Perfect secrecy

An introduction to information theory

CS349 Cryptography
Department of Computer Science
Wellesley College

The communication channel
Cryptosystems

A cryptosystem is a five-tuple \((P, C, K, E, D)\) where the following conditions are satisfied:

1. \(P\) is a finite set of possible plaintexts
2. \(C\) is a finite set of possible ciphertexts
3. \(K\), the keyspace, is a finite set of possible keys
4. For each \(K \in K\), there is an encryption rule \(e_K \in E\) and a corresponding decryption rule \(d_K \in D\). Each \(e_K : P \rightarrow C\) and \(d_K : C \rightarrow P\) are functions such that \(d_K(e_K(x)) = x\) for every element \(x \in P\).

Perfect secrecy

Caesar revisited

- The shift cipher cryptosystem is given by \(P = C = K = \mathbb{Z}_{26}\). For \(0 \leq K \leq 15\), define \(e_K(x) = (x + K) \mod 26\) and \(d_K(y) = (y - K) \mod 26\).
- English text is encrypted by setting up a correspondence between alphabetic characters and residues modulo 26 (A \(\leftrightarrow\) 0, B \(\leftrightarrow\) 1, C \(\leftrightarrow\) 2, ..., Z \(\leftrightarrow\) 25).

<table>
<thead>
<tr>
<th>Plain:</th>
<th>meet at midnight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residue:</td>
<td>12 4 4 19 0 19 12 8 3 13 8 6 7 19</td>
</tr>
<tr>
<td>Shift:</td>
<td>23 15 15 4 11 4 23 19 14 24 19 17 18 4</td>
</tr>
<tr>
<td>Cipher:</td>
<td>X P P E L E X T O Y T R S E</td>
</tr>
</tbody>
</table>

Perfect secrecy
Similarly, we define Vigenère

Let $m$ be a positive integer. Define $P = C = K = (Z_{26})^m$.

For a key $K = (k_1, k_2, \ldots, k_m)$, define

$$e_k(x_1, x_2, \ldots, x_m) = (x_1 + k_1, x_2 + k_2, \ldots, x_m + k_m)$$

and

$$d_k(y_1, y_2, \ldots, y_m) = (y_1 - k_1, y_2 - k_2, \ldots, y_m - k_m)$$

where all operations are performed in $Z_{26}$.

Key: K I N G K I N G K I N G K I N
Residue: 10 8 13 6 10 8 13 6 10 8 13
Plain: t h e m a n i n t h e m o o n
Residue: 19 7 4 12 0 13 8 13 19 7 4 12 14 14 13
Shift: 4 15 17 18 10 21 21 19 4 15 17 18 24 22 1
Cipher: E P R S K V V T E P R S Y W B

How secure is a cryptology?

- His paper addresses the $64,000 question.
Random variables

- A discrete random variable, $X$, consists of a finite set $X$ and a probability distribution defined on $X$. The probability that the random variable $X$ takes on the value $x$ is denoted $\Pr[X = x]$.*
- It must be the case that $0 \leq \Pr[x]$ for all $x$ in $X$, and $\sum_{x \in X} \Pr[x] = 1$.

*We abbreviate this to $\Pr[x]$ if the random variable $X$ is fixed.

Events

- Suppose we have a random variable $X$ defined on $X$, and $E \subseteq X$. The probability that $X$ takes on a value in the subset $E$ is computed to be
  $$\Pr[x \in E] = \sum_{x \in E} \Pr[x]$$
- The subset $E$ is called an event.
**Conditional probability**

- Suppose $X$ and $Y$ are random variables defined on finite sets $X$ and $Y$, respectively. The **joint probability** $Pr(x, y)$ is the probability that $X$ takes on the value $x$ and $Y$ takes on the value $y$.

- The **conditional probability** $Pr[x|y]$ denotes the probability that $X$ takes on the value $x$ and given that $Y$ takes on the value $y$.

**Independent events**

- Random variables $X$ and $Y$ are said to be **independent** if $Pr(x, y) = Pr[x]Pr[y]$ for all $x \in X$ and $y \in Y$.

- We would expect the random variables $X$ and $Y$ representing the values of the first and second dice respectively to be independent.
Bayes’ Theorem

If \( \Pr[y] > 0 \), then
\[
\Pr[x | y] = \frac{\Pr[x] \Pr[y | x]}{\Pr[y]}
\]

Corollary

\( X \) and \( Y \) are independent random variables if and only if \( \Pr[x | y] = \Pr[x] \) for all \( x \in X \) and \( y \in Y \).

Start with cryptosystem \((P, C, K, E, D)\)

- Keys are chosen from \( K \) using a fixed probability distribution, defining a random variable \( K \). Denote the probability that key \( K \) is chosen by \( \Pr[K=K] \).
- Similarly there is probability distribution on the plaintext space \( P \). Each plaintext element defines a random variable \( x \). The probability that plaintext \( x \) occurs is denoted by \( \Pr[x=x] \).
**Chicken scratch**

- The probability distributions on $\mathcal{P}$ and $\mathcal{K}$ induce a probability distribution on $\mathcal{C}$.
- For $K \in \mathcal{K}$, define $C(K) = \{e_K(x) : x \in \mathcal{P}\}$, the set of possible ciphertext for a given key.

**Determining the probability of plaintext $x$ given ciphertext $y$**

- For each $y \in C$,
  \[
  \Pr[y = y] = \sum_{(K:y \in C(K))} \Pr[K = K] \Pr[x = d_K(y)]
  \]
- For each and $y \in C$ and $x \in \mathcal{P}$,
  \[
  \Pr[y = y | x = x] = \sum_{(K:x=d_K(y))} \Pr[K = K]
  \]
- Using Bayes' Theorem
  \[
  \Pr[x = x | y = y] = \frac{\sum_{(K:y \in C(K))} \Pr[K = K] \Pr[x = d_K(y)]}{\sum_{(K:y \in C(K))} \Pr[K = K] \Pr[x = d_K(y)]}
  \]
**Toy example**

- Let $\mathcal{P} = \{a, b\}$ with $\Pr[a] = 1/4$, $\Pr[b] = 3/4$.
- Let $\mathcal{K} = \{K_1, K_2, K_3\}$ with $\Pr[K_1] = 1/2$, and $\Pr[K_2] = \Pr[K_3] = 1/4$.
- Let $\mathcal{C} = \{1, 2, 3, 4\}$ with encryption functions given by the following encryption matrix:

<table>
<thead>
<tr>
<th>K</th>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_1$</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$K_2$</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$K_3$</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

**Perfect secrecy**

- A cryptosystem has perfect secrecy if $\Pr[x | y] = \Pr[x]$ for all $x \in \mathcal{D}$, $y \in \mathcal{C}$.
- That is, the a posteriori probability that the plaintext is $x$, given that the ciphertext $y$ is observed, is identical to the a priori probability that the plaintext is $x$. 
The Holy Grail of cryptography

- Suppose the 26 keys in the Shift Cipher are used with equal probability 1/26. Then for any plaintext probability distribution, the Shift Cipher has perfect secrecy.
- Hence the Shift Cipher is "unbreakable" provided that a new random key is used to encrypt every plaintext character.

Characterization of perfect secrecy

Suppose $(P, C, K, E, D)$ is a cryptosystem where $|K| = |C| = |P|$. Then the cryptosystem provides perfect secrecy if and only if every key is used with equal probability $1/|K|$ and for every $x \in P$ and every $y \in C$, there is a unique key such that $e_k(x) = y$. 
One-time pad

Let $n \geq 1$ be an integer. Take $P = C = K = (\mathbb{Z}_2)^n$. For $K \in (\mathbb{Z}_2)^n$, define $e_K(x)$ to be the vector sum module 2 of $K$ and $x$. So, if $x = (x_1, x_2, \ldots, x_m)$ and $K = (k_1, k_2, \ldots, k_m)$, then $e_K(x) = (x_1 + k_1, x_2 + k_2, \ldots, x_m + k_m)$ and $d_K(y) = (y_1 - k_1, y_2 - k_2, \ldots, y_m - k_m)$.

*Equivalently, take the XOR of the two bit strings.

Perfect secrecy 5-19

Why use anything else

- Well, for one thing key has to be as long as the plaintext. Key distribution and management is a big problem.
- Reuse is tempting, but a cardinal sin.*
- Obtaining a truly random sequence is expensive.

*As the Russians learned to their dismay.