Classical Encryption Techniques

What is a Cryptosystem?

Cryptosystem

A cryptosystem is pair of algorithms that take a key and convert plaintext to ciphertext and back.

Plaintext is what you want to protect;
The design and analysis of todays cryptographic algorithms is highly mathematical.

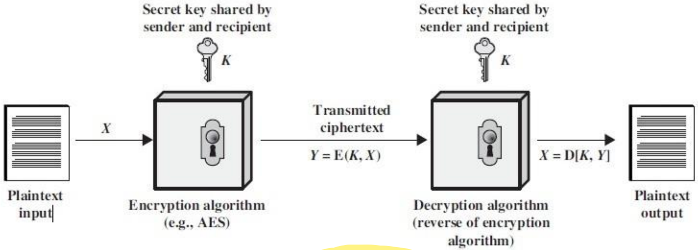
At least not at this stage

Do not try to design your own algorithms.

Some Basic Terminology

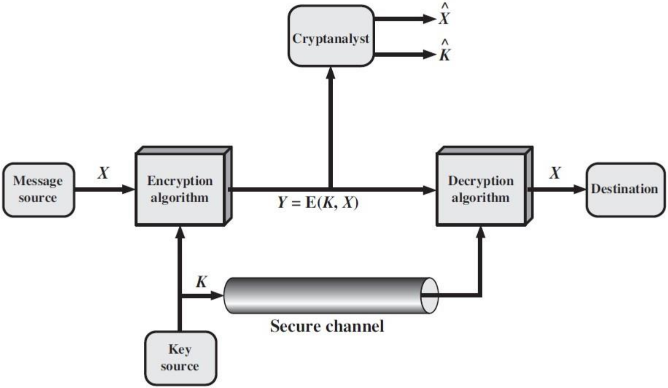
- **Plaintext** - original message
- **Ciphertext** - coded message
- **Cipher** - algorithm for transforming plaintext to ciphertext
- **Key** - info used in cipher; known only to sender/receiver; independent of the plaintext
- **Encipher (encrypt)** - converting plaintext to ciphertext
- **Decipher (decrypt)** - recovering ciphertext from plaintext
- **Cryptography** - study of encryption principles/methods
- **Cryptanalysis (code breaking)** - study of principles/ methods of deciphering ciphertext *without* knowing key
- **Cryptology** - field of both cryptography and cryptanalysis

Symmetric Cipher Model



Simplified Model of Symmetric Encryption

Symmetric Cryptosystem



Model of Symmetric Cryptosystem

Conventional Encryption

There are two requirements for secure use of conventional encryption:

1. We need a strong encryption algorithm. [Everybody knows algorithm and the cipher text]
2. Sender and receiver must have obtained copies of the secret key in a secure fashion and must keep the key secure.

Cryptosystem Classification

By type of encryption operations used

1. Substitution
2. Transposition

By number of keys used

1. Single-key or private
2. Two-key or public

By the way in which plaintext is processed

1. Block
2. Stream

Cryptanalysis

Cryptanalysis

The process of attempting to discover plaintext(X) or key (K) or both is known as cryptanalysis.

Objective: To recover key not just message

Approaches:

- Cryptanalytic attack
- Brute-force attack

Cryptanalysis (Cont.)

Two more definitions are worthy of note.

1. Unconditionally secure
2. Computationally secure

Following criteria should be met to offer *Computationally secure* algorithm.

- The cost of breaking the cipher exceeds the value of the encrypted information.
- The time required to break the cipher exceeds the useful lifetime of the information.

Substitution Technique

A substitution technique is one in which the letters of plaintext are replaced by other letters or by numbers or symbols.

- Caesar Cipher
- Monoalphabetic Ciphers
- Playfair Cipher
- Hill Cipher
- Polyalphabetic Ciphers
- One-Time Pad

Substitution Technique (Cont.)

Caesar Cipher

- Replaces each letter by 3rd letter on
- Example:

```
meet me after the toga party
PHHW PH DIWHU WKH WRJD SDUWB
```

- Can define transformation as:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Mathematically give each letter a number

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- Then have Caesar cipher as:

$$c = \mathrm{E}(\mathrm{k}, p) = (p + k) \bmod (26)$$

$$p = \mathrm{D}(\mathrm{k}, \mathrm{c}) = (\mathrm{c} - k) \bmod (26)$$

Weakness: Small key space (25 keys)

Substitution Technique (Cont.)

Monoalphabetic Cipher

- Shuffle the letters and map each plaintext letter to a different random ciphertext letter:
- Plain letters: abcdefghijklmnopqrstuvwxyz
  Cipher letters: DKVQFIBJWPESCXHTMYAUOLRGZN
- Plaintext: ifwewishtoreplaceletters
  Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA

Substitution Technique (Cont.)

Monoalphabetic Cipher Security

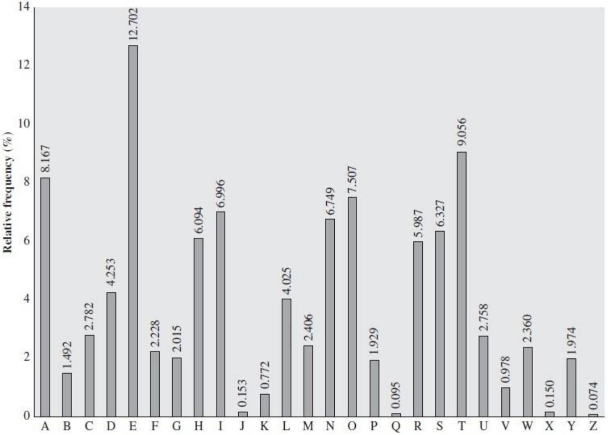
- Now we have a total of 26! keys.
- With so many keys, it is secure against brute-force attacks.
- But not secure against some cryptanalytic attacks.
- Problem is language characteristics.

Substitution Technique (Cont.)

Language Statistics and Cryptanalysis

- Human languages are not random.
- Letters are not equally frequently used.
- In English, E is by far the most common letter, followed by T, R, N, I, O, A, S.
- Other letters like Z, J, K, Q, X are fairly rare.
- There are tables of single, double & triple letter frequencies for various languages

Substitution Technique (Cont.)



Relative Frequency of Letters in English Text

Statistics for double & triple letters

- Double letters:

  th he an in er re es on, …

- Triple letters:

  the and ent ion tio for nde, …

Substitution Technique (Cont.)

Playfair Cipher

- Not even the large number of keys in a monoalphabetic cipher provides security
- In playfair cipher unlike traditional cipher we **encrypt a pair of alphabets(digraphs)** instead of a single alphabet.
- Invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair

Substitution Technique (Cont.)

The Playfair Cipher Encryption Algorithm:

The Algorithm consists of 2 steps:

1. **Generate the key Square(5×5)**
2. **Algorithm to encrypt the plain text**

Step 1: Generate the key Square

Playfair Key Matrix

- a 5X5 matrix of letters based on a keyword
- fill in letters of keyword and fill rest of matrix with other letters.
- eg. using the keyword MONARCHY

M	O	N	A	R
C	H	Y		

Step 1: Generate the key Square

Playfair Key Matrix

- a 5X5 matrix of letters based on a keyword
- fill in letters of keyword and fill rest of matrix with other letters.
- eg. using the keyword MONARCHY

M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Step 2: Algorithm to encrypt the plain text

The plaintext is **split** into pairs of two letters (digraphs). If there is an odd number of letters, a Z is added to the last letter.

For example:
Plain Text: "instrument"
After Split: 'in' 'st' 'ru' 'me' 'nt'

Step 2: Algorithm to encrypt the plain text (Diagraph Generation)

Rule 1: Pair cannot be made with same letter. Break the letter in single and **add a bogus** letter to the previous letter.

Plain Text: “hello”

After Split: ‘he’ ‘lx’ ‘lo’

Here **‘x’** is the bogus letter.

Rule 2: If the letter is standing alone in the process of pairing, then add an extra bogus letter with the alone letter

Plain Text: “helloe”

After Split: ‘he’ ‘lx’ ‘lo’ ‘ez’

Here **‘z’** is the bogus letter.

Step 2: Algorithm to encrypt the plain text (Rules for Encryption)

Rule 1: If both the letters are in the same row: Take the letter to the right of each one (going back to the leftmost if at the rightmost position).

Rule 2: If both the letters of diagraph are in the same column: Take the letter below each one (going back to the top if at the bottom).

Rule 3: If neither of the above rules is true: Form a rectangle with the two letters and take the letters on the horizontal opposite corner of the rectangle.

Step 2: Algorithm to encrypt the plain text (Rules for Encryption)

Rule 1: If both the letters are in the same row: Take the letter to the right of each one (going back to the leftmost if at the rightmost position).

Plain Text: mosque
Digraph: “mo” “sq” “ue”

Ciphertext: “ON”

M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Step 2: Algorithm to encrypt the plain text (Rules for Encryption)

Rule 1: If both the letters are in the same row: Take the letter to the right of each one (going back to the leftmost if at the rightmost position).

Plain Text: mosque

Digraph: “mo” “sq” “ue”

Ciphertext: “ON” “TS”

M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Step 2: Algorithm to encrypt the plain text (Rules for Encryption)

Rule 2: If both the letters of diagraph are in the same column: Take the letter below each one (going back to the top if at the bottom).

Plain Text: mosque
Digraph: “mo” “sq” “ue”

Ciphertext: “ON” “TS” “ML”

M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Step 2: Algorithm to encrypt the plain text (Rules for Encryption)

Rule 3: If neither of the above rules is true: Form a rectangle with the two letters and take the letters on the horizontal opposite corner of **the rectangle**.

Plain Text: Attack

Digraph: “at” “ta” “ck”

Ciphertext: “RS”

M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Hill cipher :

- Multi-letter Cipher.
- Encrypt a group of letters: digraph, trigraph or polygraph.

- ❑ Each letter is represented by a number modulo 26. Often the simple scheme A = 0, B = 1, ..., Z = 25 is used, but this is not an essential feature of the cipher.

Substitution Technique: Hill Cipher

Trigraph

This encryption algorithm takes ***m* successive** plaintext letters and substitutes for them *m* ciphertext letters.

For $m = 3$, the system can be described as

$$c_1 = (k_{11}p_1 + k_{12}p_2 + k_{13}p_3) \bmod 26$$
$$c_2 = (k_{21}p_1 + k_{22}p_2 + k_{23}p_3) \bmod 26$$
$$c_3 = (k_{31}p_1 + k_{32}p_2 + k_{33}p_3) \bmod 26$$

This can be expressed in terms of row vectors and matrices:

$$(c_1\ c_2\ c_3) = (p\ \ p_2\ \ p_3)\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \bmod 26$$

$$\mathbf{C} = \mathbf{PK} \bmod 26$$

Substitution Technique: Hill Cipher

Input : Plaintext: ACT
Key: GYBNQKURP
Output : Ciphertext: POH

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

We have to encrypt the message ‘ACT’ (n=3). The key is ‘GYBNQKURP’ which can be written as the nxn matrix.

G	Y	B
N	Q	K
U	R	P

0	2	19
A	C	T

6	24	1
13	16	10
20	17	15

Substitution Technique: Hill Cipher

Input : Plaintext: ACT
Key: GYBNQKURP
Output : Ciphertext: POH

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

$$\begin{bmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 19 \end{bmatrix} = \begin{bmatrix} 67 \\ 222 \\ 319 \end{bmatrix} = \begin{bmatrix} 15 \\ 14 \\ 7 \end{bmatrix} \text{(Mod 26)}$$

Substitution Technique (Cont.)

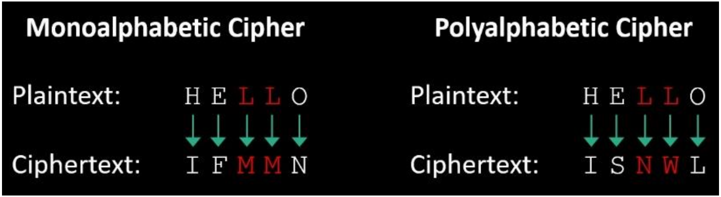
Polyalphabetic Ciphers

- another approach to improving security is to use multiple cipher alphabets
- called **polyalphabetic substitution ciphers**
- makes cryptanalysis harder with more alphabets to guess and flatter frequency distribution
- use a key to select which alphabet is used for each letter of the message
- use each alphabet in turn
- repeat from start after end of key is reached

Example

- write the plaintext out
- write the keyword repeated above it
- use each key letter as a caesar cipher key
- encrypt the corresponding plaintext letter
- eg using keyword *deceptive*

```
key:       deceptivedeceptivedeceptive
plaintext: wearediscoveredsaveyourself
ciphertext:ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```



Substitution Technique (Cont.)

One-Time Pad

- If a truly random key as long as the message is used, the cipher will be secure
- Called a One-Time pad
- Is unbreakable since ciphertext bears no statistical relationship to the plaintext
- Since for **any plaintext** & **any ciphertext** there exists a key mapping one to other
- Can only use the key **once** though
- Problems in generation & safe distribution of key

Transposition Technique

- Consider classical **transposition** or **permutation** ciphers
- these hide the message by rearranging the letter order
- without altering the actual letters used
- can recognise these since have the same frequency distribution as the original text

Transposition Technique (Cont.)

- **Rail Fence Cipher**: Write message out diagonally as:

```
m e m a t r h t g p r y
 e t e f e t e o a a t
```

- Giving ciphertext: MEMATRHTGPRYETEFETEOAAT

- **Row Transposition Ciphers**: Write letters in rows, reorder the columns according to the key before reading off .

```
Key: 4312567
Column Out 4 3 1 2 5 6 7
Plaintext: a t t a c k p
           o s t p o n e
           d u n t i l t
           w o a m x y z
Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ
```

Product Cipher

- Use several ciphers in succession to make harder, but:
  - Two substitutions make a more complex substitution
  - Two transpositions make more complex transposition
  - But a substitution followed by a transposition makes a new much harder cipher
- This is a bridge from classical to modern ciphers

Steganography

Steganography

The practice of **concealing messages or information** within other nonsecret text or data.

Image of a tree with a steganographically hidden image. The hidden image is revealed by removing all but the two least significant bits of each color component and a subsequent normalization.

Image of a cat extracted from the tree image

Source: https://en.wikipedia.org/wiki/Steganography

Summary

- The key methods for cryptography are: Substitution and transposition
- Letter frequency can be used to break substitution
- Substitution can be extended to multiple letters and multiple ciphers. Mono Mono-alphabetic = 1 cipher, Poly Poly-alphabetic = multiple ciphers
- Examples: Caesar cipher (1 letter substitution), Playfair (2-letters), Hill (multiple letters).
- Multiple stages of substitution and transposition can be used to form strong ciphers.

t

Acknowledgement

- Lawrie Browns slides supplied with William Stallings book Cryptography and Network Security: Principles and Practice, 5th Ed, 2011
- Network Security course at Department of Computer Science & Engineering, Washington University in Saint Louis.
- Network Security course at Department of Computer Science, Columbia University, New York.
- http://www.slideshare.net/mohammedarif89/cipher-techniques