CSE 5852 – Modern Cryptography: Foundations - Fall 2016

Paramik Dasgupta
Department of Computer Science and Engineering
University of Connecticut,Storrs

Lecture Note 8

September 26, 2016

1 Last Class

Last class we introduce the concept of computational security. In cryptography we have two types of security:

1. Concrete security: measures the security of protocols against current attacks and tries to predict how long an adversary will take to break the system. These numbers are very hard to obtain for a new protocol and should be judged conservatively.

2. Asymptotic security: considers a sequence of protocols and asks that the adversary gets work at breaking the protocols as the sequence proceeds (even with additional resources). The standard here is that when the adversary is given time that is a polynomial function of the sequence position their success in breaking the protocol should shrink faster than any inverse polynomial function.

2 Indistinguishable Encryption

Definition 1 (Indistinguishable)**.** *An encryption scheme* $(\mathcal{M}, K, \mathsf{Enc}, \mathsf{Dec})$ *has indistinguishable encryptions if for all* $\mathcal{A}$ *for every two messages* $m_1, m_2 \in \mathcal{M}$ *and for every ciphertext c:*

$$|\Pr_{k \in K}[\mathcal{A}(\mathsf{Enc}_k(m_1)) = 1] - \Pr_{k \in K}[\mathcal{A}(\mathsf{Enc}_k(m_2)) = 1]| < \epsilon.$$

In the Information-theoretic world, instead of 1, we had m_1 and m_2. The reason of not having m_1 and m_2 is to get rid of dependencies the adversary put on the messages.

The translation from Shannon secrecy to indistinguishable encryptions was fairly straightforward. We'll see that it is more complicated to translate Perfect secrecy. We'll make several attempts at a good definition.

Moving from Perfect secrecy to computational world

We first recall the definition of Perfect Secrecy

Definition 2. Enc*. satisfies Perfect Secrecy if for any m and message distribution M,* $\Pr[M = m | \mathsf{Enc}(K, M) = c] = Pr[M = m]$*.*

The first thing we need to do is add error (we showed in the previous class the attacker can always have some success. So this looks like:

Definition 3. Enc*. satisfies Perfect Secrecy if for any m and message distribution M,*

$$|\Pr[M = m | \mathsf{Enc}(K, M) = c] - \Pr[M = m]| < \epsilon.$$

3 An attacker trying to predict that message

We now need to introduce the concept of a machine predicting (instead of just being based on the probability). Let's take a first attempt:

Definition 4. *Let* $\mathcal{M}$ *be a message space. Let K be a distribution.* Enc *is secure if for all probabilistic polynomial time Turing machines* $\mathcal{A}$ *or PPT* $\mathcal{A}$ *if for any message distribution M over* $\mathcal{M}$*,*

$$\Pr[\mathcal{A}(C) = m] - \Pr[M = m]| < \epsilon.$$

This message distribution might be very complex. Even if we remove the ciphertext the adversary might not be able to output messages m with the correct probability (what if it takes exponential time to sample from M?). So the adversary might not be able to do this just because the message distribution is complex, not because the ciphertext is hiding anything.

Some notes:

1. It might not be possible to come up with message distribution, and makes this impossible to satisfy.
2. Haven't explicitly hidden all functions of message just the ability to predict the message. In the information-theoretic setting since the probability didn't change at all this implicitly hid all functions of the message.
3. We need to incorporate the attacker having some knowledge of the message distribution.

Attempt 2 Notice that we are now implicitly considering a sequence of encryption schemes, lets make this a little bit more formal. Define an encryption scheme that takes in security parameter. Notice this definition only talks about computing the message itself. What if some sensitive function of the message is revealed? Is this prohibited? Recall that we wanted to protect every part of the message. Notionally we want this to be true for every function of the message. Lets take a second attempt.

Definition 5. *Let $\mathcal{M}$ be a message space, K be a distribution.* Enc *is secure if for all $f : \mathcal{M} \to \{0,1\}^*$ and for all PPT A for any message distribution M over $\mathcal{M}$ and for all t*

$$\Pr[\mathcal{A}(C) = t)] - \Pr[f(M) = t]| < \epsilon.$$

This definition addresses issue 2 above. Lets consider some possible functions f:

1. The identity function.
2. A bit of the message.
3. The sum of some bits of the message.
4. An encryption of m under some message.

Attempt 3 In the case of perfect secrecy we could explain a priori message information as a new message distribution. However, in the computational setting, (even if we define a new adversary for each message distribution) its not clear they can efficiently make use of this information. We need to formally include this information.

Definition 6. *Let $\mathcal{M}$ be a message space. Let K be a distribution.* Enc *is secure if for all $f : \mathcal{M} \to \{0,1\}^*$ and for all PPT $\mathcal{A}$ if for any message distribution M over $\mathcal{M}$ and for any $h : \mathcal{M} \to \{0,1\}^*$, and for all t*

$$\Pr[\mathcal{A}(C, h(M)) = t)] - \Pr[f(M|h(M)) = t]| < \epsilon.$$

However we still have the issue that it might not be possible for the attacker to even create the output distribution of f (what if f takes exponential time to compute?)

Attempt 4 Informally, we want that the probability that the attacker guesses the function's value given the ciphertext = probability that the attacker guesses the function's value without given the ciphertext

Definition 7. *Let $\mathcal{M}$ be a message space. Let K be a distribution.* Enc *is secure if for all $f : \mathcal{M} \to \{0,1\}^*$ and for all PPT $\mathcal{A}$ there exists PPT $\mathcal{A}'$ if for any message distribution M over $\mathcal{M}$ and for any $h : \mathcal{M} \to \{0,1\}^*$,*

$$\Pr[\mathcal{A}(C, h(M)) = f(M))] - \Pr[\mathcal{A}'(h(M)) = f(M)]| < \epsilon.$$

The second attacker A' is called a *simulator* since it is trying to recreate A's attack without the ciphertext. The formal definition used is slightly different than this

Definition 8 (Semantic Security)**.** *[GM84] Let $\mathcal{M}$ be a message space. Let K be a distribution.* Enc *is* semantically-secure *if for all PPT $\mathcal{A}$ there exists a simulator $\mathcal{A}'$ such that for any message distribution M over $\mathcal{M}$ and for any $f, h : \mathcal{M} \to \{0,1\}^*$,*

$$\Pr[\mathcal{A}(C, h(M)) = f(M))] - \Pr[\mathcal{A}'(h(M)) = f(M)]| < \epsilon.$$

We'll show in the next class that an encryption scheme has indistinguishable encryptions if and only if it satisfies semantic security.

Theorem 1. *Indistinguishable Encryptions is equivalent to Semantic Security*

As a preview we will need to work with Turing machines here. We'll show that an attacker that breaks one definition can be used to break the other definition. This is called a reduction. Most of our proofs will proceed using either contradiction or the contrapositive. To prove the theorem, we need two things to prove, I.E $\rightarrow$ S.S. and S.S.$\rightarrow$ I.E.

References

[GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.