

**Cryptography and Security**  
**COMP3441 / 9441 Week 1**

**Part 1: Introduction to cryptography & history**

Reference: The Code Book, Singh

**Aims of cryptography**

Principal aims of cryptography:

- message privacy: ensuring that only the intended parties to the communication can read the message
- message integrity: ensuring that the message received is the same as the message sent

**Note:** Cryptography is a small part of computer security; there are many other issues: access control models, security kernels in operating systems, auditing, management, viruses, worms, denial of service, inference in statistical databases, implementation traps, risk analysis, etc...

**Assessment**

1. Two Assignments @ 25%
2. Final Exam 50 %

### Some Terminology

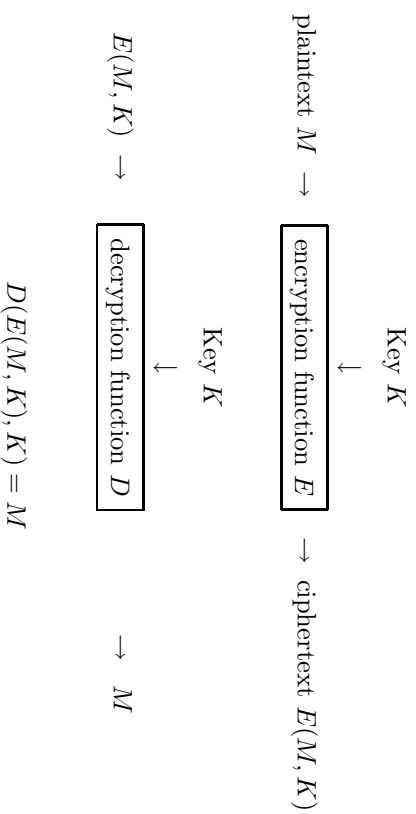
*Codes* are fixed schemes for the use of certain words or sequences of letters with specified meanings not known to the adversary.

E.g. *D-day*

*Ciphers* are general algorithms for obscuring the content of *any* given message.

*Cryptography* is the art of encrypting messages

*Cryptanalysis* is the art of discovering the content of (or *cracking*) encrypted messages



Kerckhoff's Principle (La Cryptographie Militaire, 1883): the security of a cryptosystem must not depend on keeping secret the crypto-algorithm. It must depend only on keeping secret the key.

Reasons:

- Details of crypto-algorithm can be captured/bought/reverse-engineered.
- Even if so, frequently changing the key maintains security

### Some of Julius Caesar's ciphers

(In the Gallic Wars) replaced Roman by Greek characters

Caesar's shift cipher:

$K \in \{1, \dots, 25\}$

$E(M, K)$  = shift each of  $M$  *forward* by  $K$  places in the alphabet

$D(M, K)$  = shift each of  $M$  *back* by  $K$  places in the alphabet

It is wise to omit spaces: e.g., 1-letter words are likely to be "a" or "I", two letter words "an", "at", "it", "he", "do", "go" etc.

### Substitution Cipher

A substitution cipher key  $K$  is a bijective mapping from the alphabet to itself, e.g.

$a$	$b$	$c$	$d$	$e$	$f$	$g$	$h$	$i$	$j$	$k$	$l$	$m$
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
$j$	$i$	$c$	$z$	$b$	$p$	$h$	$y$	$f$	$n$	$s$	$g$	$a$
$n$	$o$	$p$	$q$	$r$	$s$	$t$	$u$	$v$	$w$	$x$	$y$	$z$
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
$m$	$d$	$o$	$l$	$r$	$q$	$v$	$u$	$e$	$w$	$k$	$x$	$t$

$E(M,K)$  = replace each letter  $x$  in  $M$  by  $K(x)$

$D(M,K)$  = replace each letter  $x$  in  $M$  by  $K^{-1}(x)$

There are 400,000,000,000,000,000,000,000 keys!

### Types of Cryptanalytic attack

Cryptanalytic attacks on a cipher can be classified by what the cryptanalyst knows/is able to control:

- *Cipher-text only*
- *Known plain-text*
- *Chosen plain-text*
- *Adaptive chosen plain-text*
- *Rubber hose/purchase key attacks*

### Statistical Cryptanalysis of substitution ciphers

(A cipher-text only attack, described by al-Kindi 9th C.)

In sufficiently long texts letters tend to appear with an accurately predictable frequency.

In English, the most frequently occurring letters, in order of frequency, are ETAOINSHRDLU.

Letter	%	Letter	%
a	8.2	n	6.7
b	1.5	o	7.5
c	2.8	p	1.9
d	4.3	q	0.1
e	12.7	r	6.0
f	2.2	s	6.3
g	2.0	t	9.1
h	6.1	u	2.8
i	7.0	v	1.0
j	0.2	w	2.4
k	0.8	x	0.2
l	4.0	y	2.0
m	2.4	z	0.1

(Beker & Piper, *Cipher Systems: The protection of information*, computed from sources totalling 100,000 characters)

Thus, to decrypt a substitution cipher text,

1. determine the frequency of each letter in the ciphertext.
2. Try a partial substitution in which the most frequent letter of the cipher text corresponds to "E", the next most frequent to "T", etc.
3. If the ciphertext is short, or uses uncommon vocabulary, some changes of order of frequency may occur. Switch correspondences around and decrypt the cipher text according to the partial guess until it starts to look like English words, then proceed as in a crossword puzzle to figure out the rest.

### Vigenere Poly-alphabetical Cipher

Key = a word or phrase

$E(M,K)$  = align K, repeated as many times as necessary, over M. For each letter  $x$  of M,

- let  $y$  be the letter of K above  $x$ ,
  - substitute each letter of M according to the substitution cipher indicated by the row of the table starting with K
- $D(M,K)$  = align K, repeated as many times as necessary, over M. For each letter  $x$  of M,

- let  $y$  be the letter of K above  $x$ ,
- substitute each letter of M according to the inverse of the substitution cipher indicated by the row of the table starting with K

### Bitwise Vigenere = Exclusive Or

$x$	$y$	$x \otimes y$
0	0	0
0	1	1
1	0	1
1	1	0

### Cracking the Vigenere Cipher

(Babbage 1854, published by Kasiki 1863)

Step 1: Determine the key length  $l$ . Look for repeated sequences of letters. These are quite likely to be instances of the same common word (e.g. "the", "and") being encrypted with the same part of the key, yielding repeated substrings, e.g. AXG. Key length likely to be a divisor of distance between such repeats.

Step 2: For each guess of the key length  $l$ , compute frequencies for the all the letters that are at a position number equivalent to  $0 \bmod l$ . (If  $l$  is the correct key length, all these letters were encrypted using the same substitution cipher!) Similarly for the letters at  $1 \bmod l$ ,  $\dots$ , at  $l - 1 \bmod l$ . Use these frequencies to determine the substitution cipher applying to each subset.

### **One time pad**

A *one time pad* essentially a Vignere cipher in which the key length is equal to the length of the message.

Proposed by Major Mauborgne and Vernam (ATT&T) 1917 for encryption of telegraph messages (originally implemented by a punched paper tape.)

Used for encryption of cold war presidential hotline.

*Major Advantage:* cipher text only cryptanalysis is provably impossible.

Suppose  $E(M, K)$  is  $M$  encrypted by one time pad  $K$ . For *every* plaintext  $M'$  of the same length as  $M$ , there exists a key  $K'$  such that

$$E(M', K') = E(M, K)$$

All that the cryptanalyst ever learns is the length of the message!

*Major Disadvantages:* Key distribution, Key synchronization