VE475 Project1: Shannon's Theory and Cryptography

Yichen Hong,
Kaibin Wang,
Taoyue Xia, 518370910087
Wenjie Xianyu, 5183709101

June 2021

Abstract

Contents

1 Introduction **4**
1.1 Claude Shannon 4
1.2 Secrecy systems 4

2 Shannon's Theory **5**
2.1 Perfect Secrecy 5
2.2 Shannon's Theorem 5
2.3 Entropy 6
2.4 Equivocation 7
2.5 Redundancy 7
2.6 Unicity Distance 8

1 Introduction

1.1 Claude Shannon

Claude Elwood Shannon was an American mathematician, electrical engineer, and cryptographer known as "the father of information theory". Shannon is noted for having founded information theory with a landmark paper, "A Mathematical Theory of Communication", which he published in 1948. In the paper, he developed information entropy as a measure of the information content in a message, which is a measure of uncertainty reduced by the message.

In 1949, he published another notable paper "Communication Theory of Secrecy Systems", in which he developed mathematical theories of cryptography. "Perfect secrecy" is one of the most important theories.

1.2 Secrecy systems

In the paper "Communication Theory of Secrecy Systems", Shannon introduced three types of secrecy systems:

- **concealment systems**
- **privacy systems**
- **"true" secrecy systems**

Concealment systems contains methods such as using invisible ink, concealing messages in an innocent text, etc., which uses physical approaches to protect the real message from enemies.

Privacy systems contains methods such as speech inversion, which needs special equipment to recover the real message. It is a technical issue to attack the system.

"True" secrecy systems, which are the main concern of our project, uses ciphers and codes to conceal the real messages. As it is possible for enemies to intercept and record the transmitted ciphertext, we need to make the encryption algorithm as perfect as possible to prevent such occasion from happening. Below is the schematic of how this kind of systems works:

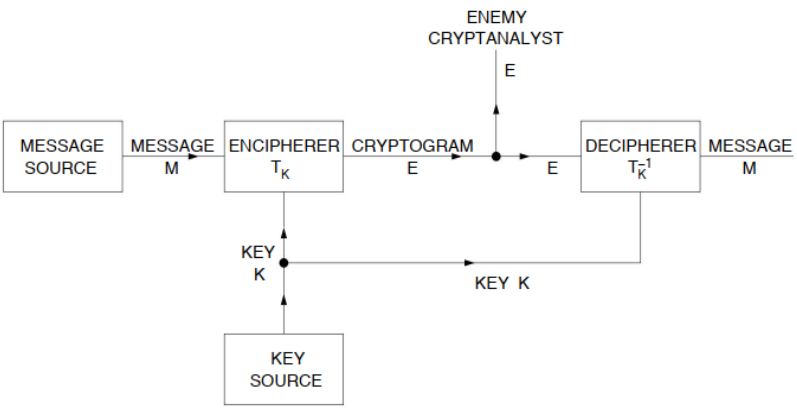

Figure 1: Schematic of a general secrecy system

2 Shannon's Theory

2.1 Perfect Secrecy

Suppose there is an adversary who is capable of learning the probability distribution of the message and the encryption scheme, as well as intercepting the ciphertext, but not the key used to encrypt the message, i.e., launching a ciphertext-only attack. Then, an encryption scheme is perfect secrecy when the adversary' s observing the ciphertext has no effect on his learning the plaintext.

The textbook Introduction to Modern Cryptography gives the definition of perfect secrecy as the following [4]:

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

$$Pr[M = m \mid C = c] = Pr[M = m] \tag{2.1}$$

The formula above has an equivalent form. That is, for any $m,\ m' \in M$, every $c \in C$:

$$Pr[Enc_K(m) = c] = Pr[Enc_K(m') = c] \tag{2.2}$$

The textbook also provides a lemma [4]:

> ***Lemma 1.*** An encryption scheme (***Gen*, *Enc*, *Dec***) with message space $\mathcal{M}$ is perfectly secret if and only if Eq. (2.2) holds for every $m,\ m' \in M$ and every $c \in C$.

In plain words, ***Lemma 1*** says that for any $m,\ m' \in M$, the probability distributions of the ciphertext are the same, thus implying nothing about the plaintext.

2.2 Shannon's Theorem

The textbook defines Shannon' s Theorem as [4]:

Let (***Gen*, *Enc*, *Dec***) be an encryption scheme with message space M, for which $|M| = |K| = |C|$. The scheme is perfectly secret if and only if:

1. Every key $k \in K$ is chosen with (equal) probability $1/|K|$ by ***Gen***.
2. For every $m \in M$ and every $c \in C$, there is a unique key $k \in K$ such that $Enc_k(m)$ outputs c.

Since it is an "if and only if" relation, we first start with the proof of the sufficiency of the conditions, i.e., if the two conditions are satisfied, then the encryption scheme is perfectly secret.

Given condition 2, we have:

$$Pr[C = c|M = m] = Pr[Enc_k(m) = c] = 1/|K|$$

Therefore, given the property |M| = |K| = |C|, by Bayes' Theorem,

$$\begin{aligned} Pr[M = m|C = c] &= \frac{Pr[C = c|M = m] \cdot Pr[M = m]}{Pr[C = c]} \\ &= \frac{\frac{1}{|K|} \cdot \frac{1}{|M|}}{\frac{1}{|C|}} \\ &= \frac{1}{|M|} = Pr[M = m] \end{aligned}$$

According to the definition of perfect secrecy and Eq. (2.1), the encryption scheme is perfectly secret.

Next we deal with the necessity of the conditions, i.e., given that the encryption scheme is perfectly secret, then the two conditions should be satisfied.

For any $c \in C$, there exists m such that $Pr[Enc_K(m) = c] \neq 0$. Then, ***Lemma 1*** gives that for any $m,\ m' \in M$,

$$Pr[Enc_K(m) = c] = Pr[Enc_K(m') = c] \neq 0$$

Let $K_i \in K$ denote the nonempty set for m_i that $Enc_{K_i}(m_i) = c$. When $m_i \neq m_j$, K_i and K_j must be disjoint, otherwise the same key for different messages will give the same ciphertext. And since $|K| = |M|$, for any $m_i \in M$ we have:

$$|K_i| = 1$$

and there stands condition 2, the existence and uniqueness of the key which gives $Enc_k(m) = c$. Again based on ***Lemma 1***, we shall have:

$$Pr[K = k_i] = Pr[Enc_K(m_i) = c] = Pr[Enc_K(m_j) = c] = Pr[K = k_j]$$

for any $m_i,\ m_j \in M$, making "every key $k \in K$ is chosen with equal probability $1/|K|$" , as in condition 1. Q.E.D.

2.3 Entropy

Traditionally, the term "*entropy*" is firstly used in thermodynamics, describing the degree of chaos in a system. The greater the entropy is, the less possible it is to predict the system's condition.

Similarly, in order to determine how many choices are involved in the selection of events, or in other words, how uncertain we are of the final output, Shannon defines "*entropy*" in information theory as [1]:

$$H(X) = -\sum_{i=1}^{n} P(x_i) \log(P(x_i))$$

Here, X is some discrete random variable with possible values $\{x_1, \cdots, x_n\}$, and $P(X)$ is the probability mass function of X.

In a secrecy system, we can use Shannon's Entropy to calculate the amount of information produced by a message, or the uncertainty to determine a specific message M:

$$H(M) = -\sum P(M) \log(P(M))$$

Also, there is an uncertainty of the choice of key K given by:

$$H(K) = -\sum P(K)\log(P(K))$$

In the perfect secrecy section above, the amount of information $H(M)$ contained in the message is at most $\log n$. Then the information can be completely concealed from enemies if the key uncertainty $H(K)$ is at least $\log n$.

2.4 Equivocation

Known as the conditional entropy, the equivocation about two random variables X and Y is given by:

$$H(X|Y) = -\sum_{x\in X, y\in Y} P(x,y)\log(P(x|y))$$

where $P(x,y)$ denotes the joint probability of X and Y.

In Shannon's theory, equivocation can be used as a theoretical secrecy index. In secrecy systems, there are two important equivocations of the message M and key K, denoted as $H(K|E)$ and $H(M|E)$, where E stands for the cryptogram [1]:

$$H(M|E) = -\sum_{M,E} P(M,E)\log(P(M|E))$$

$$H(K|E) = -\sum_{K,E} P(K,E)\log(P(K|E))$$

After some calculation, we can prove that $H(X,Y) \leq H(X) + H(Y)$. While the message and key are independent with each other, we can obtain:

$$H(M,K) = H(M) + H(K)$$

Furthermore,

$$H(M,K) = H(E,K) = H(E) + H(K|E)$$

The first equality is because the obtained information of M and K is equal to that of E and K. The second equality holds in perfect secrecy, as $P(K) = P(K|E)$.

Then we can obtain an equivocation formula of key from the two above equations,

$$H(K|E) = H(M) + H(K) - H(E)$$

Thus we can find out that if the enemy only knows the cryptogram, or ciphertext, the uncertainty of key is quite large as they can get many possible keys, but only one of them is correct.

2.5 Redundancy

To get a general knowledge of redundancy, we first introduce the concept of entropy rate, or the relativity between letters.

Firstly, the absolute rate is defined as R, where

$$R = \log|M|$$

Here $|M|$ denotes the scale of message space of a secrecy system.

Then we define the entropy of each symbol as r, to see the relativity between a symbol and the symbols already known as:

$$r = \lim_{n\to\infty} \frac{1}{n} H(X_1, \cdots, X_n) = \lim_{n\to\infty} H(X_n | X_1, \cdots, X_{n-1})$$

Then the relative redundancy D can be defined as:

$$D = \frac{R - r}{R} = 1 - \frac{r}{\log |M|}$$

When every symbol has little relativity with the previous symbol, which means every symbol in M is independent with each other, then we can get,

$$\begin{aligned} r &= \lim_{n\to\infty} \frac{1}{n} (- \sum_{m\in|M|^n} P(m) \log m) \\ &= \lim_{n\to\infty} \frac{1}{n} (- \sum_{m\in|M|^n} \frac{1}{|M|^n} \log \frac{1}{|M|^n}) \\ &= \frac{n \log |M|}{n} = \log |M| \\ \text{Then } D &= 1 - 1 = 0 \end{aligned}$$

Thus we can know that every symbol in a message can carry the amount of information of value r, and the redundancy shows the needless ratio of information. The smaller the redundancy of a message is, the better it is as it is harder to decrypt.

2.6 Unicity Distance

Unicity distance is the length of an original ciphertext needed to break the cipher by reducing the number of possible spurious keys to zero in a brute force attack [3]. The unicity distance U can be shown as:

$$U = \frac{H(K)}{D}$$

where U is the unicity distance, $H(K)$ is the key uncertainty, or the entropy of the key space, and D stands for the redundancy of plaintext in bits.

Therefore, the bigger the unicity distance the better, as it means that an enemy needs more amount of ciphertext to decrypt the system. If the unicity distance approaches infinity, the system is called an ideal secret system [1].

For example, for a one time pad with unlimited size, we have the entropy of key space unbounded as infinity, so $U \to \infty$, thus it is an ideal secret system. However, it is impossible to reach infinity, so the key size should be as large as possible.

References

[1] Shannon, Claude. “Communication Theory of Secrecy Systems”, *Bell System Technical Journal*, vol. 28(4), pp. 656–715, 1949.

[2] Shannon, Claude. ”A Mathematical Theory of Communication”, *Bell System Technical Journal*, vol. 27(3), pp. 379–423, 1948.

[3] Alfred, Menezes, Paul, Oorschot, Scott, Vanstone. “Chapter 7 - Block Ciphers”. *Handbook of Applied Cryptography*. p. 246.

[4] Katz, Jonathan, and Yehuda Lindell. “Introduction to Modern Cryptography”, 3rd ed., *CRC Press*, pp. 27-36, 2021.