# CS3235 Part1 Lec1

$_/|/|U_Ch@NgRu!$

December 16, 2021

## 1 Secuity component

1. Goals: achieve some functionality in the presense of adversaries
2. Policies(Security objective)
   (a) Confidentiality
   (b) Integrity
   (c) Availability
3. Threat model (Attacker): who is the attacker, what does he know, what can he do
4. Security mechanisms: Techniques to achieve the security goal
5. Goal: Policy cannot be violated within the threat model

IMPORTANT: Security goal is independent from the security mechanism

## 2 Kerckhoffs's principle

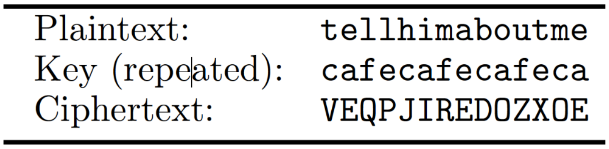
Shannon's Maxim: the enemy kn ows the system

### 2.1 Arguments in favor of the principle

1. Easier to keep key secret than algorithm
2. Easier to change key than to change algorithm
3. Standarization(Ease of deployment, public scrutiny)

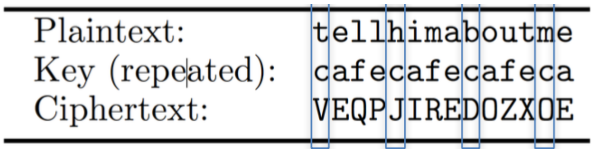
# 3 Vigenere Cipher(poly-alphabetic shift)

A short key was repeated and used as number of shift for each character of the plaintext



| | |
|---|---|
| Plaintext: | tellhimaboutme |
| Key (repeated): | cafecafecafeca |
| Ciphertext: | VEQPJIREDOZXOE |

The advantage of the Vigenere Cipher is that it smooth out the correlation of character frequency between cipher-text and plaintext and hence make statistical analysis harder
However, frequency anaysis can still be applied to guess the password length, and with the known password length, can apply frequency analysis to get the password and plaintext



| | |
|---|---|
| Plaintext: | tellhimaboutme |
| Key (repeated): | cafecafecafeca |
| Ciphertext: | VEQPJIREDOZXOE |

- When the length of the key k is known
  - Divide the ciphertext into t parts
  - Perform statistical analysis for each part
- When the length is unknown(but max length T is known)
- Brute-force the key length
- i.e. repeat the above attack T times (for t in 1,2,...,T)

# 4 Definition of Security

## 4.1 Importance of definition

Definition is important for cryptography in:

1. Design
2. Analysis
3. Sound Usage

### 4.1.1 Influence for design

- Precise definition → enforce designers to consider what they really want

- Help find what is necessary and what is not
- Unable define the feature precisely → not know when having achieved it

#### 4.1.2 Influence for analysis

Precise definitions enables meaningful analysis, evaluation and comparison of schemes

- Does one scheme satisfy certain definition
- What definition has certain scheme satisfy
- There may be multiple definitions
- Comparison between schemes can be made based of scheme's level of satisfying definitions
- A scheme that satisfy more stronger definitions is thought to be more efficient

#### 4.1.3 Influence of definition on usage

- Allows users to understand the security guarantees provided by a scheme
- Allows scheme to be part of a larger system
- Allows scheme to be substituted by another scheme that achieves the same or better safety

## 5 Assumption

Most of cryptography is based on the assumption of current computational power
e.g. One problem is hard to be solved within certain time based on certain condition(e.g. $\text{P} \neq \text{NP}$)
Principle: any such assumption should be made explicitly

#### 5.0.1 Importance of clear assumption

1. Enables studies for validating the assumption
2. Allows comparison of schemes based on different assumptions, and there is study for the minimal assumptions needed
3. Possible for study of practical implementations if the study is wrong(e.g. the study of cryptography after the application of quantum computers)
4. Enables proof of security

# 6 Proof of security

1. Provide a rigorous proof that construction satisfies certain definition under certern assumptions
2. A unbreakable proof is crucial because it works as a unbreakable guarantee

## 6.1 Limitations of cryptography

- Validity of various assumption is an active area of research
- Cryptography remains partly an art(not understand fully)
- Provably secure scheme can still be broken
  - if the definition does not correspond to the real-world threat model
  - If attacker can go ”outside the security model”(happens a lot in practice)
  - If the assumption is invalid
  - If the implementation is flawed

# 7 General form of security assumption

1. Security guarantee/goal
   - What we want to achieve
   - What we want to prevent attackers to achieve
2. Threat model
   - What (real-world) capibilities that the attacker is assumed to have

# 8 Definition and Assumption of Cryptography

## 8.1 Definition of security in cryptography

Regardless of any **prior** information the attackers have about the plain text, the cipher text should leaks no additional information about the plain text

### 8.1.1 The definition of perfect security does not enforce security on key

Consider a cipher, called one time double pad here, takes as input a message c $\in \{0,1\}^n$ and two keys $k_1, k_2 \in \{0,1\}^n$. The encryption scheme $(Enc, Dec, Gen)$ defined over $\{\mathcal{M}, \mathcal{K}, \mathcal{C}\}$ as follows:

- $c = Enc(k_1, k_2, m) = m \oplus k_1 \oplus k_2$
- $m = Dec(k_1, k_2, m) = c \oplus k_1 \oplus k_2$

In this case, even if leak of information about $k_1$ or $k_2$(not both), the perfect security still holds

## 8.2 Assumption

### 8.2.1 Threat models

1. Ciphertext-only attack: attacker is assumed to only has access to ciphertext
2. Known-plaintext attack: attacker is assumed to have access to both cipher text and plaintext
3. Chosen-plaintext attack: attacker is able to modify the plaintext and get the corresponding changed ciphertext
4. Chosen-ciphertext attack: attacker is able to modify the ciphertext the get the corresponding plaintext decrypted from the modified ciphertext

# 9 Probability Review

Random variable: variable that takes on value with certain probabilities
Probability distribution for a variable specifies the probabilities with which the variable takes on each possible value

1. Each probability must be between 0 and 1(Probability must be positive/zero)
2. The probabilities must sum to 1(Collactive exhausitive)

## 9.1 Events

Particular occurrence of experiment

- Pr[E]: probability of event E

## 9.2 Conditional probability

- probability of one event, given that the other events occur
  - Pr[A|B] = Pr[A and B]/Pr[B]

## 9.3 Independent

- Two random variables X,Y are independent if for all x,y: Pr[X=x|Y=y] =Pr[X=x]

- Two random variables X,Y are independent if for all x,y: Pr[X=x and Y=y]= Pr[X=x]Pr[Y=y]

## 10 Law of total probability/Bayes' first law

Say $E_1, ..., E_n$ are a partition of all possibilities. Then for any A:

- Pr[A]=$\sum_i Pr[A \ and \ E_i] = \sum_i Pr[A|E_i] * Pr[E_i]$

## 11 Bayes' theorem

$Pr[A|B] = Pr[B|A] * \frac{Pr[A]}{Pr[B]}$

## 12 Probability distribution

- Let $\mathcal{M}$ be the set of all plaintext, let M be the variable denoting the value of the message
  - M ranges over $\mathcal{M}$
  - In reality, the distribution of M may not be purely random, it's instead dependent on the scenario of the message.
  - $\forall M \in \mathcal{M}$, Pr[M] reflects the likelihood of the msg being sent, given the attacker's prior knowledge
- Let $\mathcal{K}$ denotes the set of all possible key or key space. Let K be the random. variabl denoting the value of key
  - K ranges over $mathcalK$
  - K is independent of M
- Let $\mathcal{C}$ denotes the set of all possible ciphertext
- $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0, 1\}^n$

Fix some encryption scheme(Gen, Enc, Dec)

- Gen defines a probability distribution for K
  (Different from M, the distribution of K is determined by G)
- Pr[K=k] = Pr[Gen output key k]
- IMPORTANT: variable M and K are independent, K should not be generated with any reference to M

The experiment is conducted as follows:

1. generate key with Gen
2. Choose plaintext M, M follows certain distribution based on the scenario of the message
3. Compute $\underbrace{c = Enc(key, M)}_{this\ defines\ distribution\ of\ ciphertext}$

## 13 Perfect secrecy of One-time pad

Based on the above Bayes' theorem probability distribution, the perfect secrecy of one-time pad can be proved as follows

1. According to Bayes' theorem, $Pr[M = m|C = c] = P[C = c|M = m] * \frac{Pr[M=m]}{Pr[C=c]}$
2. $Pr[C] = \sum_{m'} Pr[C = c|M = m']Pr[M = m'] = \sum_{m'} Pr[K = c \oplus m'|M = m']Pr[M = m'] = \frac{1}{2^n}$(Here we assume $K \in \{0,1\}^n$)
3. Fix arbitary distribution over $\mathcal{M} = \{0,1\}^n$, and arbitrary m,c $\in \{0,1\}^n$
4. $Pr[M = m|C = c] = Pr[C = c|M = m] * \frac{Pr[M=m]}{Pr[C=c]} = Pr[K = c \oplus m|M = m] * \frac{Pr[M=m]}{Pr[C=c]}$ $= 2^{-n} \frac{Pr[M=m]}{2^{-n}} = Pr[M = m]$
5. Therefore, perfect secrecy is achieved