LAURIER
Inspiring Lives.

Chapter 2

Shannon’s Theory

Shannon’s Theory

1949, “Communication theory of Secrecy Systems” in Bell Systems Tech. Journal.

Two issues:

- What is the concept of perfect secrecy? Does there any cryptosystem provide perfect secrecy?
  - It is possible when a key is used for only one encryption
- How to evaluate a cryptosystem when many plaintexts are encrypted using the same key?

Outline

- Security Categories of Cryptography
- One-time pad
- Elementary probability theory
- Perfect secrecy
- Entropy
- Entropy properties
- Product system

Security Categories (1)

- Computational security:
  - The best algorithm for breaking a cryptosystem requires at least N operations, where N is a very large number
  - No known practical cryptosystem can be proved to be secure under this definition
  - Study w.r.t certain types of attacks (ex. exhaustive key search) does not guarantee security against other type of attack

Security Categories (2)

- Provable security
  - Reduce the security of the cryptosystem to some well-studied problems that is thought to be difficult
  - Ex. RSA ⇔ integer factoring problem
- Unconditional security
  - A cryptosystem cannot be broken, even with infinite computational resources

One-Time Pad

- Unconditional security !!!
- Described by Gilbert Vernam in 1917
- Use a random key that was truly as long as the message, no repetitions

$$P = C = K = (Z_2)^n \qquad x = (x_1, \ldots, x_n) \qquad K = (K_1, \ldots, K_n)$$

$$e_K(x) = (x_1 + K_1, \ldots, x_n + K_n) \mod 2$$

For ciphertext $y = (y_1, \ldots, y_n)$

$$d_K(y) = (y_1 + K_1, \ldots, y_n + K_n) \mod 2$$

Example: one-time pad

- Given ciphertext with Vigenère Cipher:
  ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS

Decrypt by hacker 1:

Ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
Key: pxlmvmsydof**u**yrvzwc tnlebnecvgdupahfzzlmnyih
Plaintext: mr mustard with the candlestick in the hall

Decrypt by hacker 2:

Ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
Key: **p**f**t**gpmiydgaxgoufhklllmhsqdqogtewbqfgyovuhwt
Plaintext: miss scarlet with the knife in the library

Which one?

Problem with one-time pad

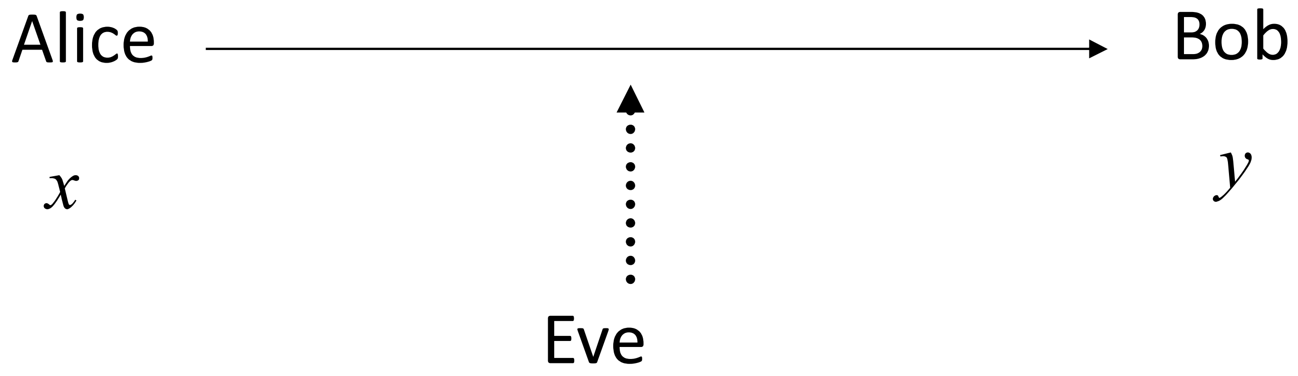
- Truly random key with arbitrary length?
- Distribution and protection of long keys
  - The key has the same length as the plaintext!
- One-time pad was thought to be unbreakable, but there was no mathematical proof until Shannon developed the concept of perfect secrecy 30 years later.

Preview of perfect secrecy (1)

- When we discuss the security of a cryptosystem, we should specify the type of attack that is being considered
  - Ciphertext-only attack
- Unconditional security assumes infinite computational time
  - Theory of computational complexity ×
  - Probability theory √

Preview of perfect secrecy (2)

- **Definition:** A cryptosystem has perfect secrecy if $\Pr[x \mid y] = \Pr[x]$ for all $x \in P$, $y \in C$
- Idea: Eve can obtain no information about the plaintext by observing the ciphertext



Outline

- Introduction
  - One-time pad
- Elementary probability theory
- Perfect secrecy
- Entropy
- Product system

Discrete random variable

- **Def:** A *discrete random variable*, say **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] or Pr[x]
- 0≤Pr[x] for all $x \in X$, $\sum_{x \in X} \Pr[x] = 1$
- Ex. Consider a coin toss to be a random variable defined on {head, tails} , the associated probabilities Pr[head]=Pr[tail]=1/2
- Ex. Throw a pair of dice. It is modeled by Z={(1,1), (1,2), …, (2,1), (2,2), …, (6,6)}
  - where Pr[(i,j)]=1/36 for all i, j.
  - sum=4 corresponds to {(1,3), (2,2), (3,1)} with probability 3/36

Joint and conditional probability

- **X** and **Y** are random variables defined on finite sets *X* and *Y*, respectively.
- **Def:** the joint probability $\Pr[x, y]$ is the probability that **X**=x and **Y**=y
- **Def:** the conditional probability $\Pr[x|y]$ is the probability that **X**=x given **Y**=y

$$\Pr[x, y] = \Pr[x|y]\Pr[y] = \Pr[y|x]\Pr[x]$$

Bayes' theorem

- If Pr[y] > 0, then $\Pr[x \mid y] = \dfrac{\Pr[x]\Pr[y \mid x]}{\Pr[y]}$

- Ex. Let **X** denote the sum of two dice.

  Y is a random variable on {D, N}, **Y**=D if the two dice are the same. (double)

$$\Pr[D \mid 4] = \frac{\Pr[4 \mid D]\Pr[D]}{\Pr[4]} = \frac{(1/6)(1/6)}{3/36} = \frac{1}{3}$$

Outline

- Introduction
  - One-time pad
- Elementary probability theory
- Perfect secrecy
- Entropy
- Product system

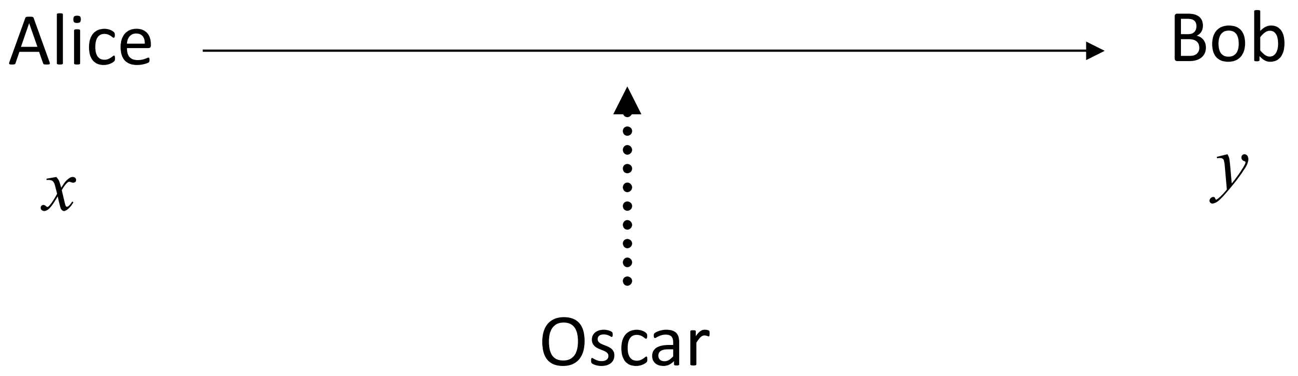
Definitions

- Assume a cryptosystem (P,C,K,E,D) is specified, and a key is used for one encryption
- Plaintext is denoted by random variable $\mathbf{x}$
- Key is denoted by random variable $\mathbf{K}$
- Ciphertext is denoted by random variable $\mathbf{y}$

$$\text{Plaintext } \mathbf{x} \xrightarrow[\mathbf{K}]{} \text{Ciphertext } \boldsymbol{y}$$

Perfect secrecy

- **Definition:** A cryptosystem has perfect secrecy if $\Pr[x|y] = \Pr[x]$ for all $x \in P$, $y \in C$
- Idea: Oscar can obtain no information about the plaintext by observing the ciphertext



Relations among **x, K, y**

- Ciphertext is a function of **x** and **K**

$$\Pr[\mathbf{y} = y] = \sum_{\{K : y \in C(K)\}} \Pr[\mathbf{K} = K] \Pr[\mathbf{x} = d_K(y)]$$

C(K): the set of possible ciphertexts if K is the key

- **y** is the ciphertext, given that **x** is the plaintext

$$\Pr[\mathbf{y} = y \mid \mathbf{x} = x] = \sum_{\{K : x = d_K(y)\}} \Pr[\mathbf{K} = K]$$

Relations among **x**, **K**, **y**

- **x** is the plaintext, given that **y** is the ciphertext

$$\Pr[\mathbf{x}=x \mid \mathbf{y}=y] = \frac{\Pr[x]\Pr[y \mid x]}{\Pr[y]}$$

$$= \frac{\Pr[\mathbf{x}=x] \times \sum_{\{K: x=d_K(y)\}} \Pr[\mathbf{K}=K]}{\sum_{\{K: y \in C(K)\}} \Pr[\mathbf{K}=K]\Pr[\mathbf{x}=d_K(y)]}$$

Ex. Shift cipher has perfect secrecy (1)

- Shift cipher: P=C=K=Z_{26} , encryption is defined as
- Ciphertext: $e_K(x) = (x + K) \bmod 26$

$$\Pr[\mathbf{y} = y] = \sum_{K \in Z_{26}} \Pr[\mathbf{K} = K] \Pr[\mathbf{x} = d_K(y)]$$

$$= \sum_{K \in Z_{26}} \tfrac{1}{26} \Pr[x = y - K]$$

$$= \tfrac{1}{26} \sum_{K \in Z_{26}} \Pr[x = y - K] = \tfrac{1}{26}$$

Ex. Shift cipher has perfect secrecy (2)

- $\Pr[y|x] = \Pr[\mathbf{K} = (y - x) \mod 26] = \frac{1}{26}$
- Apply Bayes' theorem

$$\Pr[x \mid y] = \frac{\Pr[x]\Pr[y \mid x]}{\Pr[y]}$$

$$= \frac{\Pr[x]\frac{1}{26}}{\frac{1}{26}} = \Pr[x]$$

Perfect secrecy

Perfect secrecy when |K|=|C|=|P|

- (P,C,K,E,D) is a cryptosystem where |K|=|C|=|P|, the cryptosystem provides perfect secrecy iff
  - every keys is used with equal probability 1/|K|
  - For every $x \in P$, $y \in C$, there is a unique key K such that

$$e_K(x) = y$$

- Ex. One-time pad in Z_2