CS-553/CSL-505 – Cryptography August 21, 2024

Lecture Note: Perfect Secrecy

Instructor: Dr Dhiman Saha

A cryptosystem has perfect secrecy if the interception of a ciphertext gives the cryptanalyst no information about the underlying plaintext and no information about any future encrypted messages.
Notation:

1. M: A random variable defined over the Message space $\mathcal{M}$.
2. C: A random variable defined over the ciphertext space $\mathcal{C}$.
3. K: A random variable defined over the key space $\mathcal{K}$.
4. $f(x)$: The function f is a probability density function of a random variable X. It denotes the probability that $X = x$, i.e., $Pr[X = x]$.
5. $f(m|c)$: A shorthand of the conditional probability $f(M = m|C = c)$.

Theorem 1 (Bayes's Theorem) *Let X and Y be two random variables. Then the following relation holds true for conditional probabilities $f(x|y)$ and $f(y|x)$;*

$$f(x|y) = \frac{f(y|x) \cdot f(x)}{f(y)}.$$

Bayes's formula is useful if we know the conditional probability of X on Y and want to know the reverse conditional probability of Y on X.
Before defining perfect secrecy, it is worth noting that the key is chosen before the message comes into the picture. Hence, both random variables should be assumed as independent random variables. Let us assume that there is a certain distribution in the plaintext space. Furthermore, keys are sampled according to some distribution of the user's choice. Encryption & decryption functions determine the distribution of ciphertext space using distributions of K & M as follows:

$$f(c) = \sum_{k \in \mathcal{K}: c = e_k(m) \text{ for some } m \in \mathcal{M}} f(k) f(d_k(c))$$

Now, let us look at one example of a cryptosystem that is not perfectly secure.

Example 1 *Let $\mathcal{M} = \{a, b\}$ with $f(a) = 1/4, f(b) = 3/4$. Let $\mathcal{K} = \{k_1, k_2, k_3\}$ with $f(k_1) = 1/2, f(k_2) = f(k_3) = 1/4$. Let $\mathcal{C} = \{1, 2, 3, 4\}$, and suppose the encryption functions are defined to be $e_{k_1}(a) = 1, e_{k_1}(b) = 2; e_{k_2}(a) = 2, e_{k_2}(b) = 3$;and $e_{k_3}(a) = 3, e_{k_3}(b) = 4$. This cryptosystem can be represented by the following encryption matrix:*

	a	b
k_1	1	2
k_2	2	3
k_3	3	4

Table 1: A Cryptosystem

Now, we compute the probability distribution on $\mathcal{C}$. We obtain the following:

$$f(1) = f(k_1).f(a) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$
$$f(2) = f(k_2).f(a) + f(k_1).f(b) = \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{3}{4} = \frac{7}{16}$$
$$f(3) = f(k_3).f(a) + f(k_2).f(b) = \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{3}{4} = \frac{1}{4}$$
$$f(4) = f(k_3).f(b) = \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{16}.$$

Computing conditional probability distributions on ciphertexts, given that a certain message has been observed, is straightforward using the following formula:

$$f(c|m) = \sum_{k:e_k(m)=c} f(k)$$

Now we can compute the conditional probability distributions on the plaintext, given that a certain ciphertext has been observed. This can be computed using the Bayes's formula.

$$f(a|1) = 1 \qquad f(b|1) = 0$$
$$f(a|2) = \frac{f(2|a)f(a)}{f(2)} = \frac{f(k_2)f(a)}{f(2)} = \frac{1}{7} \qquad f(b|2) = \frac{6}{7}$$
$$f(a|3) = \frac{1}{4} \qquad f(b|3) = \frac{3}{4}$$
$$f(a|4) = 0 \qquad f(b|4) = 1$$

In the above example, as we can observe that $f(a|c) \neq f(a)$ for $c = 1, 2, 4$ and $f(b|c) \neq f(b)$ for $c = 1, 2, 4$. From conditional probability distribution, it is clear that the conditional probability of the plaintext depends upon observed ciphertext which is not good for the perfect secrecy.

Definition 1 (Perfect Secrecy) *A cryptosystem has perfect secrecy if*

$$f(m|c) = f(m), \quad \forall m \in \mathcal{M} \ \& \ \forall c \in \mathcal{C}.$$

It says that the probability of any particular plaintext, $Pr(M = m)$, is independent of the ciphertext. Intuitively, this means that the ciphertext reveals no new knowledge of the plaintext.
According to the Bayes's theorem stated previously, we have the following equivalent condition of perfect secrecy:

$$f(c|m) = f(c), \quad \forall m \in \mathcal{M} \ \& \ \forall c \in \mathbb{C}, f(m) \neq 0.$$

In the example 1, the condition for perfect secrecy is satisfied for only ciphertext $c = 3$. However, it must hold for all possible ciphertext as stated by the definition 1.

Proposition 1 (Necessary Condition) *If a cryptosystem has perfect secrecy, then* $\#\mathcal{K} \geq \#\mathcal{C}^+$, *where* $\mathcal{C}^+ = \{m \in \mathcal{M} : f(m) > 0\}$ *is the set of plaintexts that have a positive probability of being selected.*

Theorem 2 *Suppose that a cryptosystem satisfies* $\#\mathcal{K} = \#\mathcal{M} = \#\mathcal{C}$, *i.e., the numbers of keys, plaintexts, and ciphertexts are all equal. Then the system has perfect secrecy if and only if the following two conditions hold:*

1. *Each key* $k \in \mathcal{K}$ *is used with equal probability.*
2. *For a given message* $m \in \mathcal{M}$ *and ciphertext* $c \in \mathcal{C}$, *there is exactly one key* $k \in \mathcal{K}$ *that encrypts* m *to* c.

Example 2 (Perfectly Secure Shift Cipher) *Consider the Shift Cipher described in previous lectures. Suppose that each of the* 26 *possible keys (shift amounts) is chosen with equal probability and that each plaintext character is encrypted using a new, randomly chosen, key. We can show that it is perfectly secure.*
Recall that $\mathcal{M} = \mathcal{C} = \mathcal{K} = \mathbb{Z}_{26}$, *and for* $0 \leq k \leq 25$, *the encryption rule* e_K *is defined as* $e_k(m) = (m + k) \mod 26$ *for any* $m \in \mathbb{Z}_{26}$.
First, we compute the probability distribution on $\mathcal{C}$. *Let* $c \in \mathbb{Z}_{26}$; *then*

$$\begin{aligned} f(c) &= \sum_{k \in \mathbb{Z}_{26}} f(k) f(d_k(c)) \\ &= \sum_{k \in \mathbb{Z}_{26}} \frac{1}{26} f(c-k) \\ &= \frac{1}{26} \sum_{k \in \mathbb{Z}_{26}} f(c-k) \\ &= \frac{1}{26} \cdot 1 = \frac{1}{26}, \text{ for any } c \in \mathcal{C} \end{aligned}$$

Next we have that for any $c \in \mathcal{C}$

$$f(c|m) = \sum_{k \in \mathbb{Z}_{26} : e_k(m) = c} f(k) = f(c - m \pmod{26}) = \frac{1}{26}$$

Now, using Bayes' theorem, it is trivial to compute

$$f(m|c) = \frac{f(c|m) f(m)}{f(c)} = \frac{1/26 \cdot f(m)}{1/26} = f(m)$$

Thus we showed that the shift cipher in aforementioned condition is perfectly secure.

Problem 1 *Suppose that a cryptosystem has two keys* k_1 *and* k_2, *three messages* m_1, m_2, *and* m_3, *and three ciphertexts* c_1, c_2, *and* c_3. *Assume that the density function for the message random variable satisfies* $f_M(m_1) = f_M(m_2) = 1/4$ *and* $f_M(m_3) = 1/2$. *The following Table 2 describes how the different keys act on the messages to produce ciphertexts:*
Show that above cryptosystem is not perfectly secure.

	m_1	m_2	m_3
k_1	c_2	c_1	c_3
k_2	c_1	c_3	c_2

Table 2: A Cryptosystem

Problem 2 *Let n be a positive integer. A Latin square of order n is an $n \times n$ array L of the integers $1, \ldots, n$ such that every one of the n integers occurs exactly once in each row and each column of L. An example of a Latin square of order 3 is shown in the Table 3: Given any Latin square L of*

1	2	3
3	1	2
2	3	1

Table 3: Latin Square

order n, we can define a related Latin Square Cryptosystem. Take $\mathcal{M} = \mathcal{C} = \mathcal{K} = 1, \ldots, n$. For $1 \leq i, j \leq n$, the encryption rule Enc is defined to be $Enc_i(j) = L(i, j)$. (Hence each row of L gives rise to one encryption rule.) Give a complete proof that this Latin Square Cryptosystem achieves perfect secrecy provided that every key is used with equal probability.

Note: This lecture note is prepared from the following two books

1. *An Introduction to Mathematical Cryptography* by Jeffrey Hoffstein, Jill Pipher Joseph, H. Silverman
2. *Cryptography: Theory and Practice* by Douglas R. Stinson, Maura B. Paterson