**CSA E0 235: Cryptography** 19th January 2016

Lecture 3

*Instructor: Arpita Patra* *Scribe: Atlanta Chakraborty*

# 1 Review

From our discussion of Secret Key Encryption (SKE) in light of Modern Cryptography, we recall the definition of Perfect Security by assuming that our adversary has unbounded computational power and also can only eavesdrop the cipher text during transit, also called as Ciphertext Only Attack (COA).

**Definition 1.1.** *(Perfectly Secret). An encryption scheme* (Gen, Enc, Dec) *over a message space $\mathcal{M}$ is perfectly secret if for every probability distribution over $\mathcal{M}$, every message $m \in \mathcal{M}$, and every ciphertext $c \in \mathcal{C}$ for which $Pr[C = c] > 0$:*

$$Pr[M = m|C = c] = Pr[M = m].$$

This basically means that the probability with which the adversary knows a plain-text remains the same before (a priori probability) and after (posteriori probability) seeing the cipher-text. Thus the adversary gets no advantage on seeing the cipher text.

Accordingly, we have seen that the Vernam Cipher (or *one-time pad*) achieves this level of security with few drawbacks.

1) The key must be as long as the message.

2) The key can be used only once to encrypt a single message securely and hence its name (OTP).

Things start getting really messy when one reuses a one-time pad. As a matter of fact, US and UK exploited this drawback to decrypt Russian plaintext in the Venona Project. So we next thought of designing schemes which would overcome these drawbacks, but unfortunately, the aforementioned drawbacks are inherent to any scheme achieving perfect secrecy.

# 2 Limitations of Perfect Secrecy

In this section, we shall prove that the following theorem is inherent to any scheme achieving perfect-security.

**Theorem 2.1.** *In any perfectly-secure encryption scheme defined by* (Gen, Enc, Dec)*, the key space $\mathcal{K}$ must be atleast as large as the message space $\mathcal{M}$ i.e. $|\mathcal{K}| \geq |\mathcal{M}|$.*

*Proof.* Assume $|\mathcal{K}| < |\mathcal{M}|$. Also assume that the message space has such a distribution wherein every message occurs with non zero probability. Let $c \in \mathcal{C}$ be a ciphertext that occurs with non-zero probability. Define a new set $\mathcal{M}(c)$ which contains all possible messages that are decryptions of $c$, i.e.,

$$\mathcal{M}(c) := \{m | m = \text{Dec}_k(c) \text{ for some } k \in \mathcal{K}\}.$$

Thus clearly, $|\mathcal{M}(c)| \leq |\mathcal{K}|$ since for each message there will be atleast one key $k \in \mathcal{K}$ for which $m = \text{Dec}_k(c)$ since Dec is deterministic, basically a particular message can be encrypted by more than one key. Also $|\mathcal{K}| < |\mathcal{M}|$ from our assumption. Together they indicate that there exists atleast one message $m' \in \mathcal{M}$ which cannot be encrypted by any key, i.e. $m' \notin \mathcal{M}(c)$. But then

$$Pr[M = m'|C = c] = 0 \neq Pr[M = m'],$$

which contradicts Definition 1.1 and implies that the scheme is not perfectly secret. □

One interesting observation is that the Vernam cipher (OTP) is optimal key length-wise. From the other limitation of key-reusability, which we shall not prove here, we can also deduce the fact that OTP is optimal key usability-wise too.

# 3 Equivalent Definitions of Perfectly Secure

We shall now give several equivalent formal definitions of perfect security, which basically captures different intuitions for achieving the same goal- perfectly secure scheme. Some of these definitions will be very handy to prove or disprove whether a SKE is perfectly secure or not.

**Definition 3.1.** *(Perfect Indistinguishability). An encryption scheme* (Gen, Enc, Dec) *over a message space $\mathcal{M}$ is perfectly secret if for every probability distribution over $\mathcal{M}$, every message $m_0, m_1 \in \mathcal{M}$, and every ciphertext $c \in \mathcal{C}$:*

$$Pr[C = c|M = m_0] = Pr[C = c|M = m_1].$$

This formulation states that the probability distribution over $\mathcal{C}$ is independent of the plaintext. Also, the ciphertext contains no information about the plaintext. It becomes impossible to distinguish an encryption of $m_0$ from an encryption of $m_1$ as the distribution over the ciphertext depends only on the choice of key and randomness of Enc when it is probabilistic, thus being the same for both messages $m_0$ and $m_1$, hence known as perfect indistinguishability.

There is another notion of perfect security given by C. E. Shannon. He then used it to prove perfect security for OTP.

**Definition 3.2.** ***(Shannon's Theorem).*** *An encryption scheme* (Gen, Enc, Dec) *over a message space $\mathcal{M}$ with $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$ is perfectly secret if and only if:*

*(i) Every key $k \in \mathcal{K}$ is chosen with equal probability $1/|\mathcal{K}|$ by algorithm* Gen.

*(ii) For every $m \in \mathcal{M}$ and every $c \in \mathcal{C}$, there exists a unique key $k \in \mathcal{K}$ such that* $\text{Enc}_k(m)$ *outputs* $c$.

**Remark:** The need for the assumption is as follows:

For a perfectly secure scheme we have $|\mathcal{K}| \geq |\mathcal{M}|$ as proved in Theorem 2.1. Also for correctness to hold we need $|\mathcal{C}| \geq |\mathcal{M}|$ otherwise we would have different ciphertexts corresponding to the same message which would thus lead to indistinguishability. Thus $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$ is the ideal optimal case.

*Proof.* We first assume that every key is obtained with probability $1/|\mathcal{K}|$ and that for every $m \in \mathcal{M}$ and $c \in \mathcal{C}$ there exists a unique key $k \in \mathcal{K}$ such that $\text{Enc}_k(m) = c$. It follows from Definition 1.1 that for every $m$ and $c$,

$$Pr[C = c|M = m] = Pr[K = k] = 1/|\mathcal{K}|$$

irrespective of the probability distribution over $\mathcal{M}$. Thus

$$\begin{aligned} Pr[C = c] &= \sum_{m \in \mathcal{M}} Pr[C = c|M = m]Pr[M = m] \\ &= 1/|\mathcal{K}| \sum_{m \in \mathcal{M}} Pr[M = m] \\ &= 1/|\mathcal{K}|. \end{aligned} \tag{1}$$

By applying Bayes' Theorem, we get

$$\begin{aligned} Pr[M = m|C = c] &= \frac{Pr[C = c|M = m]Pr[M = m]}{Pr[C = c]} \\ &= Pr[M = m], \end{aligned} \tag{2}$$

thus satisfying the definition of perfect security.

To prove the other direction, let (Gen, Enc, Dec) be as in the theorem and assume Enc to be deterministic. We now prove that if (Gen, Enc, Dec) is perfectly secret, then $(i)$ and $(ii)$ hold.

Let $\mathcal{M} = \{m_1, m_2, ...\}$ and $c$ be a ciphertext that occurs with non-zero probability for some message. Let $\mathcal{K}_i$ be the set of all keys that maps $m_i$ to some $c$ i.e. $\text{Enc}_k(m_i) = c$ if and only if $k$ belongs to $\mathcal{K}_i$. We now claim that $\mathcal{K}_i \neq \varnothing$ i.e. there exists atleast one key that can be used for encryption of a message and $\mathcal{K}_i \cap \mathcal{K}_j = \varnothing$ which means that the same key cannot be used for encrypting more than one message. We assume that $Pr[C = c|M = m] > 0$, for some $m$. For arbitrary $c$, $m_i$ and $m_j$,, from our definition of perfect security we have,

$$Pr[C = c|M = m_i] = Pr[C = c|M = m_j].$$

This implies that $\mathcal{K}_i \neq \varnothing$. Now, assume that the same key $k$ maps both the messages $m_i$ and $m_j$ to $c$. For correctness to hold good, our assumption would be contradicted since it would be impossible to decrypt it correctly. Thus $\mathcal{K}_i \cap \mathcal{K}_j = \varnothing$. Since $|\mathcal{M}| = |\mathcal{K}|$, we get that $|\mathcal{K}_i| = 1$. Thus there exists a unique key $k$ for every $m \in \mathcal{M}$ and every $c \in \mathcal{C}$ such

that $\text{Enc}_k(m) = c$. This proves condition $(ii)$. Also we know from the definition of perfect security that

$$Pr[C = c|M = m_i] = Pr[C = c|M = m_j]$$

which implies that

$$Pr[K = k_i] = Pr[K = k_j].$$

Each key in the key space is equally likely hence chosen uniformly by Gen with probability $1/|\mathcal{K}|$. This proves condition $(i)$. □

In a Vernam cipher, suppose our key $k \in \mathcal{K}$ is a $l-$bit binary string and our message $m \in \mathcal{M}$ is $m-$bit binary string, then we have $|\mathcal{K}| = 2^l$ and also $|\mathcal{M}| = 2^m$. Theorem 2.1 states that

$$|\mathcal{K}| \geq |\mathcal{M}|$$

which implies that

$$2^l \geq 2^m.$$

Hence,

$$l \geq m,$$

i.e. the length of the key must be as large as the message length, which we have already seen as the drawback of the Vernam cipher. Hence perfect secrecy is achieved only for the optimal case when $l = m$ which can be further verified using Shannon's theorem.

### 3.1 Use of Shannon's Theorem

Definition 3.2 completely characterizes a perfectly-secret encryption scheme.

1. The conditions $(i)$ and $(ii)$ of Definition 3.2 are easy to check.
2. There is no need to analyse any probability distribution in contrast to working with Definition 1.1.

There is yet another equivalent definition of perfect secrecy which is based on an experiment, call it $\text{PrivK}^{\text{eav}}_{\mathcal{A},\Pi}$ defined for $\Pi = (\text{Gen, Enc, Dec})$ over a message space $\mathcal{M}$, involving a private-key encryption setting, an eavesdropping adversary $\mathcal{A}$ who tries to break a cryptographic scheme and an imaginary tester who checks if the adversary succeeds.

The experiment is as follows:

1. The adversary $\mathcal{A}$ is given the freedom to choose any pair of messages $m_0,\ m_1 \in \mathcal{M}$.
2. The tester then chooses a random key $k$ generated by Gen and also tosses a coin to select a random bit $b \to \{0, 1\}$. Then a cipher text $c \to \text{Enc}_k(m_b)$ is computed and given to $\mathcal{A}$.
3. The adversary $\mathcal{A}$ then guesses which message was sent and outputs $b' \to \{0, 1\}$.
4. $\mathcal{A}$ succeeds only when $b' = b$, also defined as $\text{PrivK}^{\text{eav}}_{\mathcal{A},\Pi} = 1$.

Thus the adversary knows for sure that only of the messages $m_0$ or $m_1$ will be communicated with probability 1/2. It is always possible for $\mathcal{A}$ to succeed in the experiment by guessing $b'$ randomly with probability 1/2. Thus here we state that perfect secrecy is achieved if no adversary $\mathcal{A}$ can succeed with probability any better than half.

**Definition 3.3.** *(**Adversarial Indistinguishability**). An encryption scheme* (Gen, Enc, Dec) *over a message space $\mathcal{M}$ is perfectly secret if for every adversary $\mathcal{A}$ it holds that*

$$Pr[\text{PrivK}^{\text{eav}}_{\mathcal{A},\Pi} = 1] = \frac{1}{2}.$$

# 4 Conclusion

Since the limitations discussed in the introduction of the lecture are inherent to any perfectly secure encryption scheme, there is no point in giving the adversary more power in attacking a protocol. The hurdles in achieving perfect security outweighs the strength of perfect security. Therefore, we shall bring about a few relaxations in our assumptions of the adversary, namely,

1. Our assumption of the adversary having unbounded computational power is now reduced to being polynomially bounded.

2. We are also willing to accommodate for a break with a very small probability.

Hence, there is a need to find an alternative relaxed security notion different from perfect security. This gave rise to the birth of computational security. However there are a few compromises that we certainly have to make:

1. Perfectly secure schemes are very efficient.

2. Perfectly secure schemes are much faster as compared to the computationally secure protocols.

Thus we conclude by saying that although a computationally secure scheme comes with a certain price of efficiency, it is of utmost significance to examine the computational approach since it helps bypass the inherent limitations of perfect secrecy by allowing the usage of short keys and also permitting key reusability.