Lecture 3
Symmetric encryption and perfect secrecy

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License.
Lecture 3— Contents

General model of an encryption system

The guessing attack
  Success probability
  Sequential guessing

Perfect secrecy
  Definition
  One-time pad
  Necessary condition
Security goals, threats, services and mechanisms

Goals
- Confidentiality
- Integrity
- Availability
- Accountability
- Privacy

Threats
- Eavesdropping
- Forgery
- Masquerade
- Repudiation
- Denial of service
- Jamming

Services
- Secrecy
- Integrity protection
- Authentication
- Notarization
- Access control
- Jamming rejection

Mechanisms
- Encryption
- Key agreement
- Digital signature
- Authentication codes
- Intrusion detection
- Spreading

Symmetric encryption and perfect secrecy

Nicola Laurenti

October 2, 2017

3 / 17
General model of an encryption system

A \begin{array}{l}
\text{sender} \\
\end{array} \xrightarrow{u} \begin{array}{l}
E(k; u) \\
\text{encryption} \\
\end{array} \xrightarrow{k} \begin{array}{l}
\text{key} \\
\end{array} \xrightarrow{k} \begin{array}{l}
D(k; x) \\
\text{decryption} \\
\end{array} \xrightarrow{u} \begin{array}{l}
\text{secret message} \\
\text{receiver} \\
\end{array}

E \begin{array}{l}
\text{eavesdropper} \\
\end{array}
Glossary and notation

secret message (plaintext)  $u \in \mathcal{M}$  message space

transmitted message (ciphertext)  $x \in \mathcal{X}$  cipher space

encryption key  $k \in \mathcal{K}$  key space

encryption map  $E : \mathcal{K} \times \mathcal{M} \mapsto \mathcal{X}$

$E_k : \mathcal{M} \mapsto \mathcal{X} \quad E_k(u) \doteq E(k, u)$

decryption map  $D : \mathcal{K} \times \mathcal{X} \mapsto \mathcal{M}$

$D_k : \mathcal{X} \mapsto \mathcal{M} \quad D_k(x) \doteq D(k, x)$

Key and plaintext are random with probability mass distribution

key pmd  $p_k : \mathcal{K} \mapsto [0, 1]$  typically uniform:  $k \sim \mathcal{U}(\mathcal{K})$

plaintext pmd  $p_u : \mathcal{M} \mapsto [0, 1]$  not necessarily uniform

The encryption system is completely specified as:

$S = (\mathcal{M}, \mathcal{X}, \mathcal{K}, E, D, p_u, p_k)$
General assumptions

- **(perfect reliability)** The receiver must be able to recover the secret message perfectly
  \[ D_k = E_k^{-1} \quad \forall k \in \mathcal{K} \]

- **(Kerchoff’s assumption)** The eavesdropper knows the system \( S \) (in particular the maps \( E(\cdot, \cdot) \) and \( D(\cdot, \cdot) \))

Where does secrecy come from?

Secrecy is only based on the fact that the eavesdropper does not know the actual realization of \( k \) and hence the particular \( E_k(\cdot) \), \( D_k(\cdot) \) used
Example (Substitution)

\[
\mathcal{M} = \bigcup_{\ell \in \mathcal{L}} \mathcal{A}_u^\ell = \mathcal{A}_u^* \quad , \quad \mathcal{X} = \mathcal{A}_x^*
\]

\[E_k : [x_1, \ldots, x_\ell] = [e_k(u_1), \ldots, e_k(u_\ell)]\]

with \( e_k : \mathcal{A}_u \mapsto \mathcal{A}_x \) injective

Provided \( |\mathcal{A}_u| \leq |\mathcal{A}_x| \), the class \( \{e_k\} \) of all injective functions \( \mathcal{A}_u \mapsto \mathcal{A}_x \) has cardinality

\[K = \prod_{i=1}^{\mathcal{A}_u} (\mathcal{A}_x - i + 1) = \frac{|\mathcal{A}_x|!}{(|\mathcal{A}_x| - |\mathcal{A}_u|)!}\]

We can take \( \mathcal{K} = \{1, \ldots, K\} \).

If \( |\mathcal{A}_u| = |\mathcal{A}_x| = M \), we have \( K = M! \).
The guessing attack

In this attack, E wants to learn the value of $u$, and attempts a guess $\hat{u} \in \mathcal{M}$

**Ignorant guess**

By ignoring the reading of $x$, the optimal guess for E is

$$\hat{u} = \arg \max_{a \in \mathcal{M}} p_u(a)$$

and the corresponding success probability is

$$P[\hat{u} = u] = p_u(\hat{u}) = \max_{a \in \mathcal{M}} p_u(a)$$
Informed guess

By making use of her knowledge of $x$, the optimal guess for $E$ is a function of $x$

$$\hat{u} = g(x)$$

with

$$g : \mathcal{X} \mapsto \mathcal{M} \quad g(b) = \arg \max_{a \in \mathcal{M}} p_{u|x}(a|b)$$

and the corresponding success probability is

$$P[\hat{u} = u] = P[g(x) = u] = \sum_{b \in \mathcal{X}} P[g(x) = u | x = b] p_x(b)$$

$$= \sum_{b \in \mathcal{X}} p_{u|x}(g(b)|b)p_x(b) = \sum_{b \in \mathcal{X}} p_x(b) \max_{a \in \mathcal{M}} p_{u|x}(a|b)$$

In general, it is not lower than the ignorant guess, as

$$\sum_{b \in \mathcal{X}} p_x(b) \max_{a \in \mathcal{M}} p_{u|x}(a|b) \geq \max_{a \in \mathcal{M}} \sum_{b \in \mathcal{X}} p_x(b)p_{u|x}(a|b) = \max_{a \in \mathcal{M}} p_u(a)$$
Sequential guessing

If E has a means to check the correctness of her guess she can repeat guesses $\hat{u}_i, i = 1, 2, \ldots$ until she hits the correct plaintext. The optimal choice for the $i$-th ignorant guess is recursively defined as

$$\hat{u}_i = \begin{cases} \arg \max_{a \in \mathcal{M}} p_u(a), & i = 1 \\ \arg \max_{a \in \mathcal{M} \setminus \{\hat{u}_1, \ldots, \hat{u}_{i-1}\}} p_u(a), & i > 1 \end{cases}$$

while for informed guesses

$$\hat{u}_i = g_i(x), \quad g_i(b) = \begin{cases} \arg \max_{a \in \mathcal{M}} p_{u|x}(a|b), & i = 1 \\ \arg \max_{a \in \mathcal{M} \setminus \{\hat{u}_1, \ldots, \hat{u}_{i-1}\}} p_{u|x}(a|b), & i > 1 \end{cases}$$

The attack performance is evaluated in terms of

- probability of success in or before $N$ guesses:
  $$P \left[ \bigcup_{i=1}^{N} \hat{u}_i = u \right] = \sum_{i=1}^{N} P [\hat{u}_i = u]$$
- statistics of the number of attempts before success
In the ideal world model of encryption, the secret message $u$ is directly delivered, unmodified to B, and the message observed by E is generated independently from $u$. 
Perfect Secrecy

Definition

An encryption system is **perfect** if it provides **0-unconditional** security based on indistinguishability, i.e. the plaintext is statistically independent of the ciphertext

\[ p_{u|x}(a|b) = p_u(a) \quad \forall a \in \mathcal{M}, b \in \mathcal{X} \]

or equivalently

\[ p_{ux}(a, b) = p_u(a)p_x(b) \quad \forall a \in \mathcal{M}, b \in \mathcal{X} \]
\[ p_{x|u}(b|a) = p_x(b) \quad \forall a \in \mathcal{M}, b \in \mathcal{X} \]

In a system with perfect secrecy the optimal informed guessing strategy coincides with the optimal ignorant guessing
One-time pad

Let \((G, \circ)\) be a finite group, [e.g., \((\mathbb{Z}_N, + \mod N)\)]. A one-time pad (OTP) over \((G, \circ)\) is the encryption system described by

- equal spaces \(M = X = K = G\)
- uniform key \(k \sim U(G) \iff p_k(a) = \frac{1}{|G|} \forall a \in G\)
- encrypt by add \(E(a, b) = b \circ a\)
- decrypt by subtract \(D(a, c) = c \circ a^{-1}\)

Example

Let \(G = \mathbb{B}^N\), with \(\mathbb{B} = \{0, 1\}\), \(N = 5\), and \(\circ = \text{bitwise XOR}\). Then, e.g.,

\[u = 01101, \; k = 10110 \quad \Rightarrow \quad x = u \circ k = 11011\]

B can recover the message with \(k^{-1} = k = 10110\)

\[u = x \circ k^{-1} = 01101\]
Secrecy of one-time pad

**Theorem**

The one-time pad offers perfect reliability and perfect secrecy for any message distribution.

**Proof.**

Perfect reliability is guaranteed by the existence and uniqueness of $k^{-1} \in G$.

As regards perfect secrecy, we prove that $p_{u,x}(b, c) = p_u(b)p_x(c), \forall b \in \mathcal{M}, c \in \mathcal{X}$. In fact,

$$p_{u,x}(b, c) = P[u = b, x = c] = P[u = b, k = b^{-1} \circ c]$$

$$= p_u(b)p_k(b^{-1} \circ c) = p_u(b)/|\mathcal{K}|$$

$$p_x(c) = \sum_{b \in \mathcal{M}} p_{u,x}(b, c) = \sum_{b \in \mathcal{M}} p_u(b)/|\mathcal{K}| = 1/|\mathcal{K}|$$

Observe that this result holds for any $p_u(\cdot)$. 


Chain rules for (conditional) entropy

For any collection of rvs $x_1, \ldots, x_n, y_1, \ldots, y_m, z_1, \ldots, z_\ell$ the following chain rules hold:

1. $H(x_1, \ldots, x_n, z_1, \ldots, z_\ell | y_1, \ldots, y_m) \geq H(x_1, \ldots, x_n | y_1, \ldots, y_m)$, entropy increases with more conditioned variables

2. $H(x_1, \ldots, x_n | y_1, \ldots, y_m, z_1, \ldots, z_\ell) \leq H(x_1, \ldots, x_n | y_1, \ldots, y_m)$, entropy decreases with more conditioning variables

3. $H(x_1, \ldots, x_n, z_1, \ldots, z_\ell | y_1, \ldots, y_m) = H(x_1, \ldots, x_n | y_1, \ldots, y_m, z_1, \ldots, z_\ell) + H(z_1, \ldots, z_\ell | y_1, \ldots, y_m)$

They are the obvious generalization of, respectively:

1. $H(x, z) \geq H(x)$

2. $H(x | z) \leq H(x)$

3. $H(x, z) = H(x | z) + H(z)$
Necessary condition for perfect secrecy

Theorem

A necessary condition for perfect secrecy and decodability is that

\[ H(k) \geq H(u) \]

Proof.

Assume perfect secrecy holds, that is \( u \) is independent of \( x \). Then,

\[
\begin{align*}
H(u) &= H(u|x) & \text{by independence of } u, x \\
&\leq H(u,k|x) & \text{by chain rule 1} \\
&= H(u|x,k) + H(k|x) & \text{by chain rule 3} \\
&= H(k|x) & \text{by perfect decodability} \\
&\leq H(k) & \text{by chain rule 2}
\end{align*}
\]
Corollary

In a system with perfect secrecy for all message distributions $p_u$ we have

$$\log_2 |\mathcal{K}| \geq H(k) \geq \log_2 |\mathcal{M}|$$

Proof.

$H(k) \leq \log_2 |\mathcal{K}|$ is the upper bound for entropy.
From the previous theorem $H(k) \geq H(u)$ must hold for any $p_u$.
In particular, for uniform $u \sim \mathcal{U}(\mathcal{M})$, where $H(u) = \log_2 |\mathcal{M}|$. 

Corollary

In a system with $\mathcal{M} = \mathcal{A}^{\ell_u}$, $\mathcal{K} = \mathcal{A}^{\ell_k}$, and perfect secrecy, it is $\ell_k \geq \ell_u$.

So, in order to have perfect secrecy, the key must be “at least as long as” the message.