CS276: Cryptography September 22, 2015

Commitment Schemes and Introduction to Encryption

Instructor: Alessandro Chiesa *Scribe: Pratyush Mishra*

1 Commitment Schemes Continued

1.1 Commitment Scheme from a One-Way Permutation

Here we complete the proof that one can obtain a commitment scheme from a one-way permutation.

Theorem 1 *The existence of one-way permutations implies the existence of a single bit commitment scheme (i.e. $l(k) = 1$).*

Proof: Let $f_k \colon \{0,1\}^{n(k)} \to \{0,1\}^{n(k)}$ be a one-way permutation with hardcore bit $b_k \colon \{0,1\}^{n(k)} \to \{0,1\}$.

Let the commitment scheme be:

$$\mathsf{C}(1^k, s, m) := f_k(s), b_k(s) \oplus m$$

$$\mathsf{R}(1^c k, s, c) := \begin{cases} b_k(s) \oplus c_2, & f_k(s) = c_1 \\ \perp, & f_k(s) \neq c_1 \end{cases}$$

- **Binding:**
  $\forall c = (c_1, c_2)$ $\exists s, m$ that yield (c_1, c_2): $s = f_k^{-1}(c_1), m = b_k(s) \oplus c_2$.
- **Hiding:** Suppose $\exists$ distinguisher D such that the advantage

  $$\delta(k) = \left| \Pr\left[D(f_k(U_{n(k)}), b_k(U_{n(k)})) = 1\right] - \Pr\left[D(f_k(U_{n(k)}), \overline{b_k(U_{n(k)})}) = 1\right] \right|$$

  is non negligible. Then we can define a distinguisher D' for b_k:

Algorithm 1: Machine D' that breaks b_k

Machine $D'(y)$
  $\sigma \xleftarrow{\$} \{0,1\}$
  $b \leftarrow D(y, \sigma)$
  if $b = 1$ **then**
    Output σ
  else
    Output $\overline{\sigma}$

Now,

$$\begin{aligned}\Pr\big[A(f(U_{n(k)})) = b(U_{n(k)})\big] &= \frac{1}{2}\Pr\big[A(f(U_{n(k)})) = b(U_{n(k)}) \mid \sigma = b_k(U_{n(k)})\big] + \\ &\quad \frac{1}{2}\Pr\Big[A(f(U_{n(k)})) = b(U_{n(k)}) \mid \sigma = \overline{b_k(U_{n(k)})}\Big] \\ &= \frac{1}{2}\Pr\big[D(f(U_{n(k)}), b(U_{n(k)})) = 1\big] + \\ &\quad \frac{1}{2}\Pr\Big[D(f(U_{n(k)}), \overline{b(U_{n(k)})}) = 0\Big] \\ &= \frac{1}{2} + \frac{\delta(k)}{2}\end{aligned}$$

Hence we can break the "hardcoreness" of b.

□

1.2 Expanding the Message Space of Commitment Schemes

Theorem 2 *Let* (C, R) *be a commitment scheme with seed length* $r(k)$ *and message length* $l(k)$.

Then, $\forall$ *polynomials* p, $\exists(\vec{(}c), \vec{(}r))$ *that is a commitment scheme withseed length* $p(k) \cdot r(k)$ *and message length* $p(k) \cdot l(k)$

Proof:

$$\begin{aligned}\vec{\mathsf{C}}(1^k, \vec{s}, \vec{m}) &= \big(\mathsf{C}(1^k, s, m_1), \ldots, \mathsf{C}(1^k, s_{p(k)}, m_{p(k)})\big) \\ \vec{\mathsf{R}}(1^k, \vec{s}, \vec{c}) &= \bigwedge_i \big(\mathsf{R}(1^k, s, c_i)\big)\end{aligned}$$

- **Binding:**
  This property for $\vec{\mathsf{C}}$ extends simply from the binding property of C.
- **Hiding:**
  Suppose that $\vec{\mathsf{C}}$ does not hide the commitments. So then we can suppose that $\exists \left\{m_k^{(0)}\right\}_k, \left\{m_k^{(1)}\right\}_k$, ppt D such that

  $$\delta(k) = \left|\Pr\Big[D\Big(\vec{\mathsf{C}}\Big(U_{p\cdot r}, \vec{m}_k^{(0)}\Big)\Big) = 1\Big] - \Pr\Big[D\Big(\vec{\mathsf{C}}\Big(U_{p\cdot r}, \vec{m}_k^{(1)}\Big)\Big) = 1\Big]\right|$$

  is non-negligible in k, i.e. we have a distinguisher for $\vec{\mathsf{C}}$ that has non-negligible advantage. Now, we define hybrids

  $$H_k^{(i)} = \vec{\mathsf{C}}\Big(U_{p\cdot r}, m_{k,1}^{(0)}, \ldots, m_{k,i}^{(0)}, m_{k,i+1}^{(1)}, \ldots, m_{k,p}^{(1)}\Big)$$

  Note that $H_k^{(p)} = \vec{\mathsf{C}}\Big(U_{p\cdot r}, \vec{m}_k^{(0)}\Big)$ and $H_k^0 = \vec{\mathsf{C}}\Big(U_{p\cdot r}, \vec{m}_k^{(1)}\Big)$. Then by the hybrid argument there exists a i such that the advantage of D on $H_k^{(i)}$ vs. $H_k^{(i+1)}$ is non-negligible. With this knowledge we can define a distinguisher D' that attacks (C, R) (Algorithm 2).
  If $\vec{c} \sim H_k^{(i)}$ then $c \sim m_k^{(0)}$. Otherwise, $c \sim m_k^{(1)}$.

Algorithm 2: Machine D' that breaks (C, R)

Machine $D'(y)$

for $j = 1, \ldots, i$ **do** $c_j \xleftarrow{\$} \mathsf{C}(\mathcal{U}_r^{(p)}, m_{k,j}^{(0)})$

for $j = i+2, \ldots, p$ **do** $c_j \xleftarrow{\$} \mathsf{C}(\mathcal{U}_r^{(p)}, m_{k,j}^{(1)})$

$\vec{c} = (c_1, \ldots, c_i, c, c_{i+2}, \ldots, c_p)$

Output $D(\vec{c})$

□

2 Encryption Schemes

Informally, our setting has two parties, Alice and Bob that wish to communicate securely. They share some randomness σ that is known only to them. This is the private key that they use to encrypt their communications. Eve is a passive adversary that observes all messages that pass between Alice and Bob, but cannot tamper with these messages in any way. We would like our encryption scheme to have the following properties:

- **Functionality**: If Alice sends message m, Bob receives exactly m.
- **Security**: Eve learns nothing about m.

Definition 3 *An* encryption scheme *is a tuple of PPT algorithms* $(\mathsf{Enc}, \mathsf{Dec})$ *that satisfies the following properties:*

1. *Completeness:*
   $\forall\ k \in \mathbb{N}, \forall\ \mathsf{sk} \in \{0,1\}^{l(k)}, \forall\ m \in \{0,1\}^{n(k)}$*, we have that*
   $$\Pr[\mathsf{Dec}(1^k, \mathsf{sk}, \mathsf{Enc}(1^k, \mathsf{sk}, m)) = m] = 1$$
2. *Security:*
   $\forall\ \left\{m_k^{(0)}\right\}, \left\{m_k^{(1)}\right\}$ *such that* $m_k^{(b)} \in \{0,1\}^{n(k)}$*, we have*
   $$\left\{\mathsf{Enc}(1^k, U_{l(k)}, m_k^{(0)})\right\} \overset{\text{P/s/c}}{=} \left\{\mathsf{Enc}(1^k, U_{l(k)}, m_k^{(1)})\right\}$$

Theorem 4 $\exists\ (\mathsf{Enc}, \mathsf{Dec})$ *that satisfies completeness and perfect message indistinguishability.*

Proof:

Consider the **One Time Pad** (OTP):

$\mathsf{Enc}(1^k, \mathsf{sk}, m) := sk \oplus m$
$\mathsf{Dec}(1^k, \mathsf{sk}, c) := sk \oplus c$

Completeness is easy to see. Security follows from the fact that

$$U_{l(k)} \oplus m = U_{l(k)} = U_{l(k)} \oplus m'$$

□

However this construction has some important limitations that make it impractical:

1. Keys are large: $|\mathsf{sk}| \geq |m|$.
2. Key can only be used once. Say that we reuse the same secret key sk to encrypt two different messages m_1 and m_2: $c_1 = \mathsf{sk} \oplus m_1$, $c_2 = \mathsf{sk} \oplus m_2$. Then the adversary can XOR the two ciphertexts to obtain some information about each message: $c_1 \oplus c_2 = m_1 \oplus m_2$.

In fact, the "largeness" of keys is inherent to all encryption schemes that aim for perfect message indistinguishability.

Claim 5 *For every encryption scheme* $(\mathsf{Enc}, \mathsf{Dec})$ *that satisfies completeness and perfect message indistinguishability, it holds that* $l(k) \geq n(k)$.

Proof: Suppose $l(k) < n(k)$. Pick $m^{(0)} \in \{0,1\}^{n(k)}, \mathsf{sk} \in \{0,1\}^{l(k)}$ and let $c = \mathsf{Enc}(1^k, \mathsf{sk}, m^{(0)})$.

Now, pick $m^{(1)} \in \{0,1\}^{n(k)}$ such that $\forall \mathsf{sk}' \in \{0,1\}^{l(k)}$ we have $\mathsf{Dec}(1^k, \mathsf{sk}', c) \neq m^{(1)}$. We know that such an $m^{(1)}$ exists because $\left|\cup_{\widetilde{\mathsf{sk}}} \mathsf{Dec}(1^k, \widetilde{\mathsf{sk}}, c)\right| \leq 2^{l(k)} < 2^{n(k)}$. By completeness of the scheme, we have $\mathsf{Enc}(1^k, \mathsf{sk}', m^{(1)}) \neq c$, thus breaking the perfect message indistinguishability property. □

Thus, we see that perfect message indistinguishability, while offering an ideal level of security, has some serious drawbacks. Thus, we settle for **computational message indistinguishability**. However, this notion does not intuitively model the real world. Thus we define **semantic security**, and in the next lecture we will show that these two notions of security are in fact equivalent

2.1 Semantic Security

Consider the same setup as before, with Alice sending a message m drawn from some message distribution $\mathcal{M}$ to Bob. Eve is a passive adversary that wants to learn some function f of m while possessing some partial information $I(m)$ about the message. Intuitively, this definition of security says that Eve should not be able to do much better than if she didn't have the ciphertext at all. Rigorously, we have

Definition 6 ***Semantic Security*** *of an Encryption Scheme* $(\mathsf{Enc}, \mathsf{Dec})$*:*

$\forall$ *message distributions* $\mathcal{M}_k \in \Delta(\{0,1\}^{n(k)})$, $\forall$ *goal functions* $f_k : \{0,1\}^{n(k)} \to \{0,1\}^*$,
$\forall$ *partial information* $I_k : \{0,1\}^{n(k)} \to \{0,1\}^*$, $\forall$ *PPT* A, $\exists$ *ppt* S_A *such that*

$$\left|\Pr[A(1^k, I_k(\mathcal{M}_k), \mathsf{Enc}(1^k, \mathcal{U}_{l(k)}, \mathcal{M}_k)) = f_k(\mathcal{M}_k)] - \Pr[S_A(1^k, I_k(\mathcal{M}_k)) = f_k(\mathcal{M}_k)]\right| \leq \mathit{negl}(k)$$