Cryptography

Lecture 2

Announcements

- HW1 due Wednesday, 2/7 at beginning of class
- Discrete Math Readings/Quizzes due Wed, 1/31 @ 11:59pm

Agenda

- Last time:
  - Historical ciphers and their cryptanalysis (K/L 1.3)

- This time:
  - Formal definition of symmetric key encryption (K/L 2.1)
  - Definition of information-theoretic security (K/L 2.2)
  - Variations on the definition and proofs of equivalence (K/L 2.2)
  - One-Time-Pad (OTP) (K/L 2.2)

Formally Defining a Symmetric Key Encryption Scheme

Perfect Secrecy: Claude Shannon

Sender

"Offline"

Gen → K

Receiver

K

K

m

c

$c \leftarrow Enc(K, m)$

$Enc_K(m)$

Eve

$m = Dec(K, c)$

$= Dec_K(c)$

Correctness: $Dec_K(Enc_K(m)) = m$

Syntax

- An encryption scheme is defined by three algorithms
  - Gen, Enc, Dec
- Specification of message space $\boldsymbol{M}$ with $|\boldsymbol{M}| > 1$.

  ex: $\{0,1\}^\ell$

  $\{0,1\} \times \{0,1\} \cdots \cdots \{0,1\}$
- Key-generation algorithm Gen:
  - Probabilistic algorithm
  - Outputs a key k according to some distribution.
  - Keyspace $\boldsymbol{K}$ is the set of all possible keys
- Encryption algorithm Enc:
  - Takes as input key $k \in \boldsymbol{K}$, message $m \in \boldsymbol{M}$
  - Encryption algorithm may be probabilistic
  - Outputs ciphertext $c \leftarrow Enc_k(m)$
  - Ciphertext space $\boldsymbol{C}$ is the set of all possible ciphertexts
- Decryption algorithm Dec:
  - Takes as input key $k \in \boldsymbol{K}$, ciphertext $c \in \boldsymbol{C}$
  - Decryption is deterministic
  - Outputs message $m \coloneqq Dec\_k(c)$

Distributions over K, M, C

- Distribution over $\boldsymbol{K}$ is defined by running Gen and taking the output.
  - For $k \in \boldsymbol{K}, \Pr[K = k]$ denotes the prob that the key output by Gen is equal to k.
- For $m \in \boldsymbol{M}, \Pr[M = m]$ denotes the prob. That the message is equal to m.
  - Models a priori knowledge of adversary about the message.
  - E.g. Message is English text.
- Distributions over $\boldsymbol{K}$ and $\boldsymbol{M}$ are independent.
- For $c \in \boldsymbol{C}, \Pr[C = c]$ denotes the probability that the ciphertext is c.
  - Given Enc, distribution over $\boldsymbol{C}$ is fully determined by the distributions over $\boldsymbol{K}$ and $\boldsymbol{M}$.

Gen
1. Key Gen
2. sample m from msg dist.
3. $c \leftarrow Enc_k(m)$

Definition of Perfect Secrecy

- An encryption scheme (Gen, Enc, Dec) over a message space $\boldsymbol{M}$ is perfectly secret if for every probability distribution over $\boldsymbol{M}$, every message $m \in \boldsymbol{M}$, and every ciphertext $c \in \boldsymbol{C}$ for which $\Pr[C = c] > 0$:

$$\Pr[M = m \mid C = c] = \Pr[M = m].$$

a posteriori

a priori

All info that "Eve" knows about the sender's msg a priori

An Equivalent Formulation

- Lemma: An encryption scheme (Gen, Enc, Dec) over a message space $\boldsymbol{M}$ is perfectly secret if and only if for every probability distribution over $\boldsymbol{M}$, every message $m \in \boldsymbol{M}$, and every ciphertext $c \in \boldsymbol{C}$:

$$\Pr[C = c \mid M = m] = \Pr[C = c].$$

If $p \rightarrow q$.

$p \leftrightarrow q$

$p \rightarrow q$

$q \rightarrow p$

Intuitive: The ciphertext is independent of the message.

How? From the perspective of Eve who doesn't know the key.

Proof: We will do: $p \to q$

① Fix an arbitrary dist over $\mathcal{M}$

② " " " message $m \in \mathcal{M}$

③ " " " ciphertext $c \in \mathcal{C}$

(must use assumption somewhere)

$$\Pr[C=c \mid M=m] = \frac{\Pr[M=m \mid C=c] \cdot \Pr[C=c]}{\Pr[M=m]}$$ By Bayes' Law

$$= \frac{\Pr[M=m] \cdot \Pr[C=c]}{\Pr[M=m]}$$ By def of perfect secrecy

$$= \Pr[C=c]$$

☑

Basic Logic

- Usually want to prove statements like $P \to Q$ ("if P then Q")
- To prove a statement $P \to Q$ we may:
  - Assume P is true and show that Q is true.
  - Prove the contrapositive: Assume that Q is false and show that P is false.

Basic Logic

- Consider a statement $P \leftrightarrow Q$ (P if and only if Q)
  - Ex: Two events X, Y are independent if and only if $\Pr[X \wedge Y] = \Pr[X] \cdot \Pr[Y]$.
- To prove a statement $P \leftrightarrow Q$ it is sufficient to prove:
  - $P \rightarrow Q$
  - $Q \rightarrow P$

Proof (Preliminaries)

- Recall Bayes' Theorem:
  - $\Pr[A \mid B] = \frac{\Pr[B|A] \cdot \Pr[A]}{\Pr[B]}$
- We will use it in the following way:
  - $\Pr[M = m \mid C = c] = \frac{\Pr[C=c \mid M=m] \cdot \Pr[M=m]}{\Pr[C=c]}$

Proof

Proof: $\rightarrow$

- To prove: If an encryption scheme is perfectly secret then

"for every probability distribution over $\boldsymbol{M}$, every message $m \in \boldsymbol{M}$, and every ciphertext $c \in \boldsymbol{C}$:

$$\Pr[C = c \mid M = m] = \Pr[C = c]."$$

Proof (cont'd)

- Fix some probability distribution over $\boldsymbol{M}$, some message $m \in \boldsymbol{M}$, and some ciphertext $c \in \boldsymbol{C}$.
- By perfect secrecy we have that

$$\Pr[M = m \mid C = c] = \Pr[M = m].$$

- By Bayes' Theorem we have that:

$$\Pr[M = m \mid C = c] = \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\Pr[C = c]} = \Pr[M = m].$$

- Rearranging terms we have:

$$\Pr[C = c \mid M = m] = \Pr[C = c].$$

Perfect Indistinguishability

- Lemma: An encryption scheme (Gen, Enc, Dec) over a message space M is perfectly secret if and only if for every probability distribution over M, every $m_0, m_1 \in M$, and every ciphertext $c \in C$: $\Pr[C = c \mid M = m_0] = \Pr[C = c \mid M = m_1]$.

Proof (Preliminaries)

- Let $F, E_1, \dots, E_n$ be events such that $\Pr[E_1 \vee \cdots \vee E_n] = 1$ and $\Pr\left[E_i \wedge E_j\right] = 0$ for all $i \neq j$.
- The E_i partition the space of all possible events so that with probability 1 exactly one of the events E_i occurs. Then

$$\Pr[F] = \sum_{i=1}^{n} \Pr[F \wedge E_i]$$

Proof Preliminaries

- We will use the above in the following way:
- For each $m_i \in M$, E_{m_i} is the event that $M = m_i$.
- F is the event that $C = c$.
- Note $\Pr\left[E_{m_1} \vee \cdots \vee E_{m_n}\right] = 1$ and $\Pr\left[E_{m_i} \wedge E_{m_j}\right] = 0$ for all $i \neq j$.
- So we have:
  - $\Pr[C = c] = \sum_{m \in M} \Pr[C = c \wedge M = m]$

$$= \sum_{m \in M} \Pr[C = c | M = m] \cdot \Pr[M = m]$$

Proof

Proof:$\rightarrow$

Assume the encryption scheme is perfectly secret. Fix messages $m_0, m_1 \in M$ and ciphertext $c \in C$.

$$\Pr[C = c|M = m_0] = \Pr[C = c] = \Pr[C = c \mid M = m_1]$$

Proof

Proof $\leftarrow$

- Assume that for every probability distribution over M, every $m_0, m_1 \in M$, and every ciphertext $c \in C$ for which $\Pr[C = c] > 0$:
  $$\Pr[C = c \mid M = m_0] = \Pr[C = c \mid M = m_1].$$
- Fix some distribution over M, and arbitrary $m_0 \in M$ and $c \in C$.
- Define $p = \Pr[C = c \mid M = m_0]$.
- Note that for all m:
  $\Pr[C = c \mid M = m] = \Pr[C = c \mid M = m_0] = p$.

Proof

- $\Pr[C = c] = \sum_{m \in M} \Pr[C = c \wedge M = m]$

$$= \sum_{m \in M} \Pr[C = c | M = m] \cdot \Pr[M = m]$$

$$= \sum_{m \in M} p \cdot \Pr[M = m]$$

$$= p \cdot \sum_{m \in M} \Pr[M = m]$$

$$= p$$

$$= \Pr[C = c \mid M = m_0]$$

Since m was arbitrary, we have shown that
$\Pr[C = c] = \Pr[C = c \mid M = m]$ for all $c \in C, m \in M$.

So we conclude that the scheme is perfectly secret.

The One-Time Pad (Vernam's Cipher)

- In 1917, Vernam patented a cipher now called the one-time pad that obtains perfect secrecy.
- There was no proof of this fact at the time.
- 25 years later, Shannon introduced the notion of perfect secrecy and demonstrated that the one-time pad achieves this level of security.

The One-Time Pad Scheme

1. Fix an integer $\ell > 0$. Then the message space M, key space K, and ciphertext space C are all equal to $\{0,1\}^{\ell}$.
2. The key-generation algorithm Gen works by choosing a string from $K = \{0,1\}^{\ell}$ according to the uniform distribution.
3. Encryption Enc works as follows: given a key $k \in \{0,1\}^{\ell}$, and a message $m \in \{0,1\}^{\ell}$,output $c \coloneqq k \oplus m$.
4. Decryption Dec works as follows: given a key $k \in \{0,1\}^{\ell}$, and a ciphertext $c \in \{0,1\}^{\ell}$, output $m \coloneqq k \oplus c$.

OTP: Keyspace: $\{0,1\}^{\ell=3}$ Message Space: $\{0,1\}^3$

Gen: Choose random K from Keyspace $\rightarrow$ $011 = K$

$Enc(K=011, m=101)$

$$\begin{array}{r} K = 011 \\ m = 101 \\ \hline c = 110 \end{array}$$

$\oplus$ bitwise XOR

$Dec(K=011, c=110)$

$$\begin{array}{r} K = 011 \\ c = 110 \\ \hline m = 101 \end{array}$$

$\oplus$ bitwise XOR

Security of OTP

Theorem: The one-time pad encryption scheme is perfectly secure.

Proof

Proof: Fix some distribution over M and fix an arbitrary $m \in M$ and $c \in C$. For one-time pad:

$$\Pr[C = c | M = m] = \Pr[M \oplus K = c \mid M = m]$$

$$= \Pr[m \oplus K = c] = \Pr[K = m \oplus c] = \frac{1}{2^\ell}$$

Since this holds for all distributions and all m, we have that for every probability distribution over M, every $m_0, m_1 \in M$ and every $c \in C$

$$\Pr[C = c \mid M = m_0] = \frac{1}{2^\ell} = \Pr[C = c | M = m_1]$$

Example Quiz Question for Lecture 2 Material:

Interestingly, all our definitions of perfect secrecy did not explicitly involve the random variable K, corresponding to random choice of key. Consider the following attempted definition of perfect secrecy.

An encryption scheme (Gen, Enc, Dec) over message space $\boldsymbol{M}$ is perfectly secret if for every probability distribution over $\boldsymbol{M}$, every message $m \in \boldsymbol{M}$, and every ciphertext $c \in \boldsymbol{C}$ for which $\Pr[C = c] > 0$,

$\Pr[K = k \mid C = c] = \Pr[K = k]$.

The ciphertext does not contain information about the key

1. Explain the attempted definition in plain English using 1-2 sentences.
2. Why is this a bad definition? Can you describe an encryption scheme that leaks all the information about the message but still satisfies the definition?

$\mathcal{K} = \{101\}$

OTP but with keyspace consisting of a single key.