Overview of One-Time Pads

Justin Wyss-Gallifent

March 30, 2022

0.1 Introduction

A one-time pad is a key whose length is at least as long as the message being sent. There are several ways for this to manifest in encryption but here are two:

- With a message of 1s and 0s using XOR to encrypt and decrypt.
- With a message of `A` through `Z` using addition and subtraction mod 26 to encrypt and decrypt.

The basic notion here is that since the key is at least as long as the message a given ciphertext could be literally anything.

Example:

If the message is `PEACE` and the key is `EFRGH` then the ciphertext is `TJRIL`.

If the message is `FIGHT` and the key is `OBLBS` then the ciphertext is `TJRIL`.

The point here is that if Eve intercepts `TJRIL` she has no way of knowing what the message is.

0.2 Issues

There are several things of importance with one-time pads:

1. The key must be shared ahead of time and since the key is at least as long as the message, this is a bit expensive.
2. The key should only be used once. If Eve manages to decrypt a single ciphertext to obtain the corresponding message then she has the entire key. Moreover, it's possible with some analysis to figure out several messages when several messages are given and it's known that the key was shared.
3. The key should be as "random" as possible, meaning if there are N possible keys then the chance of using one key should be $1/N$. If this is not the case then the choice of key will influence the ciphertext. In the example above if for some reason having a `B` in the key is highly unlikely then the second option is less likely than the first option and so the ciphertext has given us some information about the plaintext.

0.3 Two Theorems

There are two theorems which tie one-time pads to perfect secrecy.

0.3.1 More Keys than Messages

Theorem:

If a cryptosystem has perfect secrecy then the number of possible keys must be at least as large as the number of possible messages.

Proof:

Recall that a cryptosystem has perfect secrecy iff for all messages m and all ciphertexts c we have:

$$P(M = m \,|\, C = c) = P(M = m)$$

Let K be the set of possible keys and M be the set of possible messages. Suppose $|M| > |K|$. Let c be a ciphertext. For all $k \in K$ let $D_k(c)$ be the result of applying decryption with key k to c. Each of these is a message but since there are more possible messages than possible keys we must miss some message m_0. Thus we have $P(M = m_0 \,|\, C = c) = 0$. However we also know $P(M = m_0) \neq 0$ since there's a chance of choosing that message. Thus since these are not equal, we don't have perfect secrecy.

0.3.2 The One-Time Pad has Pefect Secrecy

Probability Lemma 1:

If an event A equals the union of disjoint events A_1 and A_2 then $P(A) = P(A_1) + P(A_2)$ and this expands to an arbitrary finite number of disjoint events.

Probability Lemma 2:

The conditional probability $P(A|B)$ satisfies:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Theorem:

The one-time pad does yield perfect secrecy.

Proof:

We claim that for all messages m and all ciphertexts c we have:

$$P(M = m \,|\, C = c) = P(M = m)$$

Let's look at the bitwise case because it's easier to write down.

Suppose there are N possible keys each with probability $1/N$. For a fixed m_0 we have, for any ciphertext c and key k:

$$c \text{ is the corresponding ciphertext iff } c = k \oplus m_0$$

It follows that given the message m_0 the probability of c being the ciphertext equals the probability of k being the key, which is $1/N$. Thus each ciphertext has equal probability $1/N$ as well and so:

$$P(C = c \,|\, M = m_0) = \frac{1}{N}$$

Since this is true for all m_0 we can say that for all m and for all c we have:

$$P(C = c \,|\, M = m) = \frac{1}{N}$$

Now then, for a given ciphertext c first note that the event $C = c$ is made up of the disjoint events $C = c \cap M = m$ taken over all possible m. In other words:

$$\text{Event}(C = c) = \text{Event}(C = c \cap M = m_1) \,\dot{\cup}\, \text{Event}(C = c \cap M = m_2) \,\dot{\cup}\, ...$$

It follow that:

$$\begin{aligned}
P(C = c) &= \sum_i P(C = c \cap M = m_i) \\
&= \sum_i P(C = c \,|\, M = m) P(M = m_i) \\
&= \sum_i \left(\frac{1}{N}\right) P(M = m_i) \\
&= \frac{1}{N} \sum_i P(M = m_i) \\
&= \frac{1}{N}(1) = \frac{1}{N}
\end{aligned}$$

From here we have:

$$
\begin{aligned}
P(M = m \,|\, C = c) &= \frac{P(M = m \cap C = c)}{P(C = c)} \\
&= \frac{P(C = c \cap M = m)}{P(C = c)} \\
&= \frac{P(C = c \,|\, M = m)P(M = m)}{P(C = c)} \\
&= \frac{(1/N)P(M = m)}{1/N} \\
&= P(M = m)
\end{aligned}
$$