Codes and Cryptography

Jorge L. Villar

MAMME, Fall 2015

PART IX

Outline

The Setting (I)

Perfect environment: No storage or communication errors or excessive message delivery delays

Users: divided into

- good guys (honest)
- bad guys (corrupted by an adversary)

Alternative model: Rational Cryptography (from Game Theory). Only selfish guys (not necessarily honest, can collude).

Simplest case: One honest user, one bad user.
E.g.: Secure binary data storage.

The Setting (II)

Adversarial behavior:

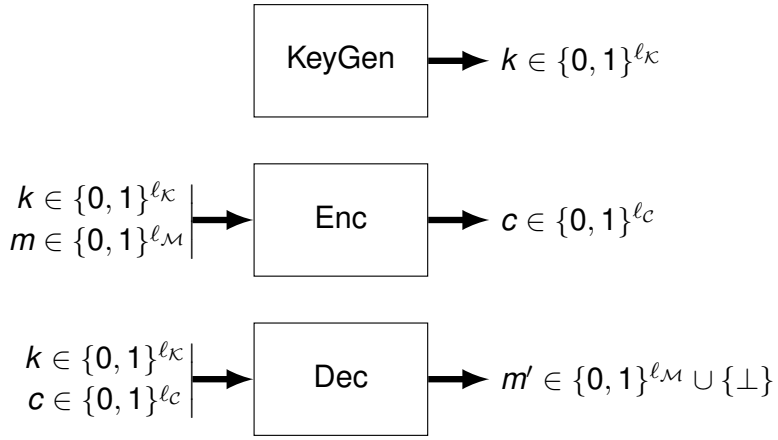
- **static:** corrupted users are fixed before starting the actual attack
- **dynamic:** corrupted users are decided on-the-fly

- **passive:** corrupted users only try to learn more than they are allowed to
- **active:** corrupted users deviate from the protocols in unpredictable ways

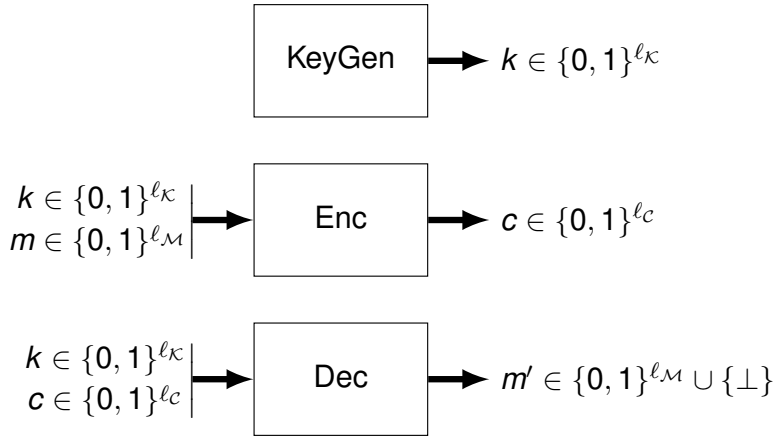
- **bounded:** the adversary has limited resources (computational power, memory)
- **unbounded:** the adversary has unlimited resources

Outline

Symmetric Encryption: Syntax



Symmetric Encryption: Correctness



$\forall m \in \{0,1\}^{\ell_\mathcal{M}},\ \forall k \in \{0,1\}^{\ell_\mathcal{K}},\quad m = \mathrm{Dec}(k, \mathrm{Enc}(k, m))$

Symmetric Encryption: Privacy

Informal definition:
"Impossible to find m from c without k"

More formally:

For any fixed $c \in \{0,1\}^{\ell_\mathcal{C}}$, and for uniformly distributed $k \in \{0,1\}^{\ell_\mathcal{K}}$, $\Pr[c = \mathrm{Enc}(k, m)]$ is the same for all $m \in \{0,1\}^{\ell_\mathcal{M}}$.

or:

Definition (Perfect Privacy)

For any probability distribution (source) of $M \in \{0,1\}^{\ell_\mathcal{M}}$ and for a uniformly distributed $K \in \{0,1\}^{\ell_\mathcal{K}}$, **the random variables** M **and** $\mathrm{Enc}(K, M)$ **are independent.**

Bounds for Perfect Symmetric Encryption

Theorem

For any correct and perfectly private symmetric encryption scheme $\ell_{\mathcal{C}} \geq \ell_{\mathcal{M}}$ *and* $\ell_{\mathcal{K}} \geq \ell_{\mathcal{M}}$

Proof: (A simple combinatorial argument) details...

Caveat: We assume that all elements in $\{0,1\}^{\ell_{\mathcal{M}}}$ are possible messages (e.g., consider compression codes for redundant sources)

No perfect solution for binary private storage!

The key can only be used once!

Bounds for Perfect Symmetric Encryption: Details

Plot of Enc function:
$G = \{(m, k, c = \mathrm{Enc}(k, m))\}_{m \in \{0,1\}^{\ell_{\mathcal{M}}}, k \in \{0,1\}^{\ell_{\mathcal{K}}}}$

Correctness implies
$(m, k, c), (m', k, c) \in G \quad \Rightarrow \quad m' = m$
Then fixing k, we obtain $\ell_{\mathcal{C}} \geq \ell_{\mathcal{M}}$

In addition, Perfect Privacy implies
$(m, k, c) \in G \quad \Rightarrow \quad \forall m' \; \exists k' \quad (m', k', c) \in G$
Then fixing c, we obtain $\ell_{\mathcal{K}} \geq \ell_{\mathcal{M}}$

go back...

A Generalization for Redundant Sources

Replace the sets $\{0,1\}^{\ell_{\mathcal{M}}}$, $\{0,1\}^{\ell_{\mathcal{K}}}$, $\{0,1\}^{\ell_{\mathcal{C}}}$ by probability distributions M, K, C on some finite sets $\mathcal{M}$, $\mathcal{K}$, $\mathcal{C}$.

Replace binary length by a measure of the average information given by a random variable (Shannon's entropy)

Definition (Perfect Privacy)

For any probability distribution (source) of $M \in \{0,1\}^{\ell_{\mathcal{M}}}$ and for a uniformly distributed $K \in \{0,1\}^{\ell_{\mathcal{K}}}$, **the random variables** M **and** $\mathrm{Enc}(K, M)$ **are independent.**

Theorem (Shannon)

For any correct and perfectly private symmetric encryption scheme $H(C) \geq H(M)$ *and* $H(K) \geq H(M)$

Shannon's Entropy of a Discrete Random Variable (I)

It is a measure of the average amount of information in a random variable.

Definition (Shannon's Entropy)

Let X be a random variable with range a finite set $\mathcal{X}$.
$H(X) = -\sum_{x \in \mathcal{X}} \Pr[X = x] \log_2 \Pr[X = x]$

Main properties:

- For any joint distribution (X, Y, Z)
  $H(X, Y, Z) + H(Z) \leq H(X, Z) + H(Y, Z)$ details...
- For a trivial (deterministic) random variable $H(0) = 0$
- For a uniform distribution $H(U_{\mathcal{X}}) = \log_2 |\mathcal{X}|$

Shannon's Entropy Main Inequality (Submodularity)

The map $t \mapsto t \log_2 t$ is convex in $[0, +\infty)$

By discrete Jensen's inequality for convex functions, for positive numbers a_i, t_i, $i \in \mathcal{I}$

$$\sum_{i \in \mathcal{I}} a_i t_i = 1 \quad \Rightarrow \quad \sum_{i \in \mathcal{I}} a_i t_i \log_2 t_i \geq -\log_2 \sum_{i \in \mathcal{I}} a_i$$

Then use

$$t_{x,y,z} = \frac{\Pr[X = x, Y = y, Z = z] \Pr[Z = z]}{\Pr[X = x, Z = z] \Pr[Y = y, Z = z]}$$

$$a_{x,y,z} = \frac{\Pr[X = x, Z = z] \Pr[Y = y, Z = z]}{\Pr[Z = z]}$$

go back...

Shannon's Entropy of a Discrete Random Variable (II)

Other properties:

- $0 \leq H(X) \leq \log_2 |\mathcal{X}|$ (with r.h.s. equality for the uniform distribution)
- $H(X) \leq H(X, Y) \leq H(X) + H(Y)$ (with r.h.s. equality for independent variables)
- $H(g(X)) \leq H(X)$ (with equality for injective maps)
- $H(X, g(X)) = H(X)$

Proof of Shannon's Theorem

Assume that $C = \mathrm{Enc}(K, M)$, and K and M are independent

Correctness implies
$H(M, K, C) = H(\mathrm{Dec}(K, C), K, C) = H(K, C) \leq H(K) + H(C)$

Perfect Privacy implies
$H(M, C) = H(M) + H(C)$

Therefore,
$H(K) + H(C) \geq H(M, K, C) \geq H(M, C) = H(M) + H(C) \quad \Rightarrow$
$H(K) \geq H(M)$

Moreover,
$H(K) + H(M) = H(K, M) = H(K, \mathrm{Dec}(K, C)) \leq H(K, C) \leq$
$H(K) + H(C) \quad \Rightarrow \quad H(M) \leq H(C)$

The One-Time Pad

For fixed length binary strings, $\ell_{\mathcal{M}} = \ell_{\mathcal{K}} = \ell_{\mathcal{C}} = \ell$, and
$\mathrm{Enc}(k, m) = k \oplus m$ and $\mathrm{Dec}(k, c) = k \oplus c$

For an abelian (additive) group $\mathcal{G}$, let $\mathcal{M} = \mathcal{K} = \mathcal{C} = \mathcal{G}$, and
$\mathrm{Enc}(k, m) = m + k$ and $\mathrm{Dec}(k, c) = c - k$

Perfect secrecy is guaranteed if k is uniformly distributed in $\mathcal{G}$

Normally used as an "information theoretical" piece in other protocols

Weakening Secrecy

To overcome the previous limitations, consider only **computationally bounded adversaries**

Definition (Perfect Privacy)

For any probability distribution (source) of $M \in \{0,1\}^{\ell_{\mathcal{M}}}$ and for a uniformly distributed $K \in \{0,1\}^{\ell_{\mathcal{K}}}$, **the random variables** M **and** $\mathrm{Enc}(K, M)$ **are independent.**

Definition (Informal Computational Privacy)

For any probability distribution (source) of $M \in \{0,1\}^{\ell_{\mathcal{M}}}$ and for a uniformly distributed $K \in \{0,1\}^{\ell_{\mathcal{K}}}$, **the random variables** M **and** $\mathrm{Enc}(K, M)$ **behave as if they were independent for a bounded adversary.**

Cryptography: The setting
Symmetric Encryption (I)

Weakening Secrecy

Definition (Informal Computational Privacy)

For any probability distribution (source) of $M \in \{0,1\}^{\ell_{\mathcal{M}}}$ and for a uniformly distributed $K \in \{0,1\}^{\ell_{\mathcal{K}}}$, **the random variables** M **and** $\mathrm{Enc}(K, M)$ **behave as if they were independent for a bounded adversary.**

Based on efficient statistical tests a computationally bounded adversary can run

Needs some extra assumptions from Complexity Theory (similarly as working with the Riemann Hypothesis)

Jorge L. Villar CODES & CRYPTO

Codes and Cryptography

Jorge L. Villar

MAMME, Fall 2015

END OF PART IX

Jorge L. Villar CODES & CRYPTO