Perfect Secrecy and Ideal Cipher

Mario Cagalj

University of Split

(Symmetric) Vernam cipher

- Plaintext space: $\mathcal{P} = \{0,1\}^n$, plaintext $P \in \mathcal{P}$
- Key space: $\mathcal{K} = \{0,1\}^n$, key $K \in \mathcal{K}$
- Ciphertext space: $\mathcal{C} = \{0,1\}^n$, ciphertext $C \in \mathcal{C}$
- Encryption algorithm: $C = P \oplus K$
- Decryption algorithm: $P = C \oplus K$

Note, P, K, C are all *random variables*. Also, the decryption function is an *inverse* function of the encryption function:

$$P = C \oplus K = P \oplus K \oplus K = P$$

Vernam cipher - example

```
      Plaintext P: Test

  Plaintext P (hex): 54657374
        Key K (hex): 00010203
 Ciphertext C (hex): 54647177

  Plaintext P (bin): 01010100 01100101 01110011 01110100
        Key K (bin): 00000000 00000001 00000010 00000011
 Ciphertext C (bin): 01010100 01100100 01110001 01110111
```

Vernam cipher - key reuse

By reusing the same encryption key over and over renders the Vernam cipher completely insecure. Indeed:

$$C_1 = P_1 \oplus K$$
$$C_2 = P_2 \oplus K$$

Then, by *xoring* two **public ciphertexts** C_1 and C_2 we have: ,

$$C_1 \oplus C_2 = P_1 \oplus K \oplus P_2 \oplus K = P_1 \oplus P_2$$

Observe:

- $C_1 \oplus C_2 = \mathbf{0}$ implies $P_1 = P_2$
- $P_2 = C_1 \oplus C_2 \oplus P_1$, where P_1 might be easily guessable
- Finally, $H(P_2) = H(P_1)$ (equal entropies)

Vernam cipher - key reuse

```
                Key K: abf1021df4

         Plaintext P1: Hello
   Plaintext P1 (hex): 48656c6c6f
  Ciphertext C1 (hex): e3946e719b

         Plaintext P2: world
   Plaintext P2 (hex): 776f726c64
  Ciphertext C2 (hex): dc9e707190

            P1 xor P2: 3f0a1e000b
            C1 xor C2: 3f0a1e000b

P1 xor C1 xor C2 (hex): 776f726c64
P1 xor C1 xor C2 (utf): world
```

Perfect secrecy

We show how to convert Vernam cipher into an *ideal cipher*.

Definition (Perfect secrecy)

A cipher is said to have a *perfect secrecy* property if:

$$Pr(P = p|C = c) = Pr(P = p), \quad \forall p \in \mathcal{P}, c \in \mathcal{C}.$$

Here, $Pr(P|C)$ denotes conditional *posterior probability*, and $Pr(P)$ *prior probability* of a plaintext.

Intuitively, a given cipher perfectly protects message confidentiality if resulting ciphertexts, once captured by the attacker, do not help him/her to gain additional insights into encrypted plaintexts.

Perfect secrecy

Let us restate the previous definition using the language of information entropy.

Definition (Perfect secrecy)

A cipher is said to have a *perfect secrecy* property if:

$$H(P|C) = H(P), \quad P \in \mathcal{P}, C \in \mathcal{C}.$$

Here, $H(P|C)$ is conditional entropy of a plaintext given ciphertext.

Intuitively, a given cipher perfectly protects message confidentiality if resulting ciphertexts, once captured by the attacker, do not decrease attacker's *apriori* uncertainty about encrypted plaintexts.

One-time pad

Definition (One-time pad)

One-time-pad is identical to Vernam cipher with the important difference that the encryption key K is selected *anew and uniformly at random* for each new plaintext to be encrypted. Thus, $Pr(K = k) = \frac{1}{|\mathcal{K}|}, \forall k \in \mathcal{K}$.

Theorem

One-time-pad has a perfect secrecy property.

Observe:

- While ideal, this cipher is not practical as it implies that the number of keys is equal to the number of plaintext messages
- It still can happen that two plaintexts are by chance encrypted using the same key, but the attacker does not know which ones

One-time-pad example

Let $P, K, C \in \{0, 1\}$:

Consider two plaintext messages: $P_1 = 1$ and $P_2 = 1$. Using *one-time-pad*, we encrypt them as follows:

- Generate randomly K_1, then calculate $C_1 = P_1 \oplus K_1$
- Generate randomly K_2, then calculate $C_2 = P_2 \oplus K_2$

After *xor*-ing two public ciphertexts C_1 and C_2 and rearranging:

$$P_2 = C_1 \oplus C_2 \oplus P_1 \oplus K_1 \oplus K_2$$

Now, even if the attacker knows P_1, he still cannot calculate P_2 since he cannot predict $K_1 \oplus K_2$ better than randomly guessing. Hence, C_1 and C_2 are useless to the attacker, i.e., $H(P|C) = H(P)$.