CS 4501-6501 Topics in Cryptography 26 Jan 2018

Lecture 2: Perfect Secrecy and its Limitations

Lecturer: Mohammad Mahmoody *Scribe: Mohammad Mahmoody*

1 Introduction

Last time, we informally defined encryption schemes. This time, we start by defining encryption schemes more formally. For now, we focus on a simplified version in which the encryption and decryption algorithms do not use any extra randomness, and we also aim for a weaker notion of security that only deals with encryption of only *one* message. We will extend the following definition to the more general cases later on.

Definition 1.1 ((Deterministic) Encryption)**.** A (deterministic private-key) encryption scheme $\mathcal{SKE} = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ for key space $\mathcal{K} = \{0,1\}^n$, message space (i.e., plain-text space) $\mathcal{M}$, and cipher-text space $\mathcal{C}$ consists of three algorithms.

- Gen is a randomized algorithm, and outputs a "uniformly" chosen random $k \in \mathcal{K}$.
- $\mathrm{Enc}(k,m)$, takes $k \in \mathcal{K}, m \in \mathcal{M}$, and outputs $c \in \mathcal{C}$.
- $\mathrm{Dec}(k,c)$, takes $k \in \mathcal{K}, c \in \mathcal{C}$, and outputs some $m \in \mathcal{C}$.

Completeness condition. We say that $\mathcal{SKE}$ is *complete* (i.e., it is correct) if for all $k \in \mathcal{K}, m \in \mathcal{M}$, it holds that

$$\mathrm{Dec}(k, \mathrm{Enc}(k,m)) = m.$$

The more general version of encryption that we will work with later on will allow the encryption and decryption algorithms to be randomized algorithms, so that even encryption a single message m might end up with many possible cipher-texts, but for now, we will only work with deterministic encryption and decryption.

Secrecy. In the following, we will aim at defining perfect secrecy for encryption. The intuition is that the cipher-text, in eyes of Eve who does not know k, should look completely *irrelevant* to m. To formally define this, we need to go over some basics of probability theory.

2 Basics of Probability Theory

Probability theory allows us to formally talk about uncertainty and chances by which an adversary succeeds. Here we review the basic tools from probability theory, but you can use many great sources on this subject that are freely available online.

Definition 2.1 (Finite Probability Space, Events, Random Variables)**.** A (finite) probability space (Ω, p) consists of a finite set Ω and a mapping $p \colon \Omega \mapsto [0, 1]$, such that $\sum_{x\in\Omega} p[x] = 1$. Intuitively, $p[x]$ is the chance of sampling x from this space. We use various notations $\Pr[x] = p_x = p[x] = p(x)$, to denote the same thing. An event $E \subseteq \Omega$, is a subset of Ω, and we define the probability of that event to be $\Pr[E] = \sum_{x\in E} \Pr[x]$. For a probability space (Ω, p), we can define a random variable X, and by $x \leftarrow X$ we define the random process of selecting x according to the distribution of X, which is defined by (Ω, p). Namely, if we select $x \leftarrow X$, then the probability of getting $X = x$ is exactly $\Pr[x]$. To clarify further, we sometimes write $\Pr_X[x] = \Pr[x = X]$ to denote the same thing. Sometimes, we prefer to start the definition from a random variable, in which case we can say that X is a random variable, with support set $\mathrm{Supp}(X) = \{x \mid \Pr[X = x] > 0\}$. By $X \equiv Y$ we mean that X, Y are random variables with the same distributions.

We can also talk about the probability of one event $E_!$, when we are given the guarantee that another event E_2 has already happened. The following definition formalizes this notion.

Definition 2.2 (Conditional Probability and Independent Events)**.** For probability space (Ω, p) and events E_1, E_2, where $\Pr[E_2] > 0$, the conditional probability $Pr[E_1 \mid E_2]$ is defined as $\Pr[E_1 \mid E_2] = \Pr[E_1 \cap E_2]/\Pr[E_2]$. Intuitively, if we restrict ourselves to E_2, then $\Pr[E_1 \mid E_2]$ defines the probability of E_1 in that context. We sometimes use $E_1 \wedge E_2$ (reading E_1 *and* E_2) instead of $E_1 \cap E_2$ to denote the same thing.

Sometimes, knowing whether or not E_2 has happened does not say anything about E_1. This is formalized as follows:

Definition 2.3 (Independent Events)**.** For nonzero E_1, E_2 events E_1 is independent of E_2 if $\Pr[E_1 \mid E_2] = \Pr[E_1]$

Do it yourself: Check the correctness of the following proposition.

Proposition 2.4. *E_1 is independent of E_2 if and only if $\mathbf{P}[E_1 \wedge E_2] = \Pr[E_1] \cdot \Pr[E_2]$. Therefore, by symmetry, E_1 is independent of E_2 if and only if E_2 is independent of E_1.*

Sometimes we deal with multiple objects, all from a jointly defined probability space. (For example, when we encrypt a randomly selected message, and this leads to a distribution over cipher-texts.)

Definition 2.5 (Marginal probabilities)**.** Suppose $\Omega = \Omega_1 \times \Omega_2$ be the Cartesian product of two sets Ω_1, Ω_2, and suppose p is a distribution over Ω. Then, intuitively, sampling according to p from Ω, means sampling a pair $(x_1, x_2) \in \Omega_1 \times \Omega_2$ where x_1 is distributed according to some distribution p_1 over Ω_1 and x_2 is distributed according to some distribution p_2 over Ω_2. We call p_1 the marginal probability distribution of x_1 and p_2 the marginal probability distribution of x_2. It is easy to compute p_1 and p_2 using p as follows:

$$p_1[x_1] = \sum_{x_2\in\Omega_2} p[(x_1, x_2)] \quad , \quad p_2[x_2] = \sum_{x_1\in\Omega_1} p[(x_1, x_2)].$$

Sometimes, we use a reversed way to define random variables. Namely, we say that X_1, X_2 are two random variables, and they also have a *joint* distribution $X = (X_1, X_2)$.

Now that we have defined the notion of random variables and independence of events, we can talk about independent random variables.

Definition 2.6 (Independent random variables)**.** Suppose X_1, X_2 are two random variables jointly distributed as $X = (X_1, X_2)$. We say that X_1 and X_2 are independent, if for all $x_1 \in \text{Supp}(X_1), x_2 \in \text{Supp}(X_2)$, the two events $E_1 = \{X \mid X_1 = x_1\}, E_2 = \{X \mid x_2 = x_2\}$ are independent events. Namely, for all such x_1, x_2, it holds that

$$\Pr[X_1 = x_1 \wedge X_2 = x_2] = \Pr[X_1 = x_1] \cdot \Pr[X_2 = x_2].$$

Definition 2.7 (Conditional random variables)**.** For jointly defined random variables $(X_1, X_2) = X$, and for $x_2 \in \text{Supp}(X_2)$, by the conditional random variable $X_1' \equiv (X_1 \mid x_2)$ we mean the distribution that samples x_1 with probability $\Pr_X[E_1 \mid E_2]$ where $E_1 = \{X \mid X_1 = x_1\}, E_2 = \{X \mid X_2 = x_2\}$. We simply write $\Pr[X_1 = x_1 \mid x_2]$ to denote the same probability $\Pr[E_1 \mid E_2]$.

Do it yourself: Check the correctness of the following proposition. It implies that two random variables are independent, if and only if, knowing the value of of one of them, does not change the (conditional) distribution of the other one.

Proposition 2.8 (What independence means)**.** *X_1 and X_2 are independent if and only if, for all $x_2 \in \text{Supp}(X_2)$, the conditional distribution of $(X_1 \mid x_2)$ is the same as the (marginal) original distribution X_1.*

3 Perfect Secrecy: 1st Try

Now that we have the tools of basic probability theory at hand, we can talk about the first definition of perfect secrecy, which is based on *semantic security.*

Definition 3.1 (Perfect Semantic Secrecy)**.** Suppose $\mathcal{SKE} = (\text{Gen}, \text{Enc}, \text{Dec})$ is an encryption algorithm for key space $\mathcal{K} = \{0,1\}^n$, message space (i.e., plain-text space) $\mathcal{M}$, and cipher-text space $\mathcal{C}$. We say that it has perfect semantic secrecy, if for every distribution M over the plain-text space, the distribution C that is imposed over the cipher-text space, is independent of M.

Note that, we always sample the key $k \leftarrow \mathcal{K}$ at random. So, that determines the marginal distribution of the key which we can denote by K, in which case (K, M, C) would be three random variables jointly distributed, and K, M would be independently sampled, however C is a deterministic function of both of (K, M). So, the above definition requires that M, C be independent.

3.1 One-Time-Pad

A perfectly semantically secure encryption can be obtained as follows if the key and message space are the same $\mathcal{K} = \{0,1\}^n = \mathcal{M}$ (i.e., of the same length).

Definition 3.2 (One time pad encryption)**.** OTP encryption is defined as follows:

- $OTPEnc(k, x) = k \oplus x$ where $\oplus$ is the *bit wise* exclusive OR.
- $OTPDec(k, c)$ is defined similarly.

Do it yourself: Check that the above encryption algorithm is complete. Namely, encrypting and decryption a message leads to the same thing.

Do it yourself: Prove that one-time pad is perfectly semantically secure. Hint: by Proposition 2.8, it is enough to show that for every plain-text x, the cipher-text distribution $C_x \equiv C$ remains the same.

Unfortunately, Shannon proved that a perfectly secret scheme according to semantic security suffers from a huge downside, which is evident in the OTP encryption: the key should be as long as the message itself.

Theorem 3.3 (Perfect semantic security implies long keys [Sha49])**.** *Suppose* $\mathcal{SKE} = (\text{Gen}, \text{Enc}, \text{Dec})$ *is an encryption algorithm for key space* $\mathcal{K} = \{0,1\}^n$*, message space (i.e., plain-text space)* $\mathcal{M}$*, and cipher-text space* $\mathcal{C}$*. if* $\mathcal{SKE}$ *is perfectly semantically secure, then* $|\mathcal{M}| \leq \mathcal{K}$*. Namely,* $\mathcal{M}$ *cannot contain all of* $n+1$ *bit long messages.*

Proof. Suppose M is the uniform distribution over $\mathcal{M}$. By perfect semantic secrecy, M should be independent random variable from cipher-text distribution C, so for every possible cipher-text $c \in \mathcal{C}$, we should have $(M \mid c) \equiv C$. Namely, the *conditional* distribution of M conditioned on knowing the $c = C$ should remain the same. Since M was the uniform distribution, so every $m \in \mathcal{M}$ should have the same $1/|\mathcal{M}| > 0$ probability even conditioned on $C = c$. However, note that when we know $C = c$, the total possible number of messages that we can *decrypt* to, is bounded by $|\mathcal{K}|$, because for each k, we have a unique $\text{Dec}(k, c)$. (Note that here we used the completeness of the scheme.) Therefore, the number k of messages that have non-zero probability of being the actual message, conditioned on $C = c$ is at most $|\mathcal{K}|$. On the other hand $k = |\mathcal{M}|$. Therefore, $|\mathcal{M}| \leq |\mathcal{K}|$. □

4 Perfect Secrecy: 2nd Try

Knowing the devastating limitation of perfect secrecy according to semantic security, here we aim at a different alternative definition based on a *security game.* The idea is that, we allow the Eve even to choose two particular messages, and then try to guess which one is encrypted.

Indistinguishability security game. In this game, we proceed as follows between and adversary and a challenger.

1. The adversary chooses $m_0, m_1 \in \mathcal{M}$ both from the same fixed message space. (When we deal with arbitrary length message later on, these two should be of the same length). Then Eve sends these to to the *Challenger*.
2. Challenger picks a random key $k \leftarrow \mathcal{K}$ and a bit $b \in \{0,1\}$ at random, and sends over $c = \mathrm{Enc}(k, m_b)$ to Eve.
3. Adversary outputs a bit b' and wins if $b = b'$. We say that the adversary wins if $b = b'$.

Definition 4.1 (Perfect indistinguishability)**.** Suppose $\mathcal{SKE} = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ is an encryption algorithm for key space $\mathcal{K} = \{0,1\}^n$, message space (i.e., plain-text space) $\mathcal{M}$, and cipher-text space $\mathcal{C}$. We say that $\mathcal{SKE}$ has perfect indistinguishability if for all adversaries, the probability of winning in the indistinguishability security game is at most $1/2$.

Unfortunately, even though the new definition looks different, it is equivalent to previous definition of perfect semantic secrecy. (Check it yourself.) So, how about we relax it a little bit so that we get around the impossibility result (of encrypting long messages with short keys)?

ε-indistinguishability. We say that $\mathcal{SKE} = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ is ε-ind secure, if we modify the definition and only require that the adversary wins with probability at most $(1+\varepsilon)/2$ (rather than $1/2$). Note that if we choose $\varepsilon = 2^{-100}$, this is essentially a *negligible* probability that will never affect anything in real life.

Unfortunately, a modification of Shannon's lower bound (of Theorem 3.3) still applies and says that even if we choose something like $\varepsilon = 1/10$, then still we need the key to be at least half of the message's length. (It is a good exercise to think about it, and try to prove this yourself). However, as we will see, this modification of perfect-ind definition will still come to help us later on when we will introduce also *another* relaxation as well.

Computational limitation on adversary. The second relaxation that we apply to get around the key-length barrier is to put computational limitations on the adversary. So far we assumed that the adversary can do arbitrary computation. However, this does not happen in real life. Namely, we are OK if an adversary that runs in time 2^{100} has 2^{-100} chance of doing something harmful. We will explore this direction from next time, and basically for the rest of the semester.

References

[Sha49] Claude E Shannon. Communication theory of secrecy systems. *Bell Labs Technical Journal*, 28(4):656–715, 1949.