Chapter 1

One-Time Pad Encryption and Perfect Secrecy

1.1 Introduction

The objective of this document is to give an informal overview of one-time pad encryption and its security properties.

1.2 Open-Time Pad encryption

How can Alice (A) transfer a message $M \in \{0,1\}^n$ securely to Bob (B) over a public channel? The public channel assumption implies that Eve (E) can monitor and read the channel between Alice and Bob. An encryption scheme has the property of *unconditional* security if the cipher text generated by the algorithm does not reveal any useful information to break the scheme even with access to unlimited computational power.

Unconditional security is sometimes referred to as *information-theoretic* security.

1.2.1 Encryption and Decryption

One-time pad encryption is a method of transmitting a message $M \in \{0,1\}^n$, an n bit string of binary numbers, from a user A to B. The main requirement is that the intruder E should not obtain any information about M by observing the cipher text. In this scheme, A chooses a random one time key $K \in \{0,1\}^n$, and XORs it component wise to the message M. The transformed message is then transmitted on the public insecure channel where E can monior the channel and access the communications. To decrypt the message, B should have a copy of the key K, which they need to have exchanged on a secure channel prior to the message transmission.

Let $\oplus$ denote the exclusive-or (XOR) symbol defined by

$$0 \oplus 0 = 1 \oplus 1 = 0; 0 \oplus 1 = 1 \oplus 0 = 1.$$

Example 1 *Suppose A wishes to send a message* $M = 0110111$*, and suppose they have*

previously established a shared secret key: $K = 1011011$. *The cipher text is formed by exclusive-oring the message with the key:* $C = M \oplus K = 1101100$. *Decryption is trivial as the XOR operation is associative. The message could be obtained by XORing* K *to the received cipher text* C.
$M = C \oplus K = 0110111$.

1.2.2 Perefect Security

What does it mean for an encryption scheme to be perfectly secure? We will look at the approach taken by Shannon in answering this question [2]. An encryption scheme has the property of *unconditional* security if the cipher text generated by the algorithm does not reveal sufficient information to break the scheme, even with access to an unlimited amount of computational power. In other words, the adversary cannot not obtain any knowledge to reverse the encryption by watching any amount of cipher text without access to the key.

Shannon in his seminal paper [2] in 1949 showed that one-time pad encryption is perfectly secure. We will first explore this idea in an informal way for the case when the length of the one-time pad is 1. Let a message space be formed by the symbols 0 and 1, i.e $M = \{0, 1\}$. Assume that an adversary knows that Alice will pick 0 with two out of three times and 1 one out of three times. This formally translates to *a priori* probabilities for the message, i.e $P(M = 0) = 2/3$ and $P(M = 1) = 1/3$. A key is a binary bit chosen completely randomly, i.e $P(K = 0) = P(K = 1) = 1/2$. The output of the cipher is the XOR function of the message and the key. In the above situation after observing the ciphertext C, the adversary cannot learn anything about the possible message other than the *a priori* knowledge of the input message distribution. Using the one-time pad, it is possible to prescisely formulate this idea through application of probability theory.

Suppose $C = M \oplus K$ was observes at the output of the cipher. With this condition, we shall estimate the probability that the actual message. If this probability is equal to the *a priori* probability of message $M = 0$ or $M = 1$, then the adversary will fail to obtain any new information about the actual message.

The tree in Figure 1.1 depicts the probabilities of various combinations of message, cipher texts and keys.

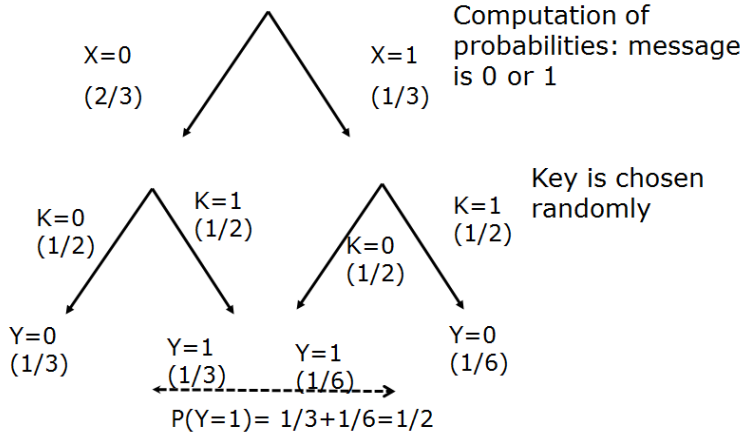

FIGURE 1.1: Probability Tree

For example, the left most edge of the tree is followed to encrypt $M = 0$ using $K = 0$

leading to ciphertext $C = 0$. Suppose the adversary observes $C = 1$. Let us consider the conditional probability of $M = 0$, with the output $C = 1$. From first principles, this is given by

$$P(M = 0|C = 1) = \frac{P(M = 0 \wedge C = 1)}{P(C = 1)} = \frac{(2/3)(1/2)}{(1/2)} = (2/3)$$

This value is same as the *a priori* probability that $M = 0$ and hence the adversary will not obtain any more information by observing the ciphertext than what he had already known. Similarly you find other probabilities:

$$P(M = 1|C = 1) = \frac{P(M = 1 \wedge C = 1)}{P(C = 1)} = \frac{(1/3)(1/2)}{(1/2)} = 1/3 = P(M = 1).$$

$$P(M = 1|C = 0) = \frac{P(M = 1 \wedge C = 1)}{P(C = 1)} = \frac{(1/3)(1/2)}{(1/2)} = 1/3 = P(M = 1).$$

$$P(M = 0|C = 0) = \frac{P(M = 0 \wedge C = 1)}{P(C = 1)} = \frac{(2/3)(1/2)}{(1/2)} = 2/3 = P(M = 0).$$

Hence from an adversary's point of view, no new knowledge about M will be gained by seeing C other than what is already known by virtue of the *a priori* distribution of M.

This is equivalent to saying that seeing the cipher text does not increase the adversary's knowledge about the underlying message. The idea holds true even for messages of aribitrary size. We have the following theorem:

Theorem 1 ([1]) *Let $M, K \in \{0,1\}^n$. Let K be a randomly chosen uniformly from $\{0,1\}^n$ and indepedent of M. Then the conditional probability of $M = x$ given that the ciphertext $C = y$ is the same as* a priori *probability that $M = x$. In otherwords, we have:*

$$P(M = x|C = y) = P(M = x). \tag{1.1}$$

Proof 1 *Consider the joint probability function $P(M = x \wedge\ C = y)$. From the property of the XOR function, we can write*

$$P(M = x \wedge C = y) = P(M = x \wedge K = (x \oplus y)).$$

As K is independent and chosen randomly from $\{0,1\}^n$ we can write

$$P(M = x \wedge C = y) = P(M = x)P(K = (x \oplus y)) = P(M = x)(1/2^n).$$

To compute $P(C = y)$, we need to add the probabilities of all the events that lead to $C = y$. So we have:

$$P(C = y) = \sum_x P(M = x \wedge C = y))$$

$$= \sum_x P(M = x)(1/2^n) = (1/2^n)\sum_x P(M = x) = (1/2^n).$$

This implies that each C is equally likely and independent of the message. Hence,

$$P(M = x|C = y) = \frac{P(M = x \wedge C = y)}{P(C = y)}$$

$$= \frac{P(M = x)2^{-n}}{2^{-n}} = P(M = x),$$

as required. Note C is equally likely and independent of the message which follows because K is a unformly distributed random variable.

Note that the above theorem depends critically on the fact that the key K is generated according to the uniform distribution and independent of M. Under these conditions, an adversary will not gain any new information about the message by watching the ciphertext. In otherwords the cipher text distribution is unformly random irrespective of any plaintext distribution.

In practice, messages may be biased which could be observed by adversaries. An encryption transformation should distribute messages to the cipher space fairly uniformly irrespective of known apriory statistics of the messages. Analysis of one-time pad encryption tells us that if we choose a uniformly distributed random secret key pad of at least the size of the message, perfect secrecy can be achieved. Hence One-time pad encryption is not practical.

1.2.3 Two time pad is bad

One-time pad is not practical. It demands a key is chosen independently to message and the key should be as long as the message which needs to exchanged in a secure manner between participants. If a particular one-time pad is used twice, it leaks the statistics of the plain text. The Germans made this mistake during World war II allowing the Alies to recover keys. Let

$$C_1 = M_1 \oplus K; C_2 = M_2 \oplus K; then$$

$$C_1 \oplus C_2 = M_1 \oplus M_2 \oplus K \oplus K = M_1 \oplus M_2,$$

which leaks input statistics through $M_1 \oplus M_2$. This means that you need a new key for every message. This idea is used in attacking the Vegenere cipher (same key-pad is added many times).

1.2.4 Vernam Cipher

The Vernam cipher is exactly the one-time pad described above except for two main differences: the message and cipher space are sequences of English characters (lower case or upper case), represented as sequences of integers modulo 26 and the $\oplus$ operation is replaced by modulo 26.

Let $M[i], K[i] \in \{0, 1, \cdots, 25\}, 0 \leq i < n$, then

```
for i:=0 to n-1 do
C[i] = M[i] + K[i] \bmod 26.
end for;
```

Similarly the decryption of C with the key K is given by

```
for i:=0 to n-1 do
M'[i] = C[i] - K[i] \bmod 26.
end for;
```

From the additive properties of integers it is clear that encryption is the inverse of decryption.

Bibliography

[1] Ronald Rivest. 6.857 Course Notes. 2002.

[2] C.E. Shannon. Communication in presence of noise. *IEEE*, 37:10–21, 1949.