} \end{cases}.$$

Claim 3 *For any* PPT *A the following is negligible:* $|\Pr_{1''}[b^* = 1] - \Pr_{1'}[b^* = 1]|$.

[2] Here, vk is again the verification key used for the challenge ciphertext; note that there must be a bit position where they differ since if $vk = vk'$ we abort anyway.

Sketch of Proof (Informal) In a given query to the decryption oracle, if all ciphertexts decrypt to the same thing then it doesn't really matter what secret key we use. The only difference between Game $1''$ and Game $1'$ occurs if the adversary queries a vector of ciphertexts $\vec{c'}$ where different ciphertexts decrypt to different messages. So the only possible way to distinguish between Game 1 and Game $1'$ is if a decryption query is ever made for which there exists two different indices i and j where the decryption of c'_i is not equal to the decryption of c'_j and yet the proof is valid (i.e., $V(r, \vec{c'}, \pi') = 1$). We argue that this event occurs with negligible probability.

Let Fake be the event that A requests a decryption query $(vk', \vec{c'}, \pi', \sigma')$ s.t. π' is a valid proof and $\exists i, j$ s.t. $\mathcal{D}'_{sk_{i,vk'_i}}(c_i) \neq \mathcal{D}'_{sk_{j,vk'_j}}(c_j)$. Note that $\Pr_{1''}[\mathsf{Fake}]] = \Pr_{1'}[\mathsf{Fake}]$ (since there is no difference between the games until Fake occurs). Furthermore, we claim that $|\Pr_{1'}[\mathsf{Fake}] - \Pr_1[\mathsf{Fake}]|$ is negligible. This is so because (as before) the only difference between these games occurs if the adversary forges a signature using vk, which happens with only negligible probability. We also claim that $|\Pr_1[\mathsf{Fake}] - \Pr_0[\mathsf{Fake}]|$ is negligible, since otherwise we can construct an adversary A' distinguishing real from simulated proofs, similar to the proof of Claim 1 (it is essential here that A' knows all secret keys, so can check when event Fake occurs). Finally, note that $\Pr_0[\mathsf{Fake}]$ is negligible by the (adaptive) soundness of the NIZK proof system. We conclude that $\Pr_{1''}[\mathsf{Fake}]$ is negligible, and this is sufficient to complete the proof of the claim. □

We construct Game 2 which is the same as Game $1''$ except that we form the challenge ciphertext by encrypting (k copies of) m_1 instead of m_0. I.e., for all i: we compute $c_i \leftarrow \mathcal{E}'_{pk_{i,vk_i}}(m_1)$

Claim 4 *For any* PPT *A the following is negligible:* $|\Pr_2[b^* = 1] - \Pr_{1''}[b^* = 1]|$.

Sketch of Proof (Informal) If A can distinguish between these two games we construct an adversary A' attacking the semantic security of the underlying encryption scheme. Actually, instead of attacking a *single* instance of the encryption scheme it will attack k instances of the encryption scheme; i.e., it gets k independently-generated public keys, outputs m_0, m_1, gets back either an encryption of m_0 (with respect to all k keys) or an encryption of m_1, and then has to guess which is the case. Note, however, that by a standard hybrid argument the semantic security of a single instance implies the semantic security of poly-many instances.

We construct our A' as follows:

```
$\underline{A'(pk_1, \cdots, pk_k):}$
$(vk, sk) \leftarrow \mathsf{SigGen}(1^k)$
$\{(pk'_i, sk'_i)\} \leftarrow \mathsf{Gen}'(1^k)$, for $i = 1, 2, \cdots, k$
$r \leftarrow \mathsf{Sim}_1(1^k)$
$pk^* = (\{pk_{i,b}\}, r)$, where $pk_{i,b} = \begin{cases} pk_i & \text{if } b = vk_i \\ pk'_i & \text{otherwise} \end{cases}$
$(m_0, m_1) \leftarrow A^{\mathcal{D}^*(\cdot)}(pk^*)$
Output $(m_0, m_1)$, get back $\vec{c}$
$\pi \leftarrow \mathsf{Sim}_2(\vec{c})$
$\sigma \leftarrow \mathsf{Sign}_{sk}(\vec{c}|\pi)$
Output whatever$A^{\mathcal{D}^*(\cdot)}(vk, \vec{c}, \pi, \sigma)$ outputs
```

It is crucial to note here that A' *can simulate the decryption oracle* $\mathcal{D}^*$ — in particular, for any ciphertext $(vk', \vec{c'}, \pi', \sigma')$ submitted by A, if $vk' = vk$ then A' just returns $\perp$ (as in the previous game), whereas if $vk' \neq vk$ then there is a bit i where they differ (i.e., $vk'_i \neq vk_i$) and A' can use the secret key $sk_{i,vk'_i} = sk'_i$ (which is knows!) to decrypt. This is by construction: A' knows exactly half the secret keys (i.e., those in positions not overlapping with vk) and can use those to decrypt.

Notice that if $\vec{c}$ is an encryption of m_1 then A is essentially interacting in Game 2, whereas if it is an encryption of m_0 then A is in Game $1''$. So, if A can distinguish between Game 1" and Game 2 then A' can distinguish the encryptions and break the semantic security of the underlying encryption scheme. □

Let Game 3 correspond to an encryption of m_1 in the real encryption scheme. We jump ahead and claim the following:

Claim 5 *For any* PPT *A, the following is negligible:* $|\Pr_3[b^* = 1] - \Pr_2[b^* = 1]|$.

Sketch of Proof (Informal) Technically, the proof would proceed by a sequence of games exactly analogous to games 1, $1'$, and $1''$ that we introduced previously. In particular, we would first revert back to decrypting using either $sk_{1,0}$ or $sk_{1,1}$; would then revert back to decrypting even if $vk' = vk$; and, finally, would go back to using a real random string/proof rather than simulated ones. Because these games (and the proofs that they all do not affect the probability that $b^* = 1$ by more than a negligible amount) are essentially the same as before, we do not repeat the arguments here. □

The above sequence of claims implies (by multiple applications of the triangle inequality) that $|\Pr_0[b^* = 1] - \Pr_3[b^* = 1]|$ is negligible; this is exactly equivalent to saying that the scheme is secure against adaptive chosen-ciphertext attacks. ■

3 Summary

We give a definition of a one-way function.

Definition 1 A function $f : \{0,1\}^* \to \{0,1\}^*$ is a *one-way function* if the following hold:

1. $f(x)$ is computable in time polynomial in $|x|$.
2. For all PPT algorithms A, the following is negligible (in k):

$$\mathsf{Pr}[x \leftarrow \{0,1\}^k ; y = f(x); x' \leftarrow A(y) : f(x') = y].$$

♢ It is not hard to show that if a one-way function exists, then $P \neq NP$. The converse (i.e., whether $P \neq NP$ implies the existence one-way functions), is not known to hold.

Since the existence of semantically-secure public-key encryption schemes implies the existence of one-way functions[3], which furthermore implies the existence of one-time strong signature schemes, we may restate the result of the previous section as follows:

[3] Prove it as an exercise!

Theorem 6 *If there exists a semantically-secure public-key encryption scheme and an adaptively-secure NIZK proof system, then there exists a CCA2-secure encryption scheme.*

Later in the course, we will show:

Theorem 7 *If there exist trapdoor permutations, then there exists an adaptively-secure NIZK proof system.*

We have shown in a previous lecture that the existence of trapdoor permutations implies the existence of semantically-secure public-key encryption. The gives the following corollary:

Corollary 8 *If there exist trapdoor permutations, then there exists a CCA2-secure encryption scheme.*

The following important question is still open:

> *Does semantically-secure public-key encryption imply CCA2-secure public-key encryption?*

In particular, can we construct adaptively-secure NIZK proof systems based on semantically-secure public-key encryption? Note that these questions are especially interesting since we do have examples of public-key encryption schemes which are not based (or, at least, so not seem to be based) on trapdoor permutations; El Gamal encryption is probably the best-known example.

References

[1] D. Dolev, C. Dwork, and M. Naor. Non-Malleable Cryptography. In *23rd ACM Symposium on the Theory of Computing*, pages 543-552, 1991.

[2] M. Naor and M. Yung. Public-Key Cryptosystems Provably Secure Against Chosen Ciphertext Attacks. In *22nd ACM Symposium on the Theory of Computing*, pages 427-437, 1990