Notes on Probabilistic Cipher Model

Probabilistic Cipher Model

2 idependent random variables: X with ranges R_X (the set of plaintexts), and Z with range R_Z (the key space).

No assumptions are made w.r.t. the probability distribution of X. It is assumed that Z is uniformly distributed.

The encryption function $E_z(x) = y$. The set of ciphertexts is the image of $X \times Z$ under $\sim$:

$$Y = (X \times Z)\big/\sim : (a,b) \sim (c,d) \iff E_b(a) = E_c(d) \ .$$

We will consider the factor space XZ with range R_{XZ}.

Information-Theoretic Security of a 1-bit XOR Cipher

$$\begin{aligned}
\Pr_{XZ}[Y=y] &= \sum_{x,z} \Pr_{XZ}[X=x, Z=z][x \oplus z = y] = \sum_{x,z} \Pr_{XZ}[X=x] \Pr_{XZ}[Z=z][x \oplus z = y] \\
&= \sum_x \Pr_{XZ}[X=x] \sum_z \underbrace{\Pr_{XZ}[Z=z]}_{\text{constant}}[x \oplus z = y] = \Pr_{XZ}[Z=z] \underbrace{\sum_x \Pr_{XZ}[X=x]}_{=1} \underbrace{\sum_z [x \oplus z = y]}_{=1} \\
&= \Pr_{XZ}[Z=z] = \frac{1}{2} \ .
\end{aligned}$$

This means that Y is uniformly distributed. The 1-bit XOR cipher is information–theoretically secure, if the ciphertexts are independent of the plaintexts: $\Pr_{XZ}[X=x|Y=y] = \Pr_{XZ}[X=x]$. This means that the ciphertext contains no information about the plaintext, and cannot leak it. Since

$$\Pr_{XZ}[X=x|Y=y] = \frac{\Pr_{XZ}[X=x, Y=y]}{\Pr_{XZ}[Y=y]} \ ,$$

and $\Pr_{XZ}[Y=y]$ is known, we need to calculate the joint probability $\Pr_{XZ}[X=x, Y=y]$.

$$\begin{aligned}
\Pr_{XZ}[X=x, Y=y] &= \sum_z \Pr_{XZ}[X=x, Z=z][x \oplus z = y] = \sum_z \Pr_{XZ}[X=x] \Pr_{XZ}[Z=z][x \oplus z = y] \\
&= \Pr_{XZ}[X=x] \sum_z \underbrace{\Pr_{XZ}[Z=z]}_{\text{constant}}[x \oplus z = y] = \Pr_{XZ}[X=x] \underbrace{\Pr_{XZ}[Z=z]}_{=\frac{1}{2}} \underbrace{\sum_z [x \oplus z = y]}_{=1} \\
&= \frac{1}{2} \Pr_{XZ}[X=x] \ .
\end{aligned}$$

The conditional probability of a plaintext given a ciphertext is

$$\Pr_{XZ}[X=x|Y=y] = \frac{\Pr_{XZ}[X=x, Y=y]}{\Pr_{XZ}[Y=y]} = \frac{\frac{1}{2}\Pr_{XZ}[X=x]}{\frac{1}{2}} = \Pr_{XZ}[X=x] \ ,$$

and therefore the 1-bit XOR cipher is information–theoretically secure.

Theorem 1. If Z is independent of X, Z is uniformly distributed and for every plaintext x and for every ciphertext y there is a unique key z such that $E_z(x) = y$, then the cipher is unbreakable.

Proof. See lecture slides. $\square$

It is possible to show that in a 1-bit XOR cipher for every pair of plaintext x and ciphertext y there is a unique key z such that $x \oplus z = y$.

Xor operation is idential to addition in $\mathbb{Z}_2$. Therefore, we can encode the encryption function as

$$y = x \oplus z \Longleftrightarrow y = x + z \pmod 2 \ .$$

For every plaintext–ciphertext pair (x, y) there is a key $z = y - x \pmod 2$ such that $x + z = x + y - x = y$. To show that such z is unique, assume there exists another key $z' \neq z$ such that $x + z' = y \pmod 2$. We have a system of two equations

$$x + z = y \pmod 2 \ , x + z' = y \pmod 2 \ .$$

Subtracting the second equation from the first one, we get $z - z' = 0 \pmod 2$ which means that $z \equiv z' \pmod 2$.

Since for every plaintext–ciphertext pair (x, y) there exists a unique key z, by Theorem 1 the cipher is information–theoretically secure.

Information–Theoretic Security of a Shift Cipher

The encryption function of a shift cipher is $y = x + z \pmod{26}$. It is possible to show that for every plaintext–ciphertext pair (x, y) there exists a unique key $z = y - x \pmod{26}$. Hence, by Theorem 1 the shift cipher is information–theoretically secure.

Inforation–Theoretic Security of a Substitution Cipher

A substitution cipher is a cipher, where the key is a mapping which puts every plaintext into one-to-one correspondence with a unique ciphertext. This is the size of the key space. The encryption function $E_z(x) = y$ is the mapped value assigned by the key for the specific plaintext $y = z(x)$. Since the key is a bijection, it is invertible, and ciphertext can be decrypted into plaintext.

Since there are 26 letters in English alphabet, there are 26! possible mappings. If we fix one specific plaintext–ciphertext pair (x_i, y_i), then there exist 25! permutations of the remaining letters. In other words, for every plaintext–ciphertext pair (x, y) there exist 25! unique keys z such that $y = z(x)$. Therefore, we cannot prove information–theoretical security using Theorem 1 above.

Instead, we will show that $\Pr\limits_{XZ}[X = x|Y = y] = \Pr\limits_{XZ}[X = x]$.

$$\begin{aligned}
\Pr_{XZ}[Y = y] &= \sum_{x,z} \Pr_{XZ}[X = x, Z = z][z : x \mapsto y] = \sum_{x,z} \Pr_{XZ}[X = x] \Pr_{XZ}[Z = z][z : x \mapsto y] \\
&= \sum_{x} \Pr_{XZ}[X = x] \sum_{z} \underbrace{\Pr_{XZ}[Z = z]}_{\text{constant}}[z : x \mapsto y] = \underbrace{\Pr_{XZ}[Z = z]}_{=\frac{1}{26!}} \underbrace{\sum_{x} \Pr_{XZ}[X = x]}_{=1} \underbrace{\sum_{z}[z : x \mapsto y]}_{=25!} \\
&= \frac{25!}{26!} = \frac{1}{26} \ .
\end{aligned}$$

This tells us that Y is uniformly distributed.

$$\begin{aligned}
\Pr_{XZ}[X=x,Y=y] &= \sum_z \Pr_{XZ}[X=x,Z=z][z:x\mapsto y] = \sum_z \Pr_{XZ}[X=x]\Pr_{XZ}[Z=z][z:x\mapsto y] \\
&= \Pr_{XZ}[X=x]\sum_z \underbrace{\Pr_{XZ}[Z=z]}_{\text{constant}}[z:x\mapsto y] = \Pr_{XZ}[X=x]\underbrace{\Pr_{XZ}[Z=z]}_{=\frac{1}{26!}}\underbrace{\sum_z [z:x\mapsto y]}_{=25!} \\
&= \frac{1}{26}\Pr_{XZ}[X=x] \ .
\end{aligned}$$

$$\Pr_{XZ}[X=x|Y=y] = \frac{\Pr_{XZ}[X=x,Y=y]}{\Pr_{XZ}[Y=y]} = \frac{\frac{1}{26}\Pr_{XZ}[X=x]}{\frac{1}{26}} = \Pr_{XZ}[X=x] \ .$$

Therefore, the substitution cipher is information–theoretically secure.