

Cryptography - Introduction

What you will learn in this section

- ◆ Overall concept of Information security
- ◆ How cryptography fits into the overall concept
- ◆ History of cryptography
- ◆ Classical cryptographic algorithms
- ◆ Theoretically perfect cryptography
- ◆ Cryptanalysis
- ◆ Work factor for exhaustive key search

The important Questions

- ◆ What are the security requirements in ICT?
- ◆ What methods are available to solve them?
- ◆ Which of these methods use cryptography?
- ◆ How effective are these methods?

An ICT System

- ◆ a collection of hardware, software, data, documents, policies, procedures and human required to do a computing task
- ◆ Therefore, there exists a multiplicity of targets for attacks
- ◆ The most vulnerable being the weakest point

Fundamental maxim



Security is like a chain, it snaps at the weakest link!

This maxim is true everywhere...

Applies everywhere :

- Home
- Organization
- National
- International



Threats, Controls, & Vulnerabilities

- ◆ Vulnerability – a security weakness in IT system or application. It may be due to failures in analysis, design, and implementation
- ◆ Threats - circumstances which have the potential to cause harm or loss especially by exploiting vulnerabilities
- ◆ Controls - protective measures (action, device, procedure, technique) which reduces threats

Main classes of threat

- ◆ **Interruption** - an asset is lost/unavailable/ cannot be utilised. eg. a database is deleted
- ◆ **Interception** - an unauthorised party (person/program) has gained access to an asset. eg. wired/wireless tapping
- ◆ **Modification** - an unauthorised party (person/program) has gained access and tampered around with it. eg. modifying a transaction in transit
- ◆ **Fabrication** - production of counterfeit objects for computing systems. eg. repeating a financial transaction

Principle of Timeliness

“Computer assets need to be protected only until they lose their value”

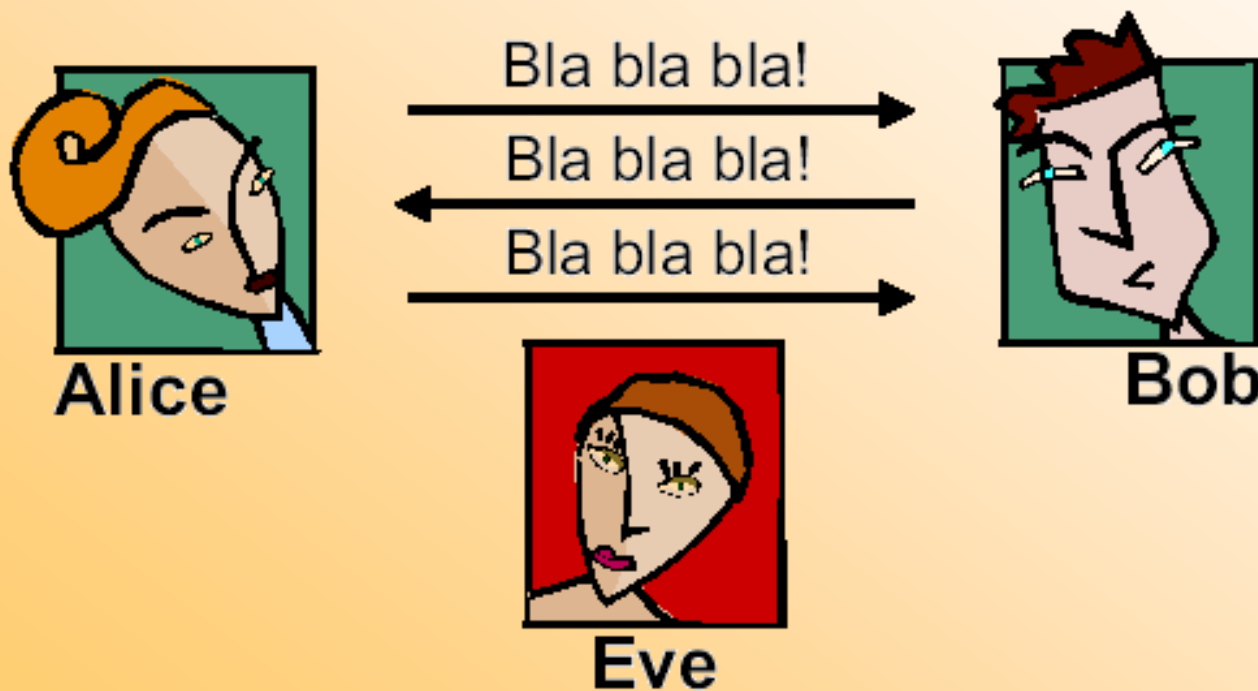
- ◆ Some assets lose their value very quickly, eg. database of news, share indices etc.
- ◆ Some take a long time, eg. military plan, secret chemical formula etc.

Principle requirements of Security

- ◆ **Confidentiality/Secrecy** - assets of computing are accessible only to authorised party. Access means reading, viewing, printing, knowing about.
- ◆ **Integrity** - assets of computing are modified only by authorised party. Modification means writing, changing, changing status, deleting or creating.
- ◆ **Availability** - assets are available to authorised party at the required time and capacity.

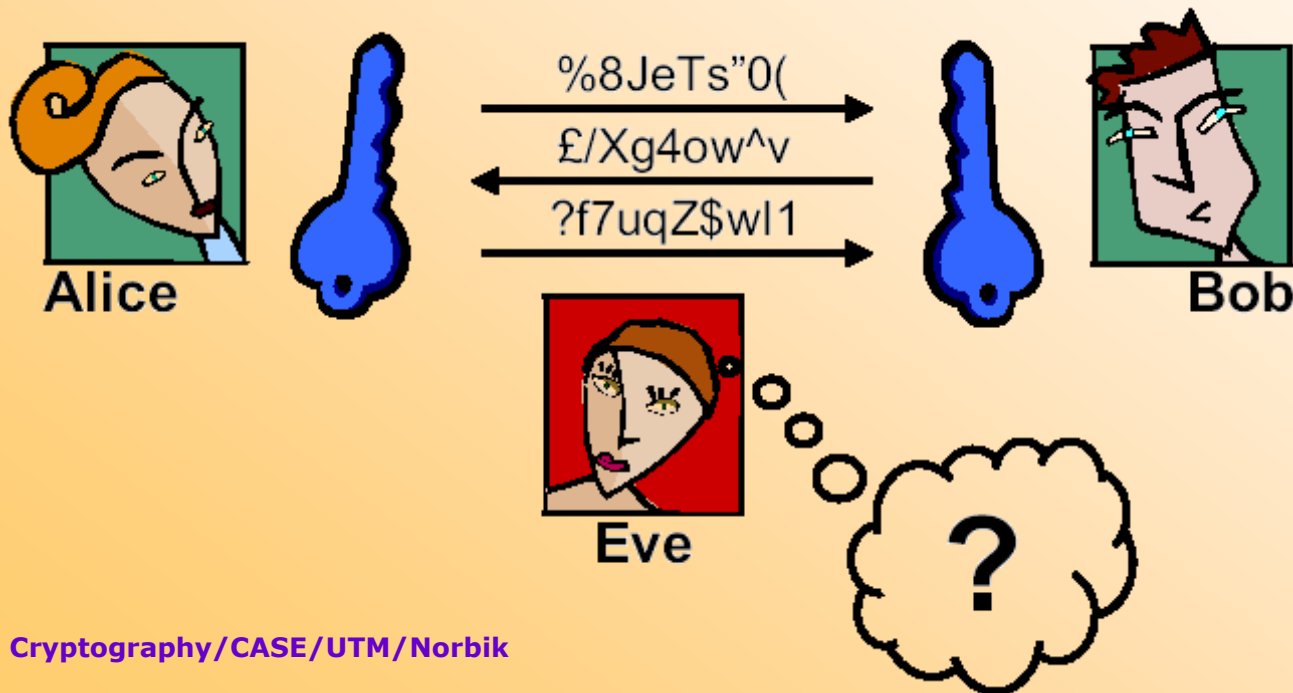
Confidentiality

- ◆ Alice and Bob communicating through insecure channel
- ◆ Eve can listen and understand what is being said



How Encryption helps...

- ◆ The insecure channel is secured by encrypting the communication
- ◆ Eve can listen but cannot understand what is being said



Integrity

- ◆ An unprotected message is susceptible to a modification that cannot be detected easily
- ◆ Eve, here, changes the time that Alice wants meet Bob
- ◆ Bob will not know that the time was modified



Alice

Let's meet at 20h



Eve

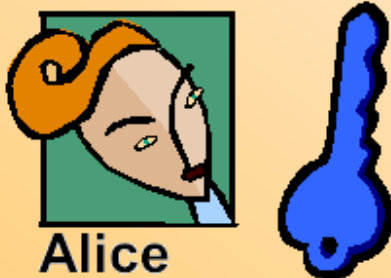


Bob

Integrity (cont.)

- ◆ Any modification to a protected message can be detected easily
- ◆ Bob will know that there is something wrong with message he has received

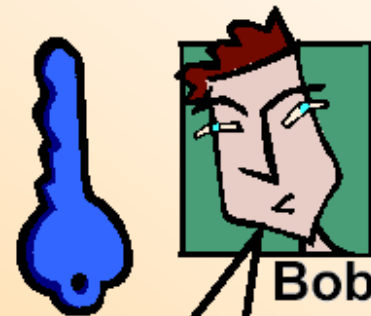
Let's meet at 02h



Let's meet at 20h



Eve



Something is wrong!

Other requirements

- ◆ **Authentication** an entity can be verified to be who he claims to be.
- ◆ **Accountability** entity can be verified to have been responsible/owned an activity/event.
- ◆ **Non-repudiation** achieved when the requirement for accountability is fulfilled.

Authentication

- ◆ The message says that it is from Alice, but there is no proof that the message is from Alice



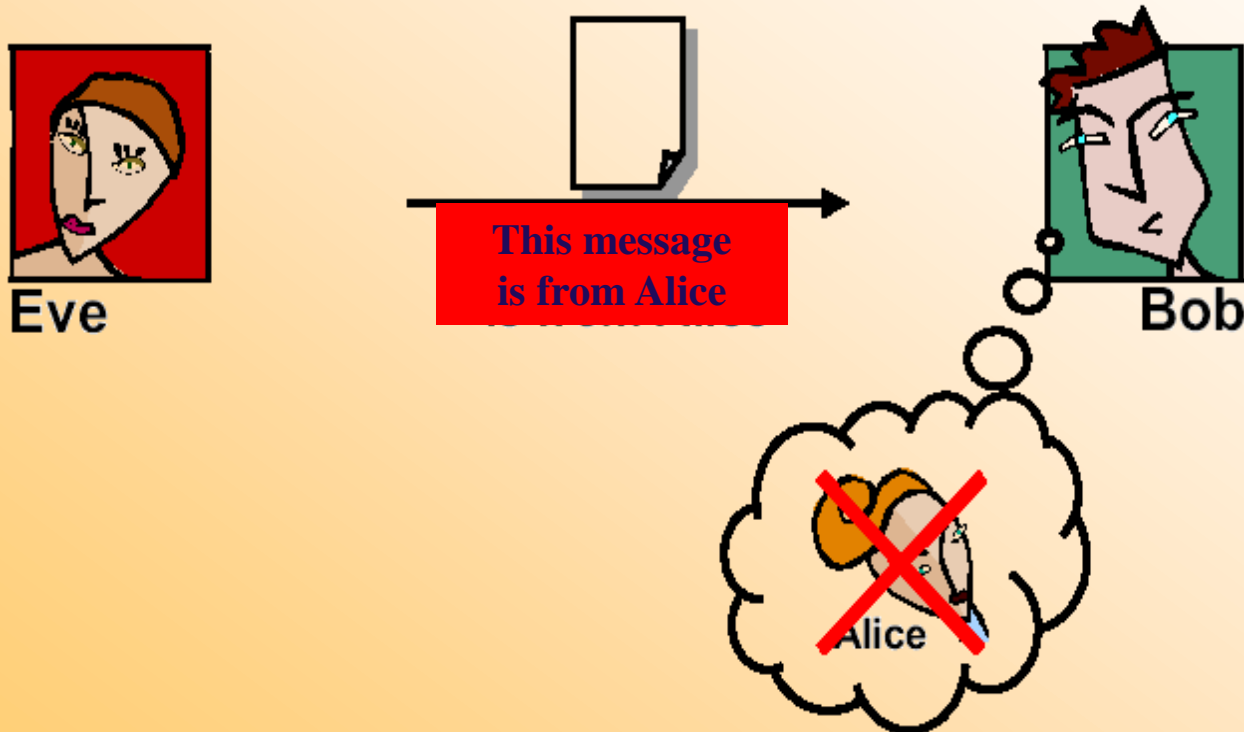
Alice



Bob

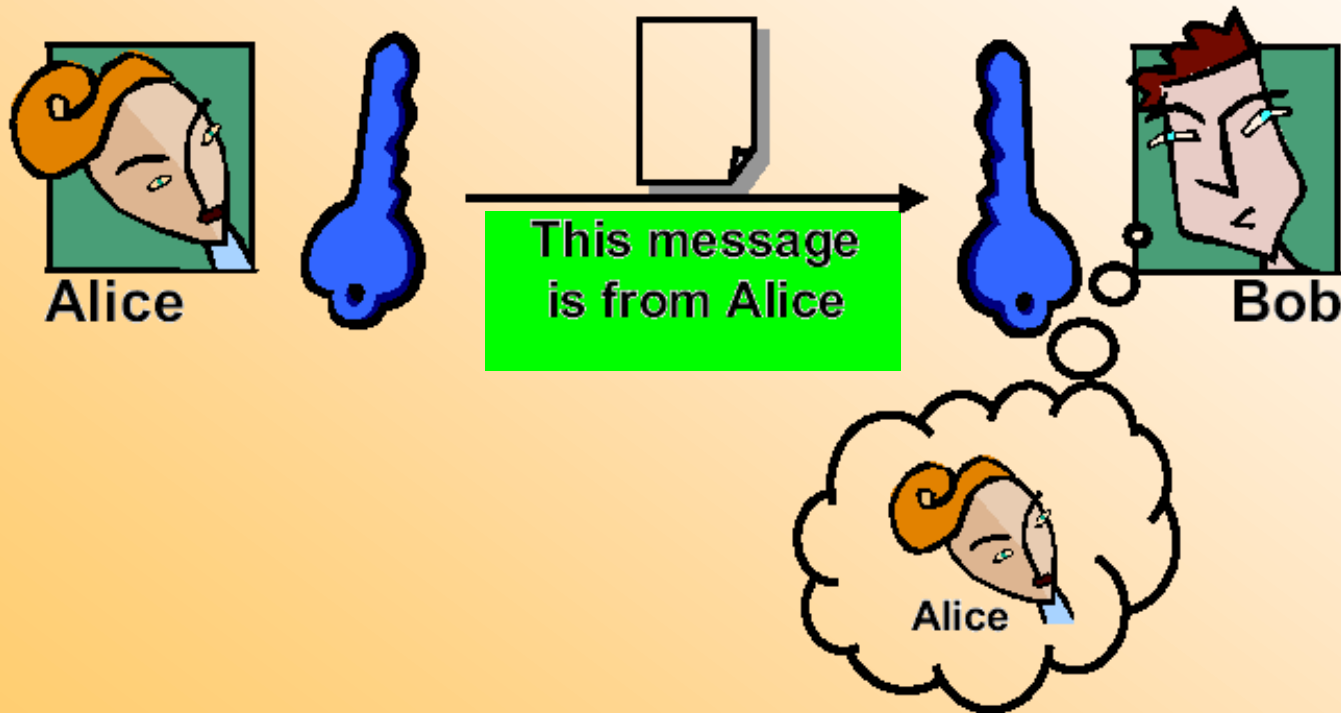
Authentication (cont.)

- ◆ Eve can impersonate Alice and send messages to Bob
- ◆ Bob still cannot prove that the message is not from Alice



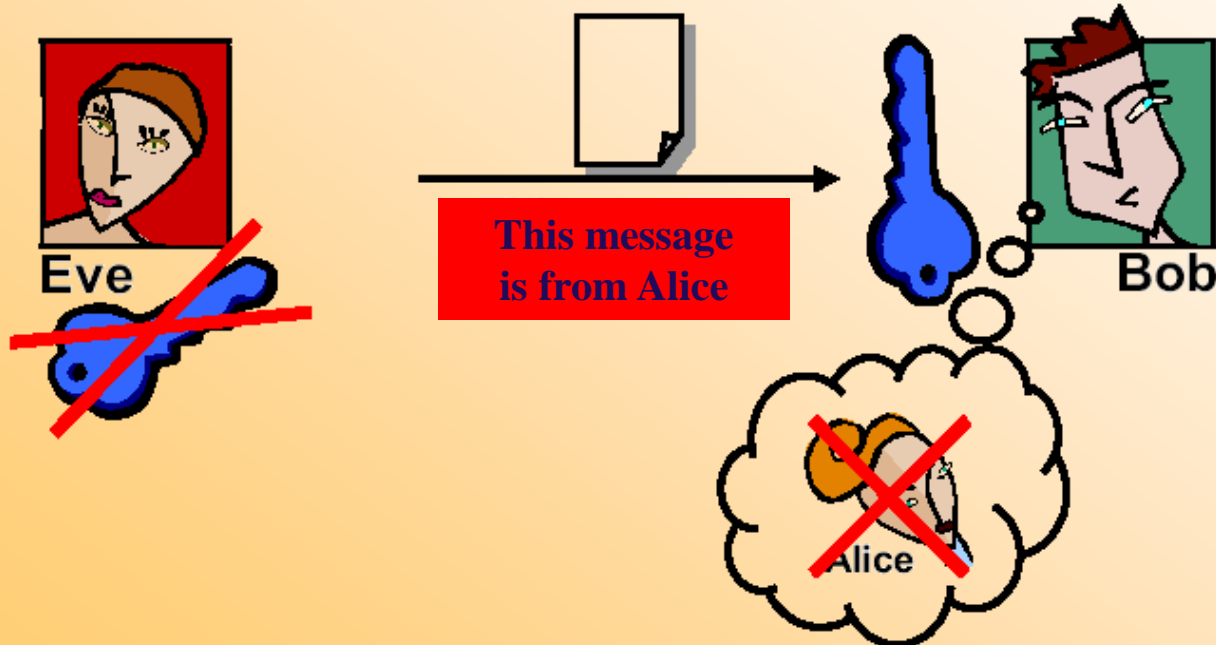
How encryption helps...

- ◆ Bob can verify that the message is from Alice since the key is shared only with her



How encryption helps...

- ◆ Eve cannot impersonate Alice since the key she uses is different from the key that Bob shares with Alice
- ◆ Bob will find easily that the message is not from Alice!



Quantifying security

- ◆ Quantifying security is not easy. Preparing for security is usually done through a technique called Risk Analysis.
- ◆ For ICT, the most common approach is to define it in terms of the risks to the 'CIA':
 - Confidentiality,
 - Integrity,
 - Availability.
- ◆ In Malaysia, Government issued HiLRA & MyRAM (Dec 2005)

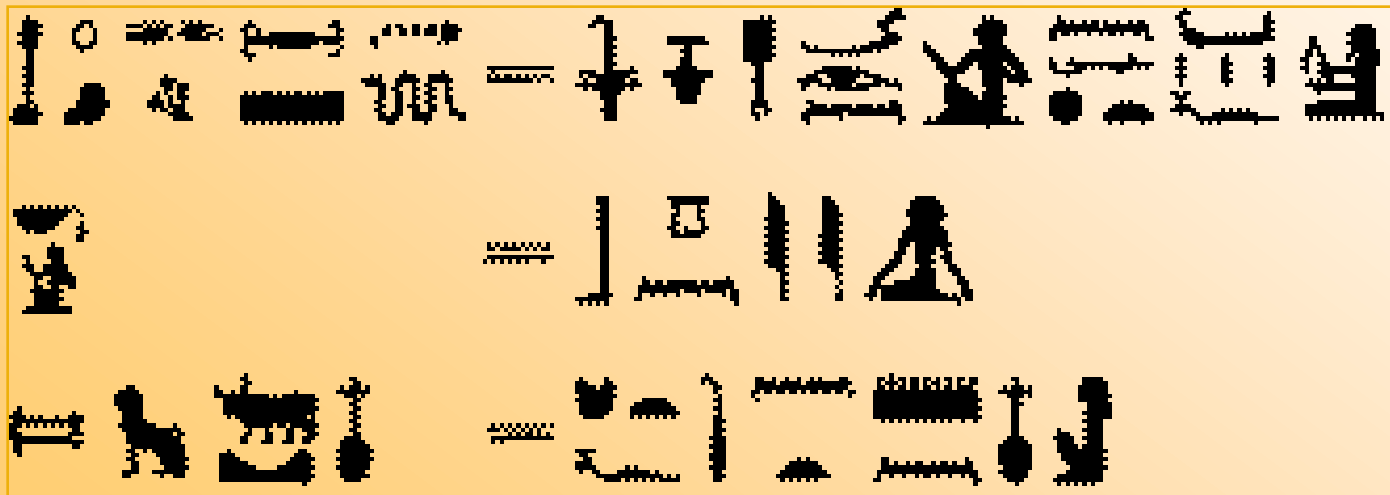
Functionality versus Assurance

- ◆ In assessing secure systems, two different aspects need considering:
 - *functionality*, i.e. what security facilities are provided, and
 - *assurance*, i.e. what guarantees are offered that the security functionality performs as claimed.
- ◆ These two aspects are reflected in the various evaluation criteria eg. the Common Criteria on crypto products

History/evolution of Cryptography

History

- ◆ Cryptography is very old
- ◆ Ancient Ciphers have a history of some 4000 years
- ◆ Ancient Egyptians encoded some hieroglyphic writings on monuments
- ◆ 2000 years ago Julius Caesar used a simple substitution cipher, now known as the Caesar cipher



Cryptography (cont.)

- ◆ Is the study of **secret** (crypto-) **writing** (-graphy)
- ◆ Is a broad subject which requires knowledge of several areas of mathematics such as general algebra, linear algebra, probability theory, complexity theory and information theory
- ◆ Concerned with developing algorithms which may be used to:
 - Conceal the context of some messages from all except the sender and recipient (**privacy** or **secrecy**), and/or
 - Verify the correctness of a message to the recipient (**authentication** or **integrity**)
- ◆ Deals with aspects of secure messaging, authentication, digital signatures, electronic money, and other applications

Cryptography (cont.)

- ◆ Cryptographic systems are generically classified along three independent dimensions:
 1. The type of operations used for transforming **plaintext** to **ciphertext** (encryption)
 - All encryption algorithms are based on two basic operations:
 - **Substitution**: each element in the plaintext (bit, letter, group of bits or letters) is mapped into another element in the ciphertext
 - **Transposition**: elements in the plaintext are rearranged
 - here, all operations must be reversible and no information is lost

Cryptography (cont.)

2. The way in which the plaintext is processed

- A **block cipher** processes the input one block of elements at a time, producing an output block for each input block
- A **stream cipher** processes the input elements continuously, producing output one element at a time, like a stream

3. The number of keys used

- If both sender and receiver use the same key, the system is referred to as symmetric, single-key, **secret-key**, or conventional encryption
- If sender and receiver each uses a different key, the system is referred to as asymmetric, two-key, or **public-key** encryption

Basic Terminology

- ◆ **Cryptography** is the art or science of keeping messages secret, and people who do cryptography are **cryptographers**
- ◆ **Cryptanalysis** is the art of **breaking** ciphers, i.e. retrieving the plaintext without knowing the **secret key**, and practitioners of cryptanalysis are **cryptanalysts**
- ◆ **Cryptology** is the branch of mathematics that studies the mathematical foundations of cryptographic methods

Basic Terminology (cont.)

- ◆ The original message is called **plaintext** or **cleartext**
- ◆ Encoding the message in such a way that hides its contents from outsiders is called **encryption**
- ◆ The encrypted message is called the **ciphertext**
- ◆ The process of retrieving the plaintext from the ciphertext is called **decryption**
- ◆ Encryption and decryption usually make use of a **key**, and the coding method is such that decryption can be performed only by knowing the **secret key**

Basic Cryptographic Algorithms

- ◆ All modern algorithms use a **key** to control encryption and decryption; a message can be decrypted only if the key matches the encryption key
- ◆ There are two classes of key-based encryption algorithms, **symmetric** (or **secret-key**) and **asymmetric** (or **public-key**) algorithms
- ◆ Symmetric algorithms use the **same key** for encryption and decryption
- ◆ Asymmetric algorithms use **different keys** for encryption and decryption

Basic Cryptographic Algorithms (cont.)

- ◆ Symmetric algorithms can be divided into stream ciphers and block ciphers
- ◆ Stream ciphers encrypt a single bit of plaintext at a time
- ◆ Block ciphers take a number of bits and encrypt them as a single unit
- ◆ Asymmetric ciphers (also called public-key algorithms) permit the encryption key to be public (it can even be published in a newspaper), allowing anyone to encrypt with the key, whereas only the proper recipient (who knows the decryption key) can decrypt the message
 - The encryption key is also called the public key and the decryption key the private key or secret key

qANQR1
DBwU4D
zyyrDy8
V58MQC
AC37dkn
ScAc/1rq
pwYcEW
sTO/Cv/o
F9KGi/cE
ISZz1+R
LSFju8E
X8e8Fs42
DV59OV
TdoG52r
CECK3a9
By4G8X
VYzzzok
KSyXbA
p5YsxBK
ebvLYBK
pBz4p8Er
esJkMJpn
llkZdWhq
iRgzxCey
/wUTTv19
cDagP++
X4YVzy
DFjifTR1
7T7+F8w
BeY/VLp
5XVoE7k
b9XDsk
W0v6PM
n5moJ+2
UL/w4n7
Dd8rv5hP
GTIKT2i
LV3xjrzH
v9AJW/O
XBLVxC
Fp1iBdU
+LkoxmV
uvxYv6W
B5y0R5U
Zll7dp2a
AaGYQzj
YT58TAz
t23G2i9b
QQ1SGE
2o5njXf7
0V3RaHi
Dt/j/K8O
B/0WHH
DQwmsy
4x3xL1lr
vgGeQh7

Classical Cryptography

Classical Encryption Techniques

- ◆ The two basic building blocks of all encryption techniques:
 - **Substitution**
 - Each element in the plaintext (bit, letter, group of bits or letters) is mapped into another element in the ciphertext
 - The earliest known use of a substitution cipher, and the simplest, was by *Julius Caesar*
 - **Transposition**
 - Elements in the plaintext are rearranged by performing some sort of permutation on the plaintext letters
 - The simplest such cipher is the *rail fence* technique

Classic Techniques (cont.)

- ◆ Systems, referred to as **product systems**, involve multiple stages of substitutions and transpositions
- ◆ These techniques may be:
 - **Monoalphabetic** - only one substitution / transposition is used, or
 - **Polyalphabetic** - where several substitutions / transpositions are used

Substitution Ciphers

- ◆ Simple substitution cipher:

- ◆ $a = p, \quad b = m, \quad c = f, \dots$

- ◆ Polyalphabetic substitution cipher

- ◆ $a = p, \quad b = m, \quad c = f, \dots$

- ◆ $a = l, \quad b = t, \quad c = a, \dots$

- ◆ $a = f, \quad b = x, \quad c = p, \dots$

Caesar Cipher

- ◆ Each letter of the alphabet is replaced with the letter standing three places further down the alphabet ($\text{letter} = \text{letter} + 3$), and the alphabet is wrapped around, so that the letter following Z is A

- Plain alphabet: ABCDEFGHIJKLMNOPQRSTUVWXYZ
- Cipher alphabet: DEFGHIJKLMNOPQRSTUVWXYZABC

- ◆ Example:

- plaintext: I CAME I SAW I CONQUERED
- Ciphertext: L FDPH L VDZ L FRQTXHUHG

General Caesar Cipher

- ◆ A shift may be any amount from 1 to 25 (i.e., replace each letter of message by a letter of fixed distance away)
- ◆ Can be described mathematically as a function by assigning a numerical equivalent to each letter (A = 0, B = 1, C = 2,, Y = 24, Z = 25); then:
 - Encryption : $C = E(P) = (P + k) \bmod 26$
 - Decryption : $P = D(C) = (C - k) \bmod 26$
 - Where P : plaintext, C : ciphertext, E : encryption, D : decryption, and k : key which takes on a value in the range 0 – 25
- ◆ Any cipher using a simple letter shift is called a *Caesar Cipher*, not just those with shift 3

Breaking Caesar Cipher

◆ Simply try all the 25 possible keys

- e.g. try to recover the original message from the following cipher:

cipher: **PHHW PH DIWHU WKH SDUWB**

- Try all possible keys (1, 2, ..., 25)

Key	PHHW PH DIWHU WKH SDUWB
1	oggv og chvgt vjg rtcva
2	nffu nf bgufs uif qbsuz
4	ldds ld zesdq sgd ozqsx
...
25	qiix qi ejxiv xli tevxc

Breaking Caesar Cipher (cont.)

- ◆ The following three important characteristics of the above problem enabled us to use a brute-force cryptanalysis:
 - the encryption and decryption algorithms are known
 - there are only 25 keys to try
 - the language of the plaintext is known and easily recognizable
- ◆ Today, what makes the brute-force cryptanalysis difficult/impractical is the use of an algorithm that employs a large number of keys
 - For example, the DES algorithm makes use of a 56-bit key, giving a key space of 2^{56} or greater than 7×10^{16} possible key

Mixed Monoalphabetic Substitution Cipher

- ◆ Shuffle the letters arbitrarily each plaintext letter maps to a different random ciphertext letter or even to 26 arbitrary symbols. Hence, key is 26 letters long
- ◆ This is known as a Mixed **Monoalphabetic Substitution Cipher = Many Many**
- ◆ e.g.,
 - Plain: **ABCDEFGHIJKLMNOPQRSTUVWXYZ**
 - Cipher: **DKVQFIBJWPESCXHTMYAUOLRGZN**
 - For example:
 - Plaintext: **IFWEWISHTOREPLACELETTERS**
 - Ciphertext: **WIRFRWAJUHYFTSDVFSFUUFYA**

Easier Monoalphabetic Substitution Ciphers

- ◆ The Mixed Monoalphabetic Cipher has a 26 letter key
- ◆ Several possibilities, one of the easiest is write keyword(s) without duplicate letters, then rest of alphabet following last key letter
- ◆ e.g., given a key “**JULIUSCAESAR**” = “**JULISCAER**”
 - Plain: **ABCDEFGHIJKLMNOPQRSTUVWXYZ**
 - Cipher: **JULISCAERTVWXYZBDFGHKMNOPQ**

General Monoalphabetic Cipher

- ◆ e.g., given the keyword “**STARWARS**”
- ◆ Remove repeated letters to get “**STARW**”
- ◆ Write out and fill in remaining letters as follows:

STARW BCDEF GHIJK LMNOP QUVXY Z →

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

- ◆ Then read off by columns to get the translation alphabet

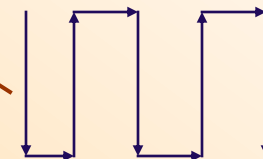
Plaintext: **ABCDEFGHIJKLMNOPQRSTUVWXYZ**

Ciphertext: **SBGLQZTCHMUADINVREJOXWFKPY**

- ◆ Can then use this to encrypt/decrypt, e.g.

Plaintext: **I KNOW ONLY THAT I KNOW NOTHING**

Ciphertext: **H UINF NIAP OCSO H UINF INOCHIT**



History of Monoalphabetic Ciphers

- ◆ Have a total of $26! \cong 4 \times 10^{26}$ keys. With so many keys, for a long time this was thought secure
- ◆ The method of breaking it using frequency analysis was discovered by Arab scientists. The earliest known description is in Abu al-Kindi's "*A Manuscript on Deciphering Cryptographic Messages*", published in the 9th century
- ◆ Stronger monoalphabetic ciphers with several replacement symbols for each letter, for common words, and nulls were developed in middle ages
- ◆ Eventually all were susceptible to analysis

History of the Monoalphabetic (cont.)

- ◆ The main practical problem with a Monoalphabetic Substitution Cipher is remembering all 26 letters/symbols in the key
- ◆ To make it easier, schemes were devised to build the shuffled alphabet from a keyword(s)
- ◆ The disadvantage of any such scheme is that it reduces the number of keys, though the total number is still very large

Polyalphabetic Substitution Ciphers

- ◆ To improve security of the simple monoalphabetic technique use **many monoalphabetic** substitution alphabets where each letter can be replaced by many others (this is called **polyalphabetic**)
- ◆ Use a key to select which alphabet is used for each letter of the message (i^{th} letter of key specifies i^{th} alphabet to use)
- ◆ Use each alphabet in turn and repeat from start after end of key is reached
- ◆ Blaise de Vigenère is generally credited as the inventor of the “polyalphabetic substitution cipher”

Vigenère Cipher

- ◆ To use Vigenère scheme, first write the plaintext and under it write the keyword repeated
- ◆ Using each key letter in turn as a *Caesar Cipher* key, encrypt the corresponding plaintext letter (encryption and decryption follow the same process)
- ◆ e.g., using the keyword ‘**CIPHER**’ as a key:
 - Plaintext: **THISPROCESSCANALSOBEEEXPRESSED**
 - Keyword: **CIPHERCIPHERCIPHERCIPHERCIPHE**
 - Ciphertext: **VPXZTIQKTZWTCVPSWFDMTETIG AHLH**

Vigenère (cont.)

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

Vigenère: Example

- ◆ Plaintext: THISPROCESSCANALSOBEEEXPRESSED
- ◆ Using the keyword “**CIPHER**”, we have the following translation alphabets:

	ABCDEFGHIJKLMNOPQRSTUVWXYZ
– C	CDEFGHIJKLMNOPQRSTUVWXYZAB
– I	IJKLMNOPQRSTUVWXYZABCDEFGHI
– P	PQRSTUVWXYZABCDEFGHIJKLMNO
– H	HJKLMNOPQRSTUVWXYZABCDEFGHI
– E	EFGHIJKLMNOPQRSTUVWXYZABCD
– R	RSTUVWXYZABCDEFGHIJKLMNO

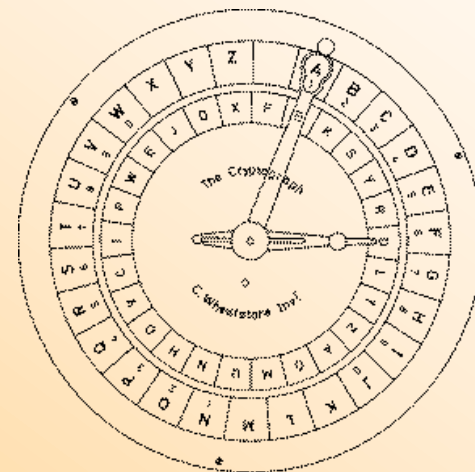
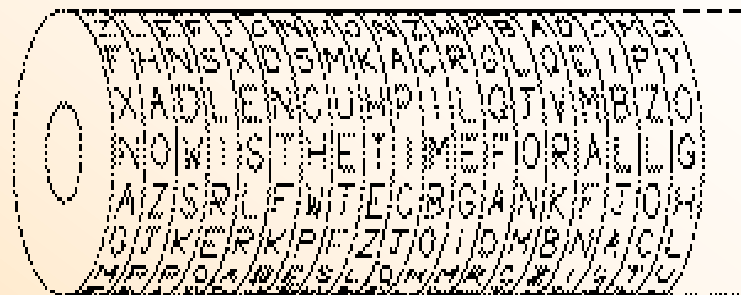
Vigenère: Example (cont.)

◆ Hence, map the plaintext message as follows:

- ‘T’ uses key ‘C’ maps to ‘V’
- ‘H’ uses key ‘I’ maps to ‘P’
- ‘I’ uses key ‘P’ maps to ‘X’ etc.
- Ciphertext: **VPXZTIQKTZWTCVPSWFDMTETIG AHLH**

Ciphers Machines

- ◆ To ease tedium of encrypting/decrypting, mechanical devices were developed
- ◆ **Jefferson cylinder**, developed in 1790s comprising 36 disks each with a random alphabet
- ◆ **Wheatstone disc**, by Wadsworth in 1817, and Wheatstone in 1860's, comprised two concentric wheels to generate a polyalphabetic cipher



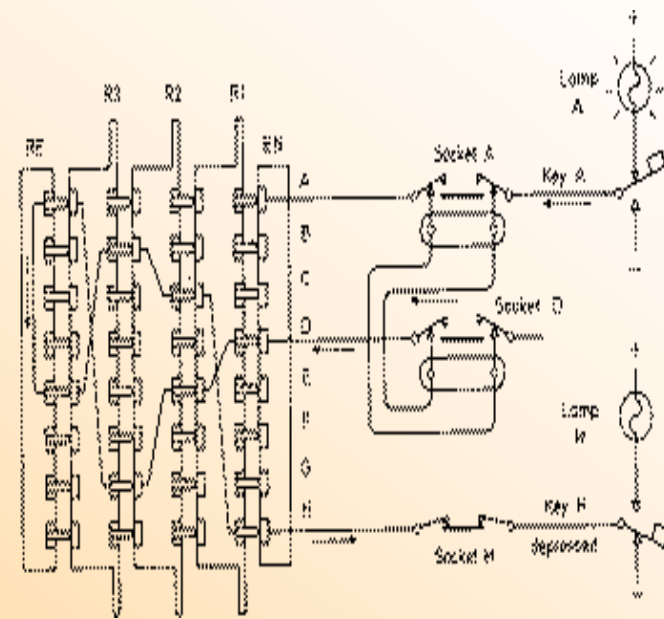
Ciphers Machines (cont.)

- ◆ The next major advance in ciphers required the use of mechanical cipher machines
- ◆ These enabled use of complex varying substitutions
- ◆ Extensively used in the WWII
- ◆ These included the German Enigma, the Swedish Hagelin (right), and the Japanese Purple



Ciphers Machines Internals

- ◆ The **Enigma Rotor machine** was a substitution using a continuously changing alphabet
- ◆ Achieved by linking: a keyboard, 3 cipher wheels substituting one letter for another, which rotated after each use, a reflector which bounced the signal back through the 3 wheels & plug-board and output lamps



Transposition Ciphers

- ◆ The plaintext remains the same, but the order of characters is shuffled around (e.g., shuffle *secret* to *etcrse*)
- ◆ Classically, arrangement was done with the aid of some type of geometric figure, usually 2-dimensional array matrix
- ◆ Transposition is not a permutation of alphabet characters but a permutation of places
 - Letters remain their identity but lose their position
 - There is a permutation of the plaintext letters

Rail Fence

- ◆ Simplest form of the transposition techniques
- ◆ Plaintext is written down as a sequence of diagonals and then read off as a sequence of rows to give the ciphertext
- ◆ e.g.,

Plaintext: **MERCHANT TAYLORS' SCHOOL**

Arrangement:

	M		R		H		N		T		Y		O		S		C		O		L
		E		C		A		T		A		L		R		S		H		O	

Ciphertext: **MRHNTYOSCOLECATALRSHO**

Columnar Transposition Cipher

- ◆ **Encryption:** plaintext is written horizontally onto the matrix of fixed width and the ciphertext is read off vertically
- ◆ **Decryption:** ciphertext is written vertically onto the same matrix of identical width and then reading the plaintext off horizontally
- ◆ e.g.: Plaintext is **RENAISSANCE** is written into a 3x4 matrix as:

R	E	N	A
I	S	S	A
N	C	E	

- ◆ The column taken-off in the order 2-4-1-3, the resulting cipher text is

ESCAARINNSE

Complexity of Transposition Cipher

- ◆ Involves no additional work beyond arranging the letters and reading off again
- ◆ The algorithm is constant in the amount of work per character
- ◆ The time for the algorithm is proportional to the length of the message
- ◆ The algorithm requires space for all characters of the message and therefore it depends directly on the length of the message

Complexity of Transposition Cipher (cont.)

- ◆ Output characters cannot be produced until all characters of the message have been read
- ◆ The delay of the algorithm depends on the length of the message
- ◆ Due to storage and delay, this algorithm is not appropriate for long messages

Product Ciphers – ADFGVX Cipher

- ◆ ADFGVX cipher is one of the most famous fractionation system
 - Substitution is first made from symbols in the plaintext to multiple symbols (usually pairs in the ciphertext, which is then super-encrypted by a transposition
- ◆ It uses a 6x6 matrix to substitution-encrypt the 26 letter of the alphabet and 10 digits into pairs of the symbols A,D,F,G,V,X
- ◆ The resulting cipher is only an intermediate cipher, it is then written into a rectangular matrix and transposed to produce the final cipher which is the one which would be transmitted

Example : *ADFGVX* Cipher

- ◆ Encrypting the phrase “Merchant Taylors” using the keyword “**subject**”

- ◆ 1st stage:

	A	D	F	G	V	X
A	S	U	B	J	E	C
D	T	A	D	F	G	H
F	I	K	L	M	N	O
G	P	Q	R	V	W	X
V	Y	Z	0	1	2	3
X	4	5	6	7	8	9

Plaintext: M E R C H A N T T A Y L O R S

Ciphertext: FG AV GF AX DX DD FV DA DA DD VA FF FX GF AA

Example : *ADFGVX* Cipher (cont.)

◆ 2nd stage

- Intermediate ciphertext can then be put in a transposition matrix based on a different key

F	G	A	V	G	F
A	X	D	X	D	D
F	V	D	A	D	A
D	D	V	A	F	F
F	X	G	F	A	A

- The final cipher is (read columns 1-5-4-2-3-6) :

FAFDF GDDFA VXAAF GXVDX ADDVG FDAFA

Vernam Cipher

- ◆ Involves arbitrarily long nonrepeating sequence of numbers that are combined with the plaintext
- ◆ Vernam uses 2 paper tape readers, one for the message and the other for the key
 - The tape contain random numbers which are combined with the characters
- ◆ Sequence of random numbers are nonrepeating and each tape is used only once
- ◆ Immune to cryptanalytic attack because the available ciphertext does not display the bit pattern of the key
- ◆ Marked the beginning of modern cryptography

Vernam Cipher : Example

- ◆ Assume that the alphabetic letters are combined by sum mod 26 with a stream of random two-digit numbers

Plaintext	V	E	R	N	A	M	C	I	P	H	E	R
numerical value	21	4	17	13	0	12	2	8	15	7	4	17
+ random value	76	48	16	82	44	3	58	11	60	5	48	88
= sum	97	52	33	95	44	15	60	19	75	12	52	105
mod 26	19	0	7	17	18	15	8	19	23	12	0	1
ciphertext	T	A	H	R	S	P	I	T	X	M	A	B

Binary Vernam Cipher

- ◆ In order to encrypt a binary string, random binary digits can be combined with bits from the plaintext binary string

- ◆ Example:

plaintext	1	0	1	1	0	1	1	0	0	1	1	0	1	0
random	0	0	0	1	1	1	0	0	0	1	0	0	1	0
ciphertext	1	0	1	0	1	0	1	0	0	0	1	0	0	0

- ◆ Binary addition mod 2 can be calculated by the bitwise XOR

One-Time Pad (OTP)

- ◆ A version of the Vernam Cipher
- ◆ Produces random output that bears no statistical relationship to the plaintext
 - Secure against frequency analysis
- ◆ The most secure cryptographic algorithm
 - Unbreakable provided
 - Pad is never reused
 - Unpredictable random numbers are used
 - Ultimate in security
 - Perfect encryption scheme

One-Time Pad (cont.)

◆ Example:

Message:

S E C R E T

Number :

18	5	3	17	5	19
----	---	---	----	---	----

OTP :

+

15	8	1	12	19	5
----	---	---	----	----	---

Mode 26 :

7	13	4	3	24	24
---	----	---	---	----	----

Ciphertext:

g m d d x x

Theoretical One-Time Pad

- ◆ Systems using perfect random, non-repeating keys which is endless and senseless
- ◆ A random key sequence added to a nonrandom plaintext produces a completely random ciphertext
 - No amount of computing power can solve it
 - The key must be random for OTP to be totally secure (pseudo-random has nonrandom properties and does not ensure total security)
- ◆ Random key used once, and only once

Theoretical One-Time Pad

◆ Unbreakable in theory

- The key neither repeats, nor recurs, nor makes sense, nor erects internal frameworks
- Perfect randomness nullifies any horizontal or lengthwise cohesion
- The only unbreakable cryptography system

One-Time Pad : Problems

- ◆ Why wouldn't it be used today?
 - The need to be able to destroy the key after every use
 - Sender and receiver need to be perfectly synchronized
 - If receiver is off by a bit (bit dropped during transmission) the plaintext will not make any sense
 - If bits are altered during transmission (noise hit) those bits will decrypt incorrectly
 - Provides only confidentiality, no authenticity

Cryptanalysis

- ◆ Is the process of attempting to **break** a cryptographic system to recover the *encrypted message* or the *key* used for encryption, or **both**
- ◆ The strategy used by the cryptanalyst depends on the nature of the encryption scheme and the information available to the cryptanalyst
- ◆ Based on the amount of information known to the cryptanalyst, the types of cryptanalytic attacks can be summarized as: (next slide)

Cryptanalysis (cont.)

Type of Attack	Known to Cryptanalyst
Ciphertext only	<ul style="list-style-type: none">Encryption algorithmCiphertext to be decoded
Known plaintext	<ul style="list-style-type: none">Encryption algorithmCiphertext to be decodedOne or more plaintext-ciphertext pairs formed with the secret key
Chosen plaintext	<ul style="list-style-type: none">Encryption algorithmCiphertext to be decodedPlaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key

Cryptanalysis (cont.)

Type of Attack	Known to Cryptanalyst
Chosen ciphertext	<ul style="list-style-type: none">• Encryption algorithm, Ciphertext to be decoded• Attacker is given Plaintext-Ciphertext pairs for some Ciphertexts he chooses. He then tries to deduce the key which will recover Plaintexts of other Ciphertexts.
Adaptive Chosen plaintext	<ul style="list-style-type: none">• Is a chosen-plaintext attack whereby the choice of next plaintext to use depends on ciphertext generated from previous plaintext used, ie after parts of decryption, attacker can deduce where the known plaintext fits, giving more clues.
Adaptive Chosen ciphertext	<ul style="list-style-type: none">• Is a chosen-ciphertext attack whereby the choice of next ciphertext to use depends on plaintext recovered from previous plaintext used

Exhaustive Key Search

- ◆ Always theoretically possible to simply try every key
- ◆ It is the most basic attack, which is directly proportional to key size
- ◆ Assume, either know or can recognize, when plaintext is found
- ◆ Tabulate for reasonable assumptions about number of operations possible:

Key Size (bits)	Time (1μs/test)	Time (1μs/10 ⁶ test)
32	35.8 mins	2.15 msec
40	6.4 days	550 msec
56	1140 years	10.0 hours
64	~500000 years	107 days
128	5 x 10 ²⁴ years	5 x 10 ¹⁸ years

Unconditional and Computational Security

- ◆ Unconditional security
 - No matter how much computer power is available, the cipher cannot be broken since the ciphertext provides insufficient information to uniquely determine the corresponding plaintext
- ◆ Computational security
 - Given limited computing resources (e.g. time needed for calculations is greater than the age of universe), the cipher cannot be broken
- ◆ Unconditional security would be nice, but the only known such cipher is the **one-time pad**. For all reasonable encryption algorithms, have to assume computational security where it either takes too long or is too expensive to bother breaking the cipher

Conclude : Intro to Cryptography

- ◆ Complex electronic systems and computers enabled the development of new and much more complex cryptographic systems
- ◆ In the early to mid 70s saw the development of the first modern **block ciphers**: Lucifer and DES
- ◆ The late 70s saw the development of **public key cryptography**: RSA
- ◆ These enabled the development of the systems we use today
- ◆ With modern systems, we can now have ciphers of vastly greater complexity than before, but still built using the same foundations