INFORMATION SECURITY

LECTURE 2

DATA ENCRYPTION/DECRYPTION - SYMMETRIC ALGORITHMS

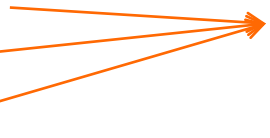
Khalid Abdullah
MSc computer and Network Security
Computer Science Dept.
University Of Zakho

LAST LECTURE

Security: some basic definitions

Security services

- Confidentiality
- Integrity
- Availability
- Authentication
- Non-repudiation
- Access control
- Availability



CIA

Attacks

Security mechanisms

OUTLINE

- **Some basic definitions**
  - Cipher
  - Breaking the code
  - Cryptology
- **Symmetric cryptography**
- **Substitution Techniques**
  - Caesar Cipher
  - **Polyalphabetic Ciphers** (Vigenere Cipher)
  - **Mono-alphabetic** (One time pad)

LESSON OBJECTIVE

- Understand basic definitions related to conventional encryption.
- Understand what is symmetric Key and algorithms
  - Understand some encryption schemes such as
    - Caesar cipher
    - Monoalphabetic ciphers
    - Vignere Cipher
    - One time pad

Learn how to break the encryption schemes in order to build better schemes

BASIC DEFINITIONS

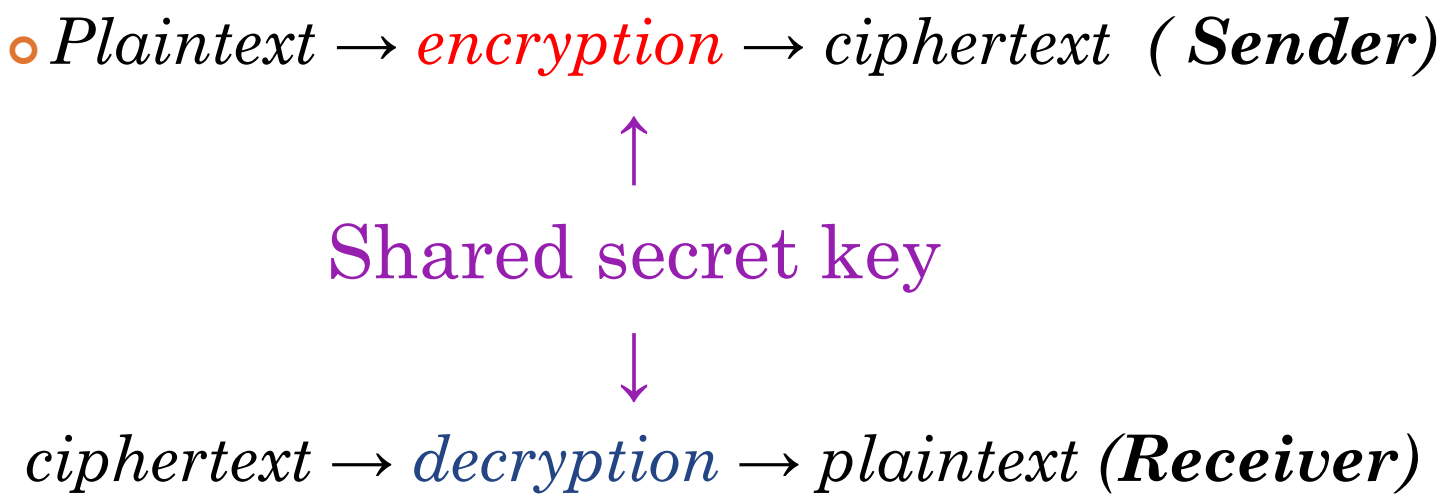
- **Plaintext:** *This is intelligible (Clear ,original) message that need to be send.*
- **Ciphertext:** *This is encrypted message that is unintelligible by human and computer.*
- **Encryption:** *The process that convert the plaintext to a ciphertext.*
- **Decryption:** *The process of restoring the plaintext from ciphertext.*
- **Cryptography:** *The area of study of different schemes used for decryption.*
- **Cipher:** Is a scheme.

Basic definitions

- **Authentication:** sender, receiver want to confirm identity of each other.
- **Access and Availability:** services must be accessible and available to authorised properly users.

What is Secret Key (Symmetric) Cryptography

- Sender and receiver share the same key used for encryption and decryption.

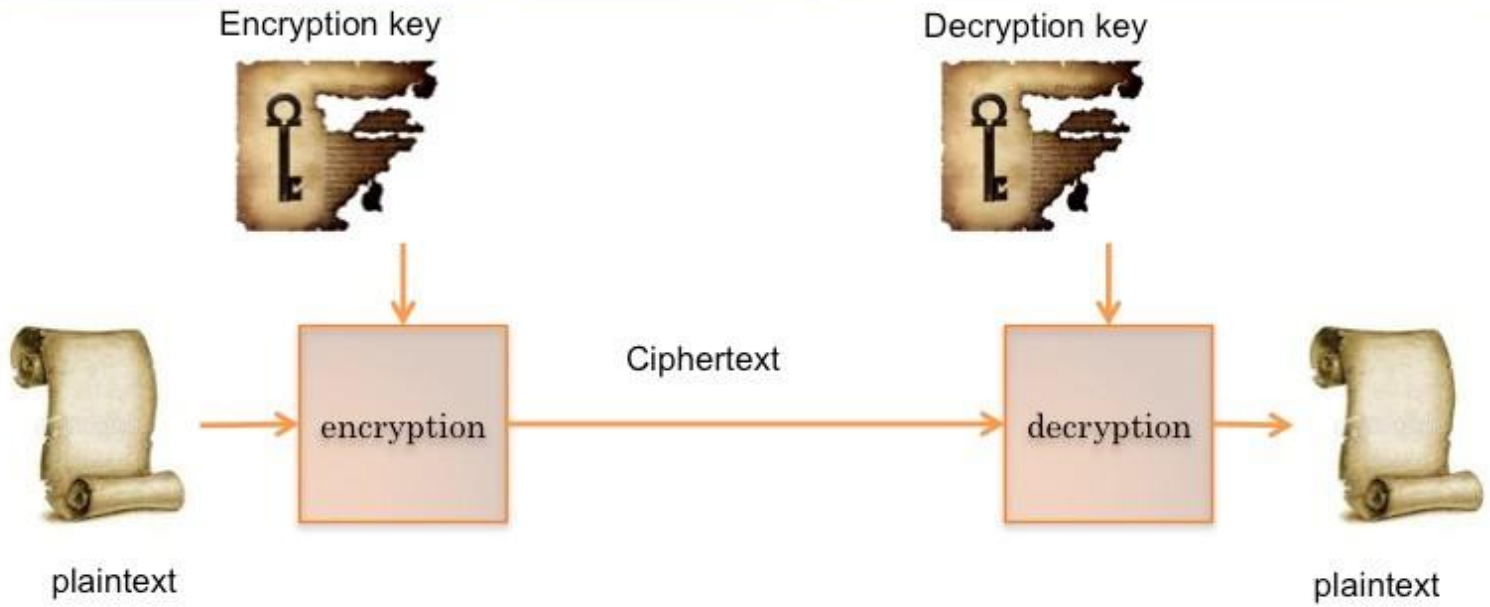


BASIC DEFINITIONS

- **Cryptanalysis**: *the techniques used to decrypt the message without the complete knowledge of the cipher details.*
- **Cryptology= Cryptography + Cryptanalysis**

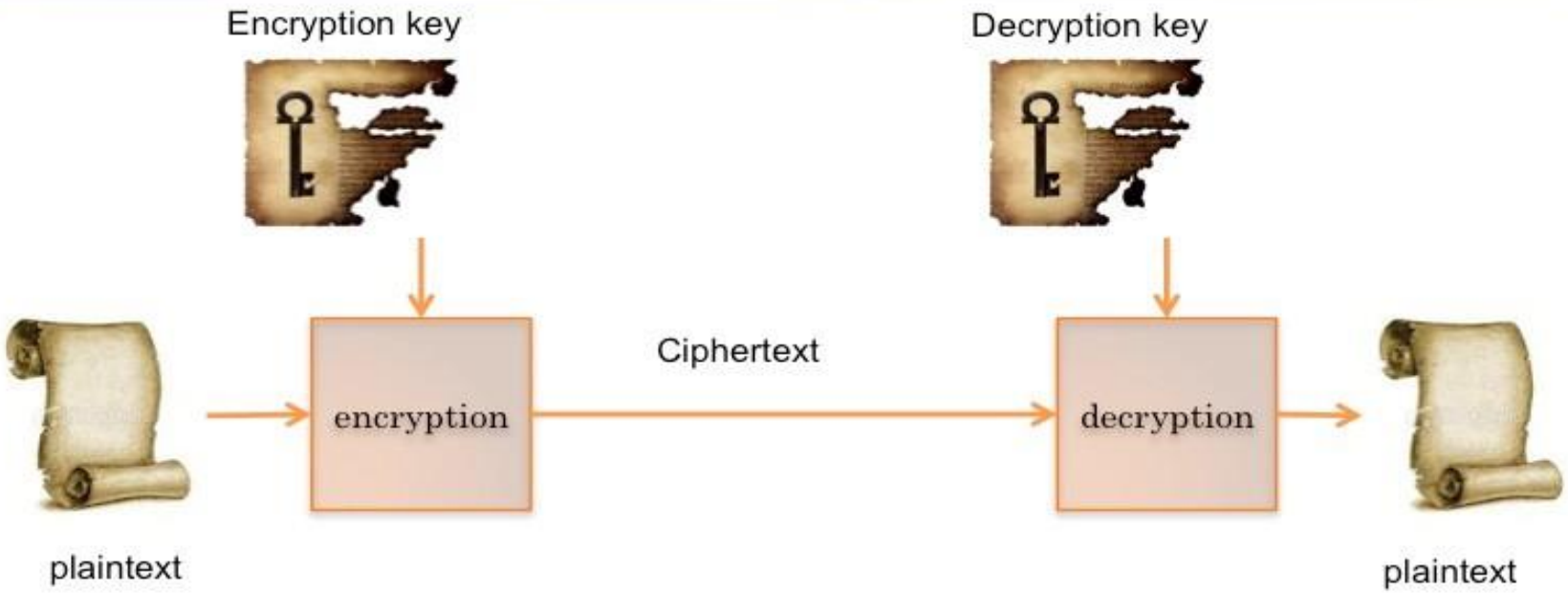
- An instance:
  - **Symmetric encryption** is often referred to as **conventional encryption** or **single key encryption**
  - Let us instantiate our basic definitions on the symmetric encryption

Conventional Encryption Principles



- **Encryption:** performs various substitution and transposition on the plain text
- **Secret key:** it is an input to the encryption algorithm. The exact substitution and transposition performed by the algorithm depend on the key. Sender and receiver have the same key.

Conventional Encryption Principles



- **Ciphertext:** A scrambled message that depends on the plaintext and secret key. For a given message two different keys produce two different ciphertexts.
- **Decryption:** the encryption algorithm running in reverse. It takes the ciphertext and the secret key and produces the original plaintext

CONVENTIONAL ENCRYPTION PRINCIPLES

The encryption algorithm should be public and the key kept secure.

WHY?

We should avoid security by obscurity

Cheap implementations

Breaking the code

- **Brute force attack**
  - All possible keys are tried until the ciphertext can be understood
- **Cryptanalysis**
  - The nature of the encryption is considered together with characteristics of the plaintext or even some plaintext-cipher couples
  - The algorithm is studied in order to deduce the plaintext or the key

AVERAGE TIME REQUIRED FOR EXHAUSTIVE KEY SEARCH

Key Size (bits)	Number of Alternative Keys	Time required at 10^6 Decryption/μs
32	$2^{32} = 4.3 \times 10^9$	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	10 hours
128	$2^{128} = 3.4 \times 10^{38}$	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	5.9×10^{30} years

CONVENTIONAL ENCRYPTION PRINCIPLES

- **Unconditionally secure:** the ciphertext does not contain enough information to determine uniquely the plaintext.

- Encryption algorithms are rarely unconditionally secure then we aim to **computationally secure.**

  - The cost of breaking a cipher exceeds the value of the encrypted information.

  - The time required to break the cipher exceeds the useful lifetime of the information.

CONVENTIONAL ENCRYPTION PRINCIPLES

What are the minimum requirements of an encryption algorithm?

- **Strong algorithm:** an attacker, who knows the algorithms and have access to some ciphertexts, should not be able to decrypt the ciphertexts and obtain the key.

Stronger encryption algorithm: The attacker cannot discover the ciphertext or the key even if he know the some ciphertexts and the related plaintext.

SUBSTITUTION TECHNIQUES

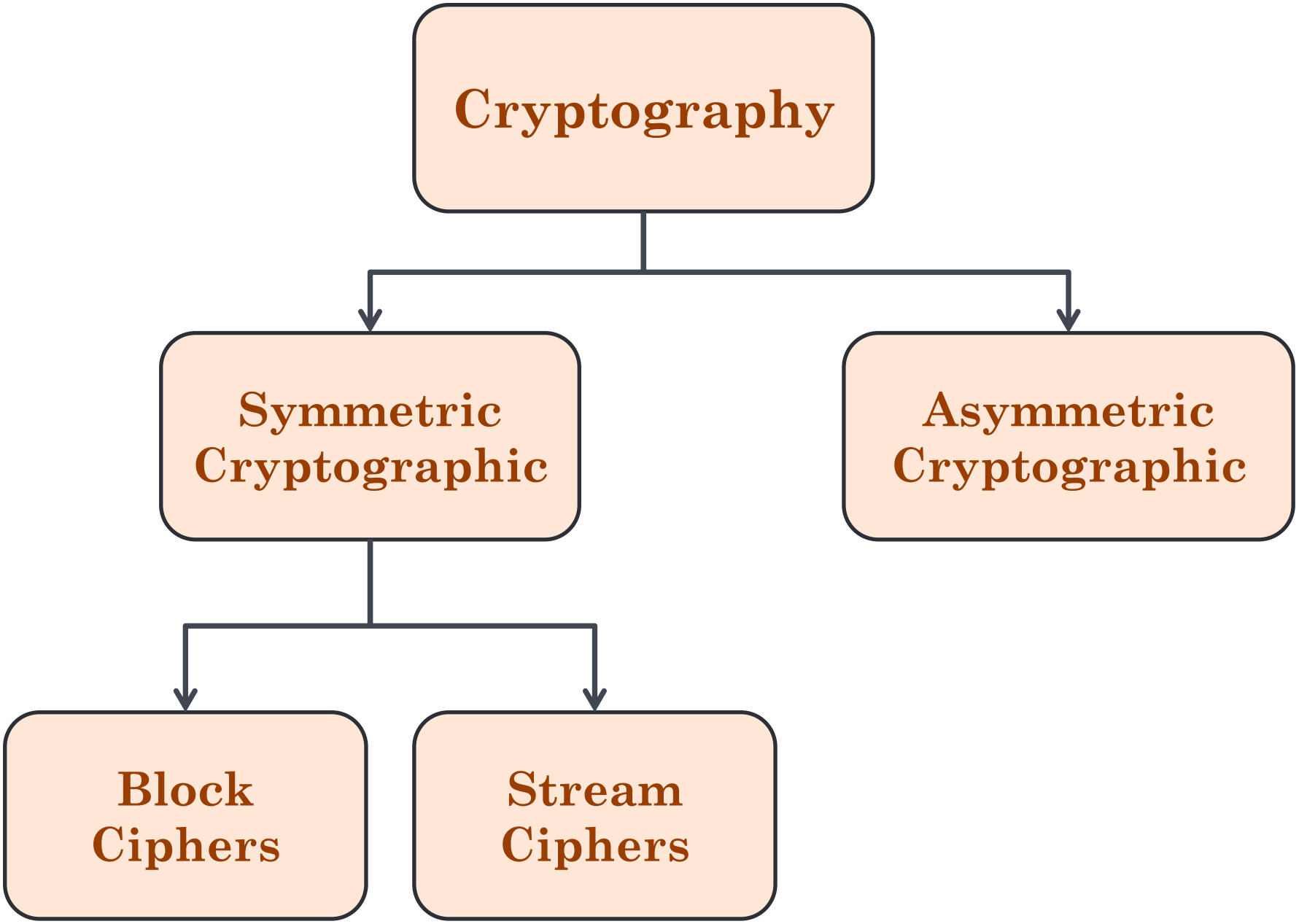
Substitution technique is one in which the letter of plaintext are replaced y other letters or by number s or symbols.

- Let us understand some techniques in nexts slides

Taxonomy of Cryptography

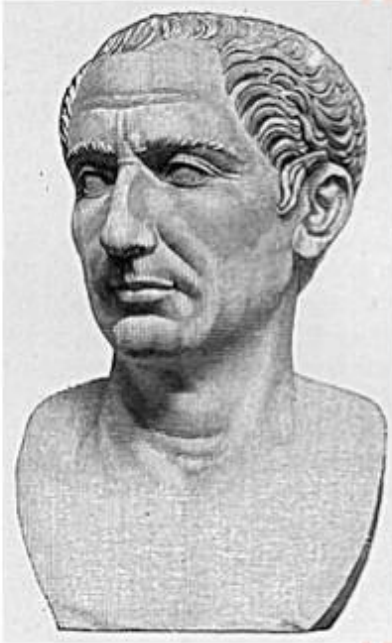
- Modern Cryptography can be divided into two main subfields of study:
- **Symmetric-key** and **Asymmetric-key** cryptography.
- Symmetric-key can be divided into:
- **block ciphers** and **stream ciphers**.
- The next figure depicts the taxonomy of cryptography.

Taxonomy of Cryptography…



CONFIDENTIALLY: CAESAR'S CODE

- The need of concealing a message is very old;
- Julius Caesar (13 July 100 BC – 15 March 44 BC) was a Roman general. He played a critical role in the gradual transformation of the Roman Republic into the Roman Empire.
- Julius Caesar invented a cipher code and used it to protect messages of military significance;



CONFIDENTIALLY: CAESAR'S CODE

- GOAL: Julius wants to send messages to his general through a guard
- THREAT: If the attacker "meets" the guard he can **intercept** the message
- Caesar implemented the **confidentiality service** using the Caesar's code (i.e., **the mechanisms**)

EVERYONE ATTACK NOW

CONFIDENTIALLY: CAESAR'S CODE

- it is a substitution chiper in which a letter in the plain text is replaced by a letter some fixed number of positions down the alphabet

EVERYONE
↓
F
↓
G
↓
HYHUBRQH

ATTACK NOW
↓
B
↓
C
↓
DWWDFN QRZ

number is equal to 3 that is the key

Confidentially: Caesar's code

- The general receives the message and decodes it back
  - Is it easy to break?
  - YES: brute force attack i.e. I try all possible keys (25)
  - The encryption scheme is not computationally secure

HYHUBRQH DWWDFN QRZ

EVERYONE ATTACK NOW

Caesar Cipher

- Then the algorithm can be expressed as follows.
  - For each plaintext letter P, substitute the ciphertext letter C.
- Caesar cipher sometimes called shifted alphabets, which shift the letters of the alphabet to the right by k positions, modulo the size of the alphabet.

$$C=E(3,P)=(P+3) \bmod 26$$

A shift may be of any amount, the general Caesar algorithm is

$$C=E(K,P)=(P+K) \bmod 26$$

Where K takes on a value in the range 1 to 25.

- The decryption algorithm is simply

$$P=D(K,C)=(C-K) \bmod 26$$

Caesar Cipher

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

- If we assign key to 3 in mean we shift the all alphabet 3 times as following

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

- Example:
  - Plain: MEET ME AFTER THE TOGA PARTY
  - Cipher: PHHW PH DIWHU WKH WRJD SDUWB

Brute-Force Cryptanalysis of Caesar Cipher

KEY	PHHW	PH	DIWHU	WKH	WRJD	SDUWB
1	oggv	og	chvgt	vjg	vqic	rctva
2	nffu	nf	bgufs	uif	uphb	qbsuz
3	meet	me	after	the	toga	party
4	ldds	ld	zesdq	sgd	snfz	ozqsx
5	kccr	kc	ydrcp	rfc	rmey	nyprw
6	jbbq	jb	xcqbo	qeb	qldx	mxoqv
7	iaap	ia	wbpan	pda	pkcw	lwnpu
8	hzzo	hz	vaozm	ocz	ojbv	kvmot
9	gyyn	gy	uznyl	nby	niau	julns
10	fxxm	fx	tymxk	max	mhzt	itkmr
11	ewwl	ew	sxlwj	lzw	lgys	hsjlq
12	dvvk	dv	rwkvi	kyv	kfxr	grikp
13	cuuj	cu	qvjuh	jxu	jewq	fqhjo
14	btti	bt	puitg	iwt	idvp	epgin
15	assh	as	othsf	hvs	hcuo	dofhm
16	zrrg	zr	nsgre	gur	gbtn	cnegl
17	yqqf	yq	mrfqd	ftq	fasm	bmdfk
18	xppe	xp	lqepc	esp	ezrl	alcej
19	wood	wo	kpdob	dro	dyqk	zkbdi
20	vnnc	vn	jocna	cqn	cxpj	yjach
21	ummb	um	inbmz	bpm	bwoi	xizbg
22	tlla	tl	hmaly	aol	avnh	whyaf
23	skkz	sk	glzkx	znk	zumg	vgxze
24	rjjy	rj	fkyjw	ymj	ytlf	ufwyd
25	qiix	qi	ejxiv	xli	xske	tevxc

MONOALPHABETIC CIPHER

- Monoalphabetic Cipher improves the Caesar cipher.
- The "cipher" line can be any permutation of the 26 alphabetic characters.

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

EVERYONE ATTACK NOW

QMQSYPZQ CBBCAI ZPU

2.MONOALPHABETIC CIPHERS

- With only 25 possible keys, the Caesar cipher is far from secure.
- A dramatic increase in the key space can be achieved by allowing an arbitrary substitution.
- rather than just shifting the alphabet
- could shuffle the letters arbitrarily
- each plaintext letter maps to a different random ciphertext letter
- hence key is 26 letters long

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

Plaintext: I f we w i s h t o r e p l a c e l e t t e r s

Ciphertext: W I R F R W A J U H Y F T S D V F S F U U F Y A

MONOALPHABETIC CIPHER

- There are about 403291461126605635584000000 possible keys.
- Brute force is not feasible anymore.
- Is that secure?

- We know the nature of the plain text that is English
- Analysts can take advantage of regularities of the language
- The relative frequency of the letters in the ciphertext can be determined and compared to a standard frequency distribution in English

MONOALPHABETIC CIPHERS..

- now have a total of 26! = 4 x 10^{26} keys
- with so many keys, might think is secure

- but would be **!!!WRONG!!!**
- problem is language characteristics

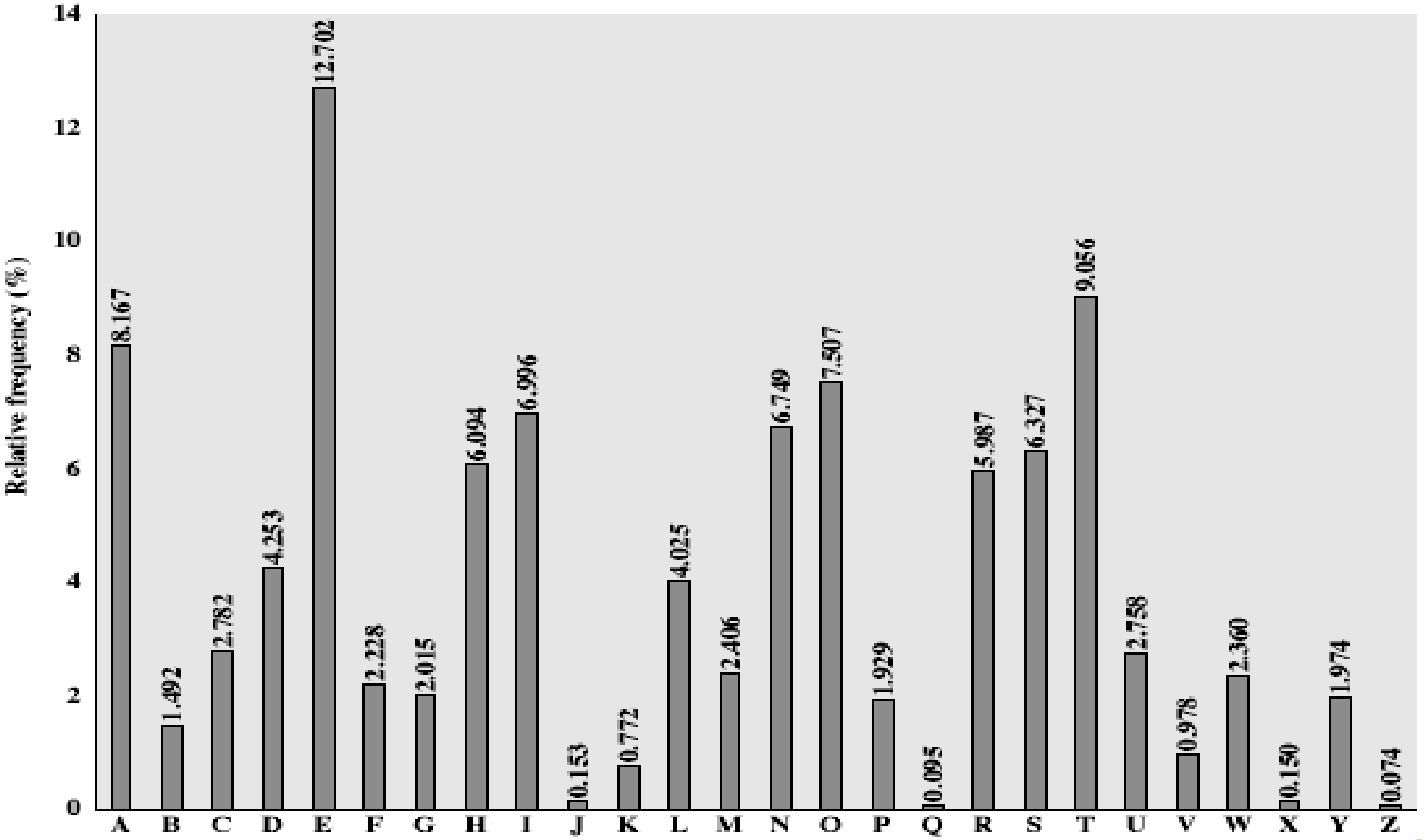
- Weakness of mono-alphabetic cipher

1) Easy to learn
2) Frequency analysis

Language Redundancy and Cryptanalysis

- human languages are **redundant**
- letters are not equally commonly 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
- have tables of single, double & triple letter frequencies for various languages

English Letter Frequencies



MONOALPHABETIC CIPHERS..

- **Example Cryptanalysis**
  - given ciphertext:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXA

IZVUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZ

UHSXEPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

  - Count relative letter frequencies
  - P frequency =15 and Z frequency =14 and so on
  - guess P & Z are e and t
  - guess ZW is th and hence ZWP is the

VIGENERE CIPHER

- The Vigenère Cipher is a polyalphabetic cipher, It consists of the alphabet written out 26 times in different rows,
- each alphabet shifted cyclically to the left compared to the previous alphabet called Vigenère table.

- The strength of this cipher is that there are multiple ciphertext letters for each plaintext letter.

- Vigenère cipher is suspected, then progress depends on determining the length of the keyword.

VIGENERE CIPHER

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

VIGENERE CIPHER

- To encrypt a message, a key is needed that is as long as the message for example:

Plaintext	**T**	**E**	**L**	**L**	**H**	**I**	**M**	**A**	**B**	**O**	**U**	**T**	**M**	**E**
Key	C	A	F	E	C	A	F	E	C	A	F	E	C	A
Cipher text	V	E	Q	P	J	I	R	E	D	A	Z	X	O	E

- The process of encryption is simple:
- Given a key letter *x and a plaintext letter y, the ciphertext letter is at the intersection of the row labeled x*

and the column labeled y; in this case the ciphertext is V.

EXAMPLE

- Decrypt the following message using vigenere cipher algorothm:

Plaintext is : **we are discovered save yourself**

and the key is ***deceptive***,

Solution:

key: *deceptivedeceptivedeceptive*

plaintext: wearediscoveredsaveyourself

ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ

VIGENERE CIPHER

- **Strength of Vivenere.**

1. there are multiple ciphertext letters for each plaintext letter.
2. the letter frequency information is obscured.

Weakness of vigenere:

1. The weakness of a "raw" vigenere cipher is it's tendency to repeat letter patterns at specific intervals .

2. **In other words** (weakness of Vigenère's cipher is the repeating nature of its key).

3. can be attacked with frequency analysis and brute force tests because it reveals some of the mathematical principles that apply in cryptanalysis.

ONE TIME PAD

- The key is as long as the message and is truly random that is with no repetitions:

$$C = P \oplus K$$

$$P = C \oplus K$$

p	q	p ⊕ q
0	0	0
1	0	1
0	1	1
0	0	0

- The key is used once
- The schema is unconditionally secure
  - If the key is random the ciphertext is random thus there no regularities that can be exploited by a cryptanalyst

ONE-TIME PAD

Why is one-time pad used in practice?

- Key distribution
- There are problems to generate a large amount of random numbers

Summary

- Some basic definitions
  - Cipher
  - Breaking the code
  - Cryptology
- **Symmetric cryptography**
- **Substitution techniques**

  - Polyalphabetic Ciphers (Vigenere Cipher)
  - Mono-alphabetic (One time pad)
  - Caesar Cipher

What next

- Vernam or one time pad
- Cryptanalysis of symmetric ciphers
- Cryptography Terminology
- Frequency distribution table
- Kasiski method