Lecture 10 - Cryptography and Information Theory

Jan Bouda

FI MU

May 18, 2012

Part I

Cryptosystem

Cryptosystem

- The traditional main goal of cryptography is to preserve secrecy of the message, i.e. to transform it in the way that no unauthorized person can read the message while it is easily readable by authorized persons.
- First applications of message secrecy are known from ancient times and served to keep secret military and diplomatic secrets, craftsmanship methods and also love letters.
- Craftsmanship secrets on earthen tablets in Ancient Summer.
- Secret love letters in Kamasutra.
- Spartian Scytale.
- Secrets hidden in a wax table or under hair of a slave.
- Caesar cipher.

Cryptosystem

Definition

A encryption system (cipher) is a five-tuple $(\mathbf{P}, \mathbf{C}, \mathbf{K}, \mathbf{E}, \mathbf{D})$, where

1. $\mathbf{P}$ is a finite set of possible plaintexts
2. $\mathbf{C}$ is a finite set of possible ciphertexts
3. $\mathbf{K}$ is a finite set of possible keys
4. For each $k \in \mathbf{K}$ there is an encryption rule $e_k \in \mathbf{E}$ and a corresponding decryption rule $d_k \in \mathbf{D}$. Each $e_k : \mathbf{P} \rightarrow \mathbf{C}$ and $d_k : \mathbf{C} \rightarrow \mathbf{P}$ are functions such that $d_k(e_k(x)) = x$ for every $x \in \mathbf{P}$.

Shift Cryptosystem

Example

Example is e.g. the **shift cryptosystem**, sometimes known as the Caesar cipher. In this case $\mathbf{P} = \mathbf{C} = \mathbf{K} = \mathbb{Z}_{26}$. For $0 \leq k \leq 25$ we define

$$e_k(x) = (x + k) \bmod 26 \tag{1}$$

and

$$d_k(y) = (y - k) \bmod 26 \tag{2}$$

for $x, y \in \mathbb{Z}_{26}$.

Perfect Secrecy

To derive a definition of perfect secret we assume that there is some a priori distribution on plaintexts described by the random variable X with distribution $P(X = x)$. The key is chosen independently from the plaintext and described by the random variable K. Finally, ciphertext is described by the random variable Y that will be derived from X and K. Also, for $k \in \mathbf{K}$ we define $\mathbf{C}_k = \{e_k(x) | x \in \mathbf{X}\}$ as the set of all ciphertexts provided k is the key.

Now we can explicitly calculate the probability distribution of Y as

$$P(Y = y) = \sum_{k:y \in \mathbf{C}_k} P(K = k) P(X = d_k(y)). \tag{3}$$

Another quantity of interest is the probability of a particular ciphertext given a particular plaintex, easily derived as

$$P(Y = y | X = x) = \sum_{k:x = d_k(y)} P(K = k). \tag{4}$$

Perfect Secrecy

Definition

We say that the cryptosystem $(\mathbf{P}, \mathbf{C}, \mathbf{K}, \mathbf{E}, \mathbf{D})$ achieves **perfect (unconditional) secrecy** if and only if for every $x \in \mathbf{X}$ and $y \in \mathbf{Y}$ it holds that

$$P(X = x|Y = y) = P(X = x). \tag{5}$$

In words, the a posteriori probability distribution of plaintext given the knowledge of ciphertext is the same as the a priori probability distribution of the plaintext.

Following our previous analysis we calculate the conditional probability of a (possibly insecure) cryptosystem as

$$P(X = x|Y = y) = \frac{P(X = x) \sum_{k:x=d_k(y)} P(K = k)}{\sum_{k:y \in \mathbf{C}_k} P(K = k) P(X = d_k(y))}. \tag{6}$$

Perfect Secrecy

Theorem

Suppose the 26 keys in the Shift cipher are used with equal probability 1/26. Then for any plaintext distribution the Shift cipher achieves perfect secrecy.

Proof.

Recall that $\mathbf{P} = \mathbf{C} = \mathbf{K} = \mathbb{Z}_{26}$. First we compute the distribution of ciphertexts as

$$\begin{aligned} P(Y = y) &= \sum_{k \in \mathbb{Z}_{26}} P(K = k) P(X = d_k(y)) \\ &= \sum_{k \in \mathbb{Z}_{26}} \frac{1}{26} P(X = y - k) \\ &= \frac{1}{26} \sum_{k \in \mathbb{Z}_{26}} P(X = y - k). \end{aligned} \tag{7}$$

Perfect Secrecy

Proof.

For fixed y the values $(y-k) \bmod 26$ are a permutation of $\mathbb{Z}_{26}$ and we have that

$$\sum_{k\in\mathbb{Z}_{26}} P(X=y-k) = \sum_{x\in\mathbb{Z}_{26}} P(X=x) = 1. \tag{8}$$

Thus for any $y \in \mathbf{Y}$ we have

$$P(Y=y) = \frac{1}{26}.$$

Next, we have that

$$P(Y=y|X=x) = P(K \equiv (y-x) \pmod{26}) = \frac{1}{26}$$

for every x and y. □

Perfect Secrecy

Proof.

Using the Bayes' theorem we have

$$
\begin{aligned}
P(X = x|Y = y) =& \frac{P(X = x)P(Y = y|X = x)}{P(Y = y)} = \frac{P(X = x)\frac{1}{26}}{\frac{1}{26}} \\
=& p(X = x)
\end{aligned} \tag{9}
$$

what completes the proof. □

The previous result shows that the shift cipher is unbreakable provided we use an independent key for each plaintext character.

Perfect Secrecy

- If $P(X = x_0) = 0$ for some $x_0 \in \mathbf{P}$, then we trivially obtain $P(X = x_0|Y = y) = P(X = x_0)$. Therefore we consider only elements such that $P(X = x) > 0$.
- For such plaintexts we observe that $P(X = x|Y = y) = P(X = x)$ is equivalent to $P(Y = y|X = x) = P(Y = y)$.
- Let us suppose that $P(Y = y) > 0$ for all $y \in \mathbf{C}$. Otherwise y can be excluded from $\mathbf{C}$ since it is useless.
- Fix $x \in \mathbf{P}$. For each $y \in \mathbf{C}$ we have $P(Y = y|X = x) = P(Y = y) > 0$. Therefore for each $y \in \mathbf{C}$ there must be some key $k \in \mathbf{K}$ such that $y = e_k(x)$. It follows that $|\mathbf{K}| \geq |\mathbf{C}|$.
- The encryption is injective giving $|\mathbf{C}| \geq |\mathbf{P}|$.

Perfect Secrecy

Theorem (Shannon)

Let $(\mathbf{P}, \mathbf{C}, \mathbf{K}, \mathbf{E}, \mathbf{D})$ *be a cryptosystem such that* $|\mathbf{P}| = |\mathbf{C}| = |\mathbf{K}|$. *Then the cryptosystem provides perfect secrecy if and only if every key is used with equal probability* $1/|\mathbf{K}|$, *and for every* $x \in \mathbf{P}$ *and every* $y \in \mathbf{C}$, *there is a unique key* k *such that* $e_k(x) = y$.

Proof.

Let us suppose the given cryptosystem achieves a perfect secrecy. As argued above for each x and y there must be at least one key such that $e_k(x) = y$. We have the inequalities

$$|\mathbf{C}| = |\{e_k(x) : k \in \mathbf{K}\}| \leq |\mathbf{K}|. \tag{10}$$

□

Perfect Secrecy

Proof.

We assume that $|\mathbf{C}| = |\mathbf{K}|$ and therefore

$$|\{e_k(x) : k \in \mathbf{K}\}| = |\mathbf{K}|$$

giving there do not exist two different keys $k_1, k_2 \in \mathbf{K}$ such that $e_{k_1}(x) = e_{k_2}(x) = y$. hence, for every x and y there is exactly one k such that $e_k(x) = y$.
Denote $n = |\mathbf{K}|$, let $\mathbf{P} = \{x_i | 1 \leq i \leq n\}$ and fix a ciphertext element y. We can name keys $k_1, k_2, \ldots, k_n$ in the way that $e_{k_i}(x_i) = y$. Using Bayes' theorem we have

$$\begin{aligned} P(X = x_i | Y = y) &= \frac{P(Y = y | X = x_i) P(X = x_i)}{P(Y = y)} \\ &= \frac{P(K = k_i) P(X = x_i)}{P(Y = y)}. \end{aligned} \tag{11}$$

Perfect Secrecy

Proof.

The perfect secrecy condition gives $P(X = x_i | Y = y) = P(X = x_i)$ and we have $P(K = k_i) = P(Y = y)$. This gives that all keys are used with the same probability. Since there are $|\mathbf{K}|$ keys, the probability is $1/|\mathbf{K}|$. Conversely, suppose the conditions are satisfied and we want to show perfect secrecy. The proof is analogous to the proof of perfect secrecy of the Shift cipher. □