Introduction of Applied Cryptography

Chun-Jen (James) Chung

Arizona State University

Outline

- Security and cryptography concepts
- Classical Encryption Techniques

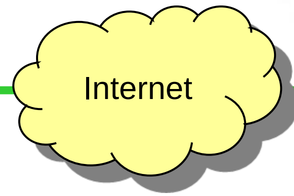
Concepts

Q: Why do we need cryptography for systems and/or networks?

An example

Insecure Channel

Alice



Bob

UGETGV

E

by Caesar with *k=2*

by DES, 3DES, AES, *etc*

SECRET

Shift *2*

UGETGV

The data has not been viewed by a 3rd party → *Confidentiality* ← Encryption

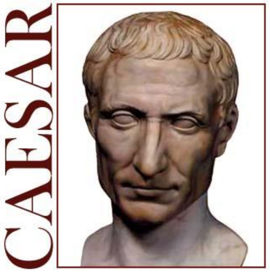
Confidentiality: *the protection of transmitted data from passive attacks (release of message contents and traffic analysis)* secret → **??** → UGETGV

Substitution
(Transposition)

Shift by 1 ? → TFDSFU

by 2 ? → UGETGV

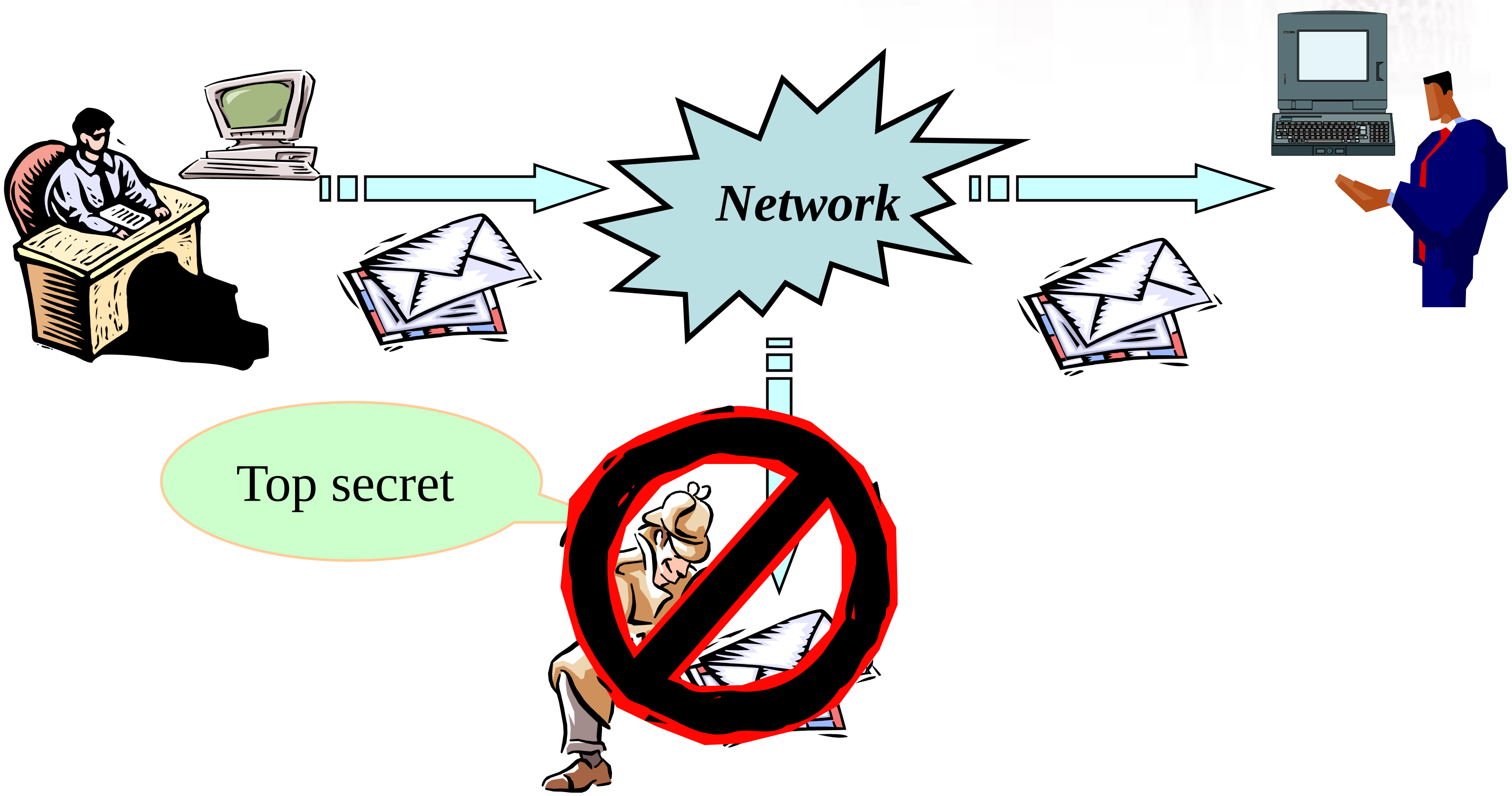
Caesar cipher



key = 2

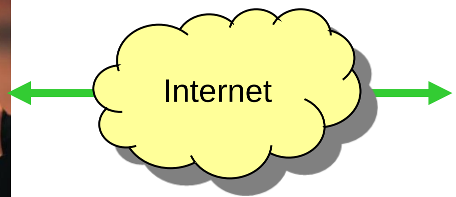
We will revisit this later

Confidentiality: Concealment of information or resources.



Use HMAC(Hashed message authentication code), such as HMAC-MD5, HMAC-SHA1

Alice

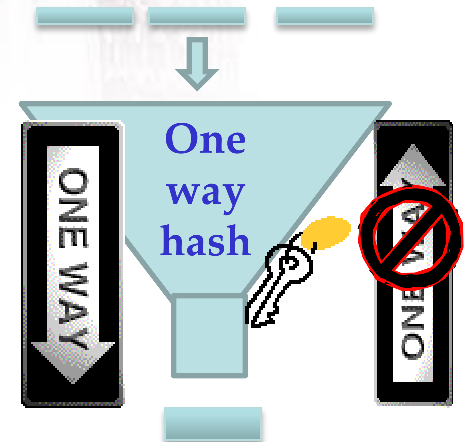


Bob

Wire transfer $ to E

E

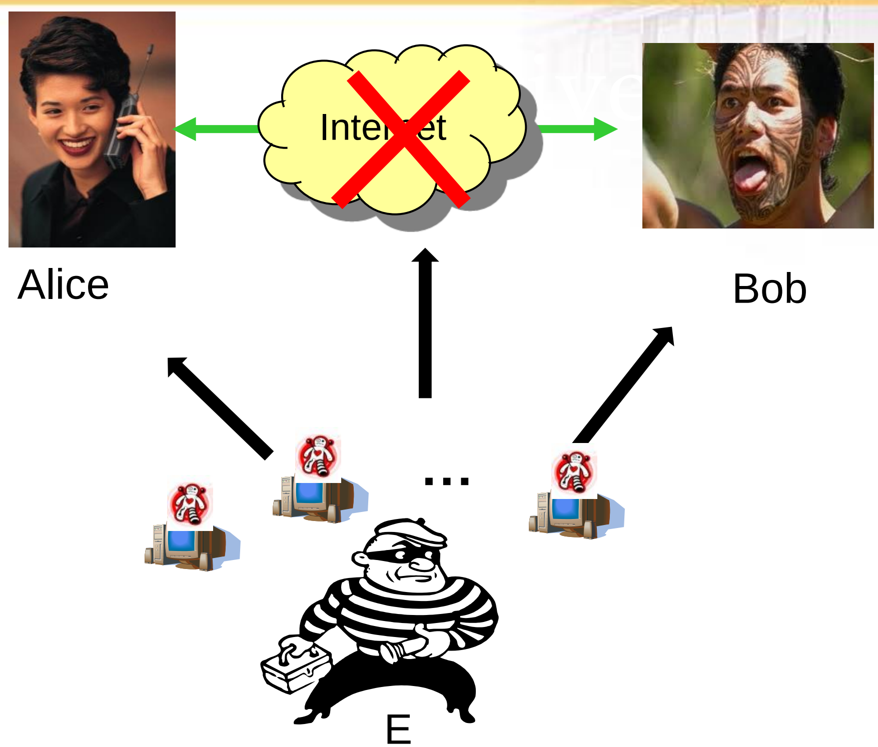
Inputs (e.g., msgs)



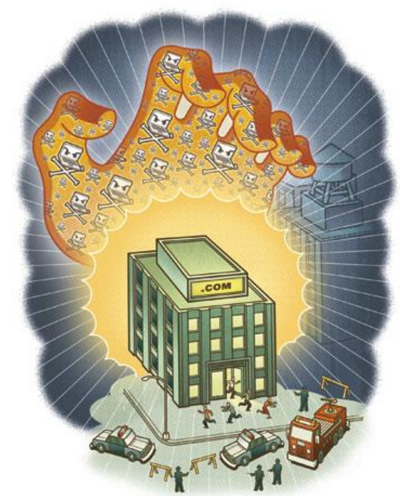
Fixed-size output

The data has not been modified in transit → *Integrity* ← Crytographic Hash func.

Integrity: the assurance that data received are exactly as sent by an authorized entity (i.e., contain no modification, insertion, deletion, or replay)



Distributed Denial of Service (DDoS) attacks

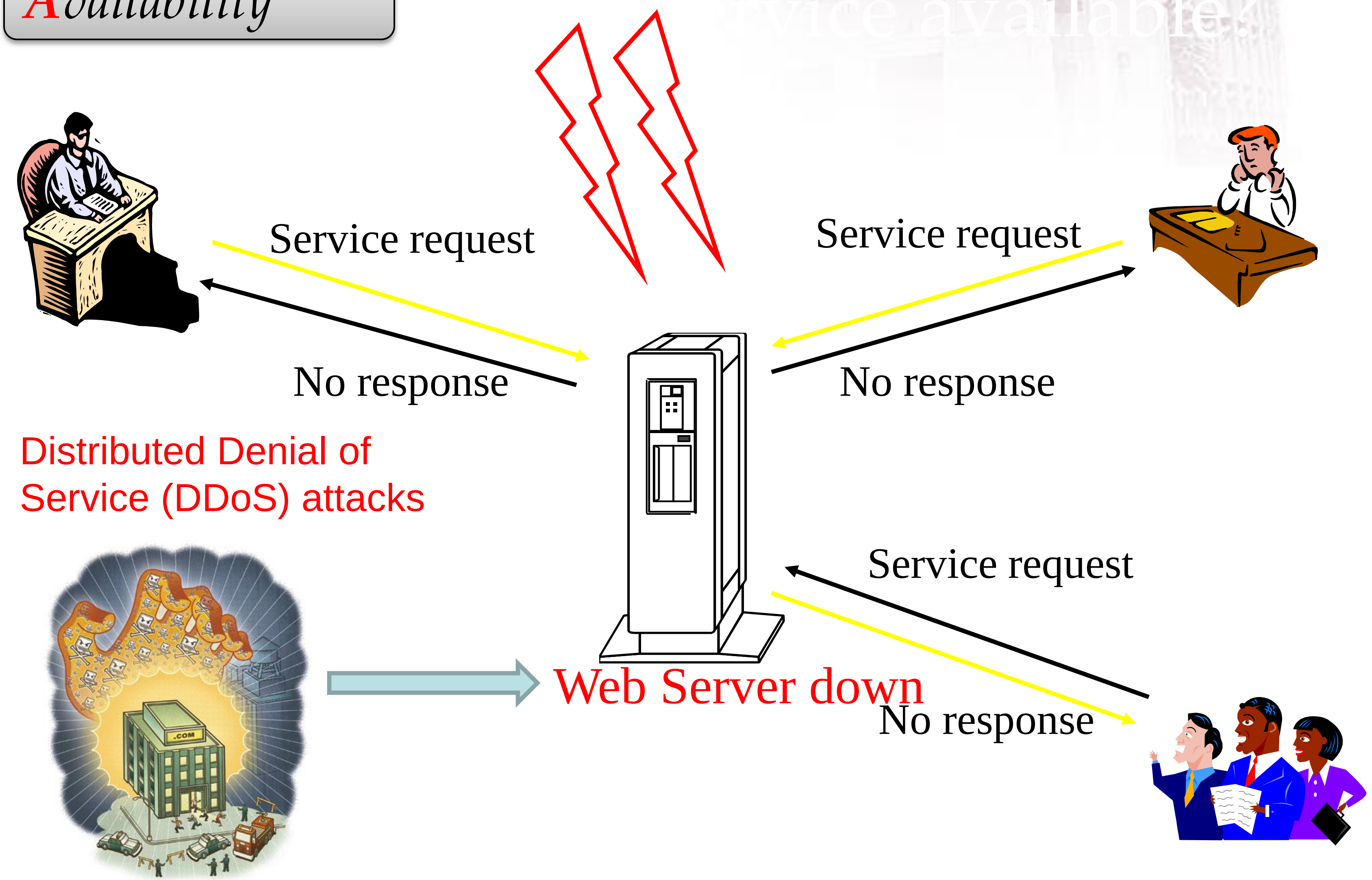


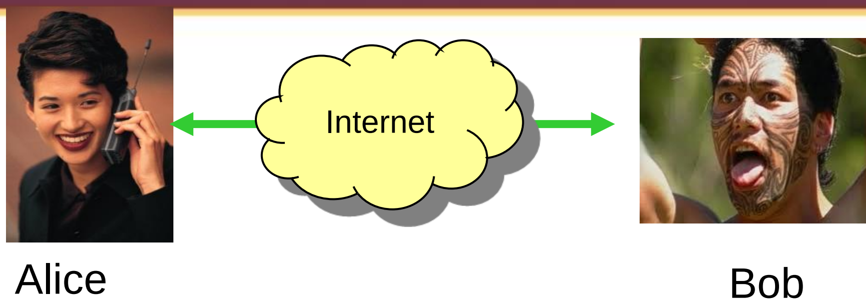
For any information system to serve its purpose, the information must be **available** when it is needed

Availability

Source: http://memeburn.com

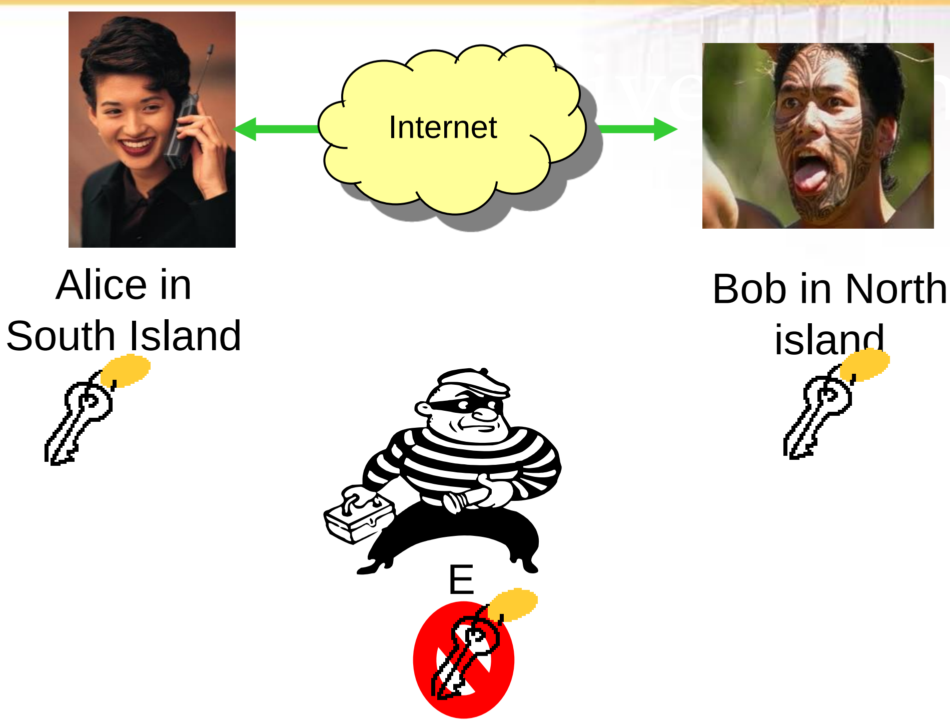
Availability





Cryptography

Requirement	Property	Cryptography
The data can not been viewed by a 3rd party	*Confidentiality*	Encryption
The data has not been modified in transit	*Integrity*	Hash func.
The data must be **available** when it is needed	**A**vailability	
	Authentication Non-repudiation	Encryption Message Authentication code (MAC) Hash func.



The claimed sender is the true sender → **Authentication** ← Encryption

Message Authentication code (MAC)

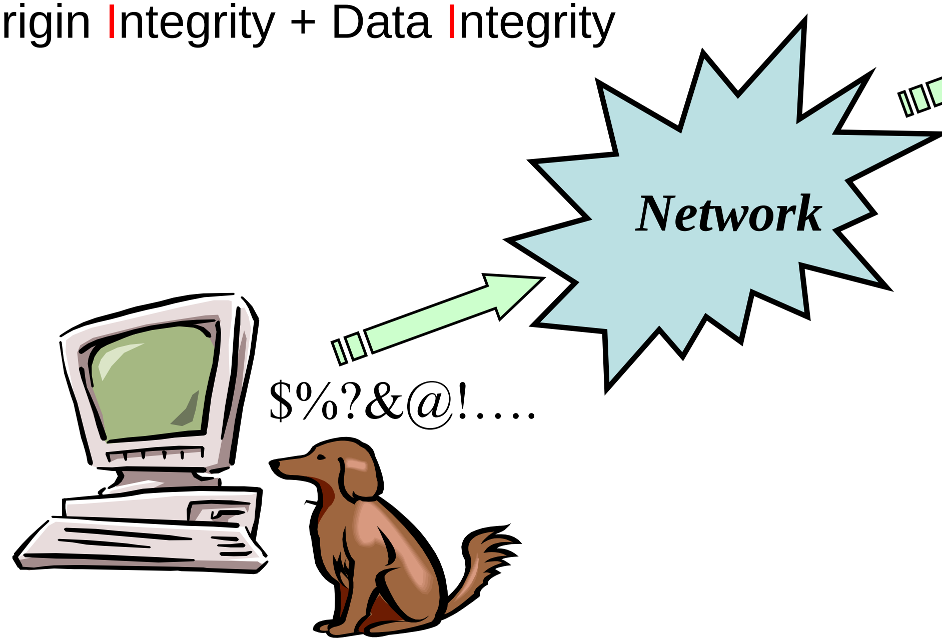
Hash func.

Authentication : the assurance that the communicating entity is the one that it claims to be.

Authentication: identification + verification

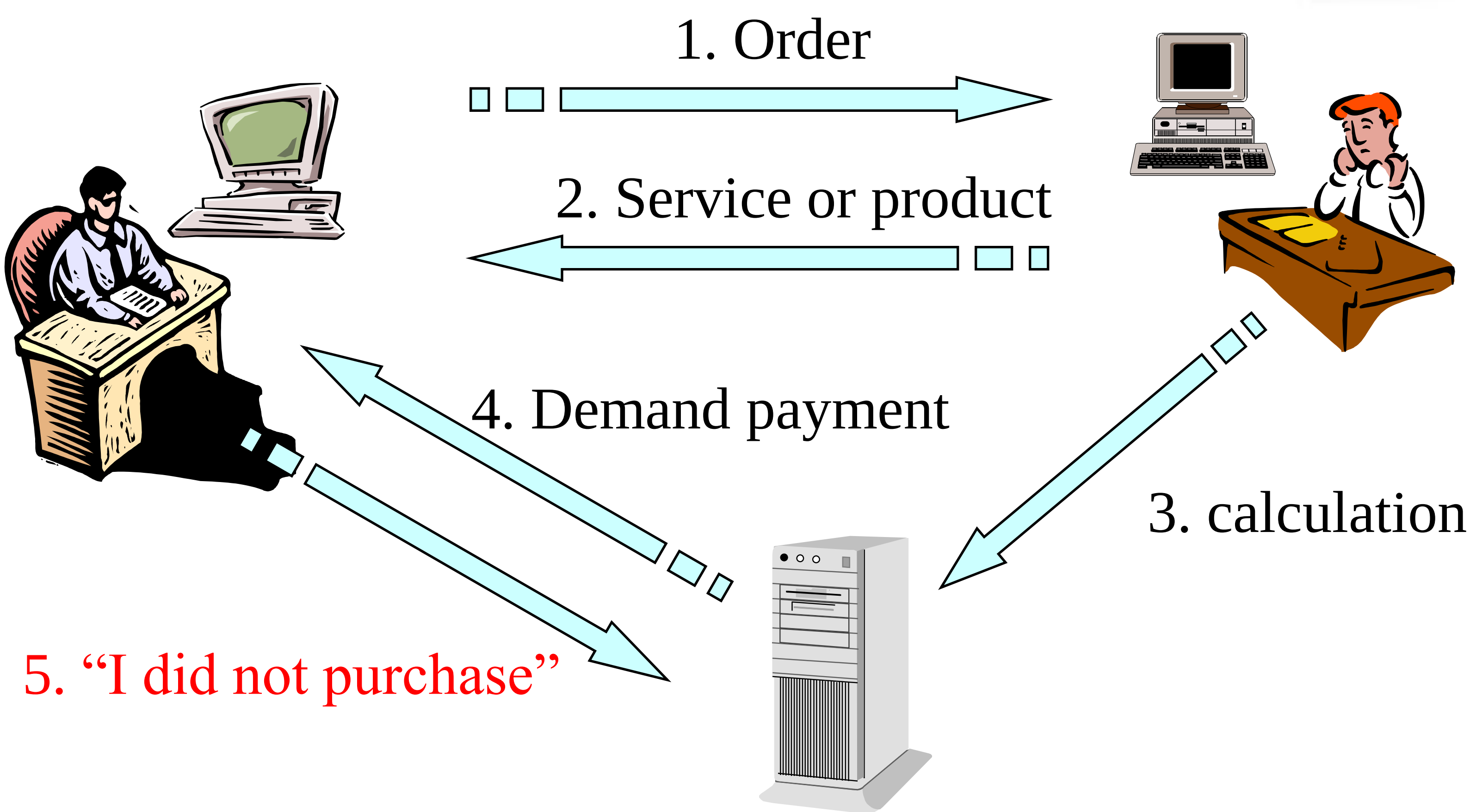
Who sent this message??

Authentication
= Origin Integrity
or
= Origin Integrity + Data Integrity



Non repudiation

Provides protections against *denial* by one of the entities involved in a communication of having participated in all or part of the communication

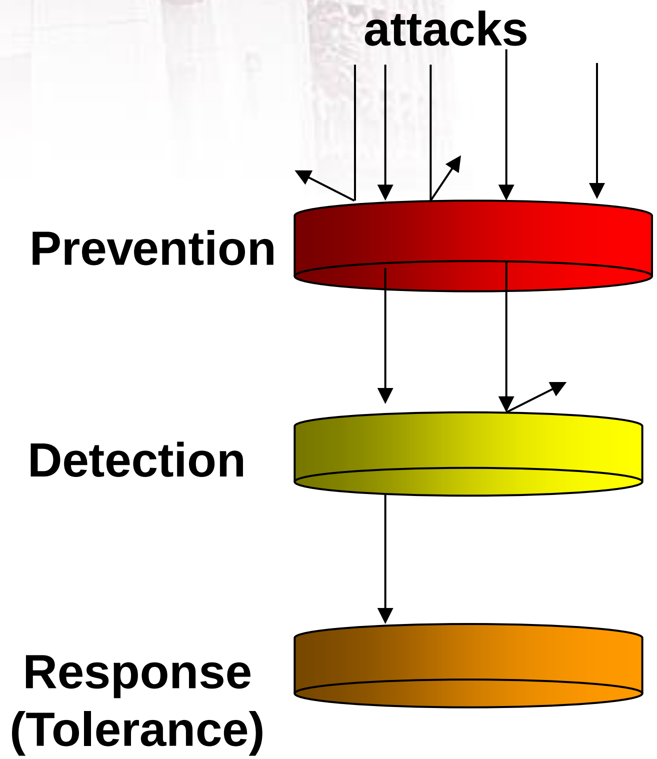


Security objectives summary

- Confidentiality
  - Prevent/detect/deter improper *disclosure* of information
- Integrity
  - Prevent/detect/deter improper *modification* of information
  - Authentication=Origin Integrity (or with Data Integrity)
- Availability
  - Prevent/detect/deter improper *denial of access to services* provided by the system
- These objectives have different specific interpretations in different contexts

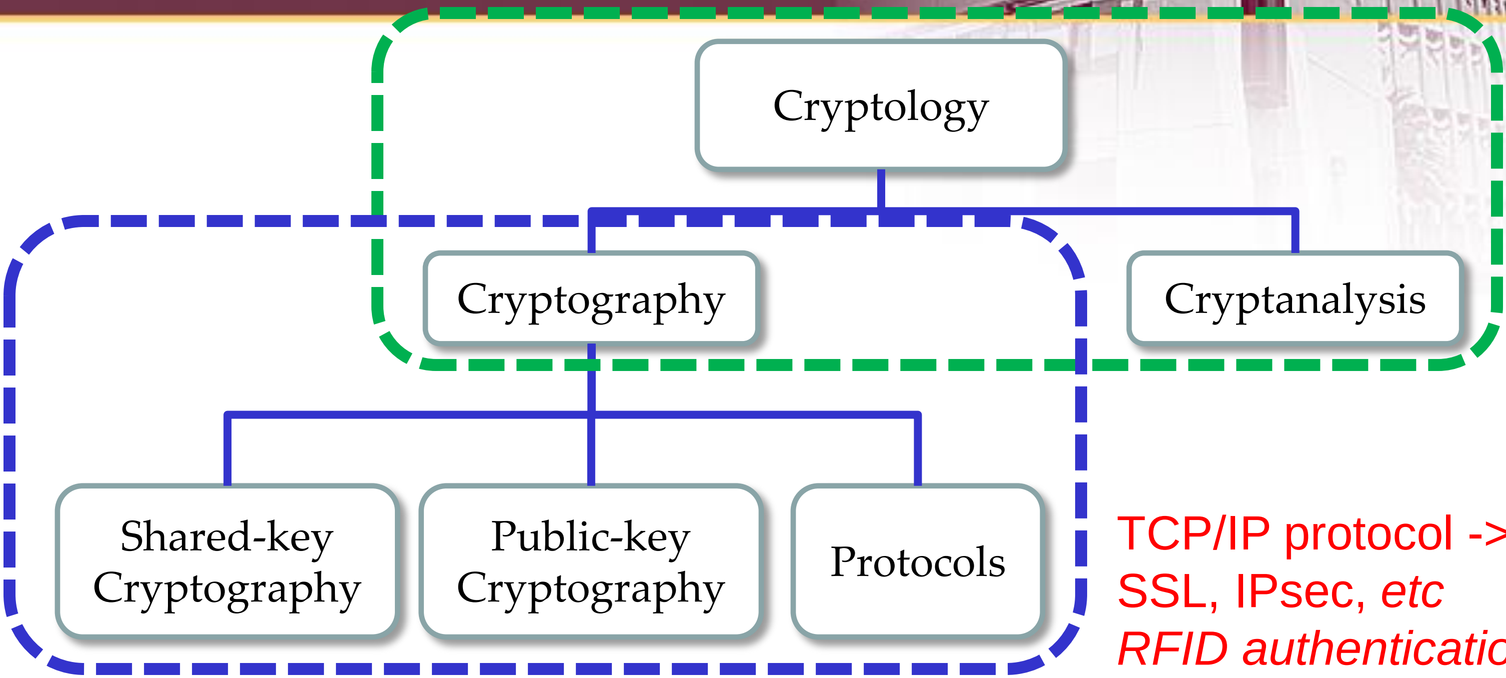
Security mechanisms

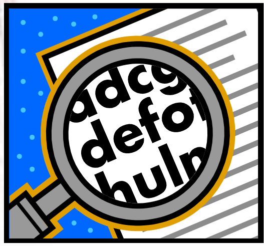
- In general, three types
  - Prevention
    - Example: access control (e.g., firewall)
  - Detection
    - Example: Auditing and intrusion detection (e.g., IDS, forensics)
  - Tolerance
    - Example: intrusion tolerance (e.g., ITS)



Cryptography underlies (almost) all security mechanisms

Introduction to Cryptography
Classical Encryption



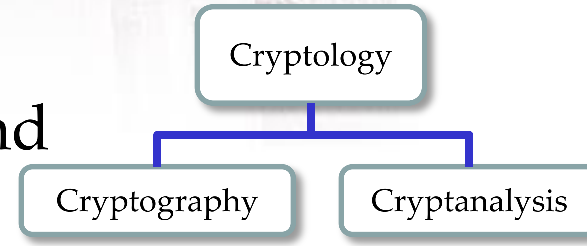


TCP/IP protocol -> SSL, IPsec, *etc*

RFID authentication protocol

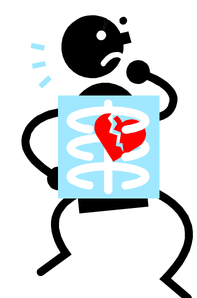
Computationally secure vs. information-theoretically secure

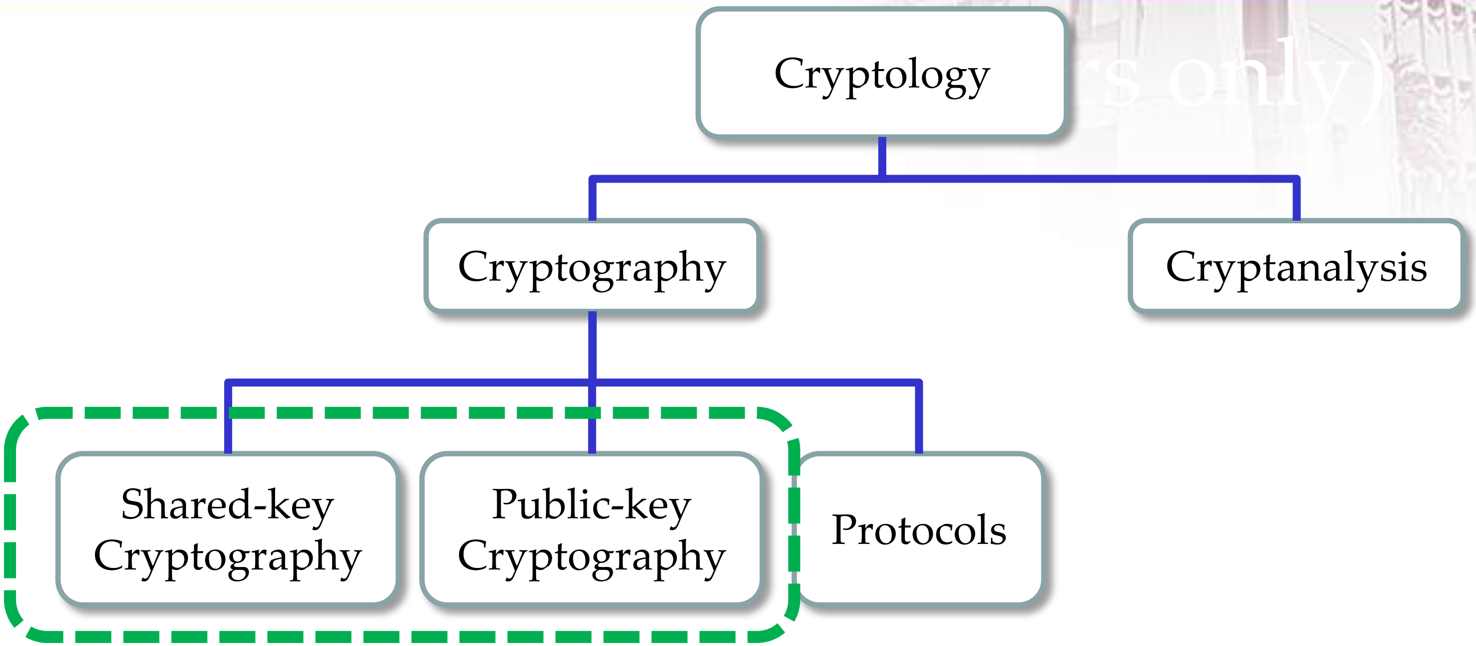
Cryptology: Cryptography + Cryptanalysis



- Cryptology
  - study of mathematical, linguistic, and other coding patterns and histories
  - An umbrella term for cryptography and cryptanalysis
  - **cryptologist**
- Cryptography
  - study of *encryption* principles/methods,
  - **cryptographer**

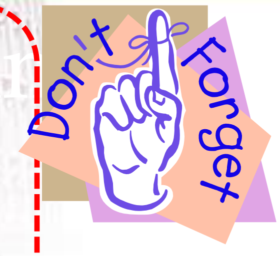
- Cryptanalysis (codebreaking)
  - the study of principles/methods of *deciphering* ciphertext without knowing key
  - **cryptanalyst**





Source: Prof. Baktier’s lecture at WPI

conventional, secret-key, single key



Encryption algorithms

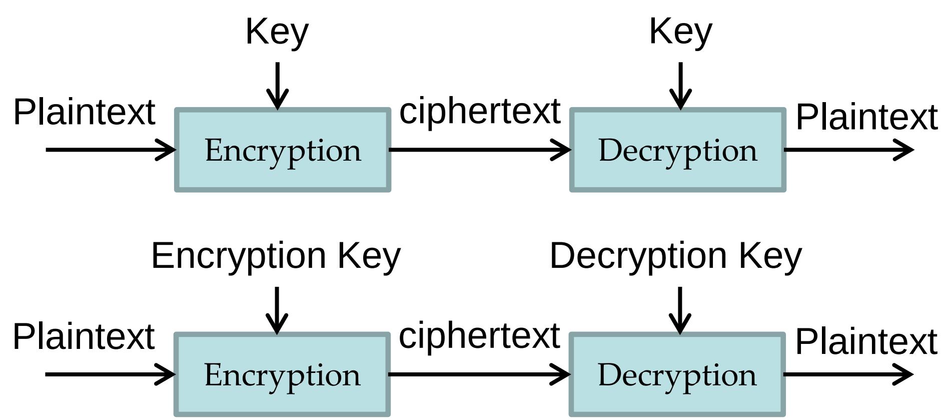
- Symmetric key
  - Block
    - Feistel
    - SPN
    - etc
  - Stream
    - random
    - Pseudo random
- Asymmetric key
  - *Later…*

64 bits, 128 bits, etc

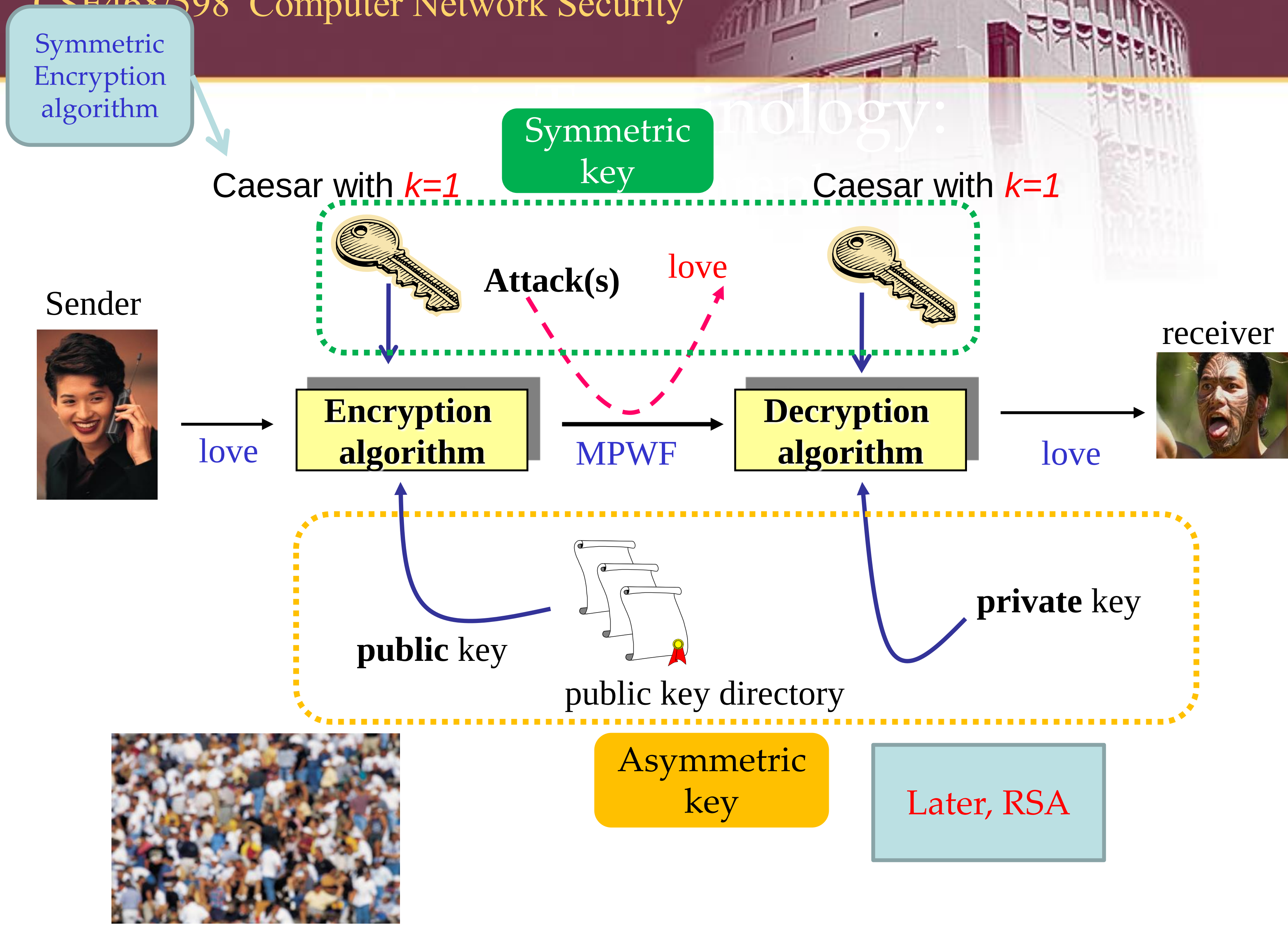
1 bit (or 1 byte)

Feistel: e.g., DES, Triple-DES SEED, Blowfish

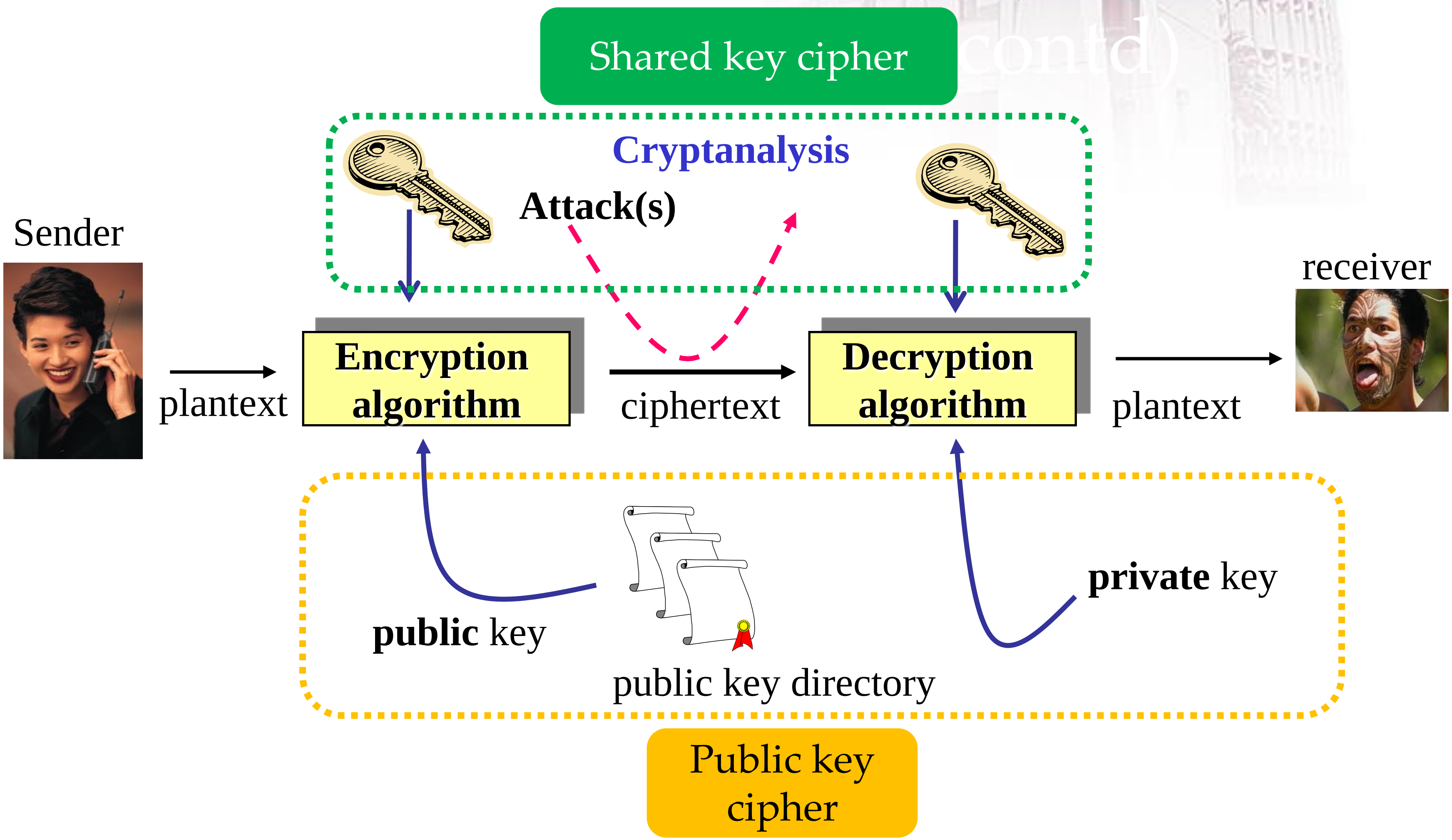
SPN: e.g., AES



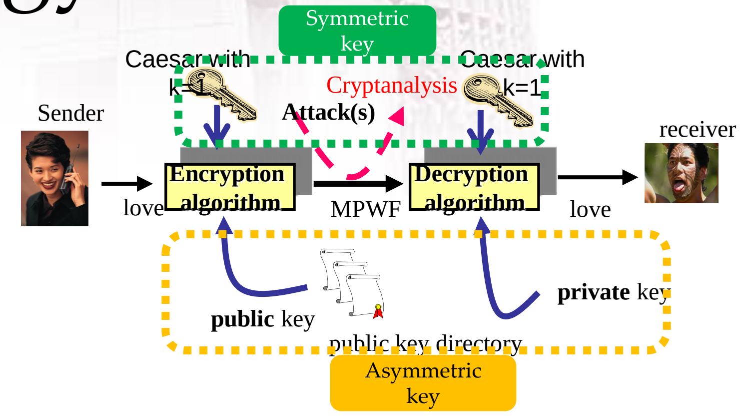
- Block ciphers take a number of bits and encrypt them as a single unit, padding the plaintext so that it is a multiple of the block size.
- Stream ciphers encrypt the digits (typically bytes) of a message one at a time.



Symmetric Encryption algorithm
Symmetric key
Caesar with k=1
Caesar with k=1
Attack(s)
love
Sender
receiver
love
Encryption algorithm
MPWF
Decryption algorithm
love
public key
private key
public key directory
Asymmetric key
Later, RSA

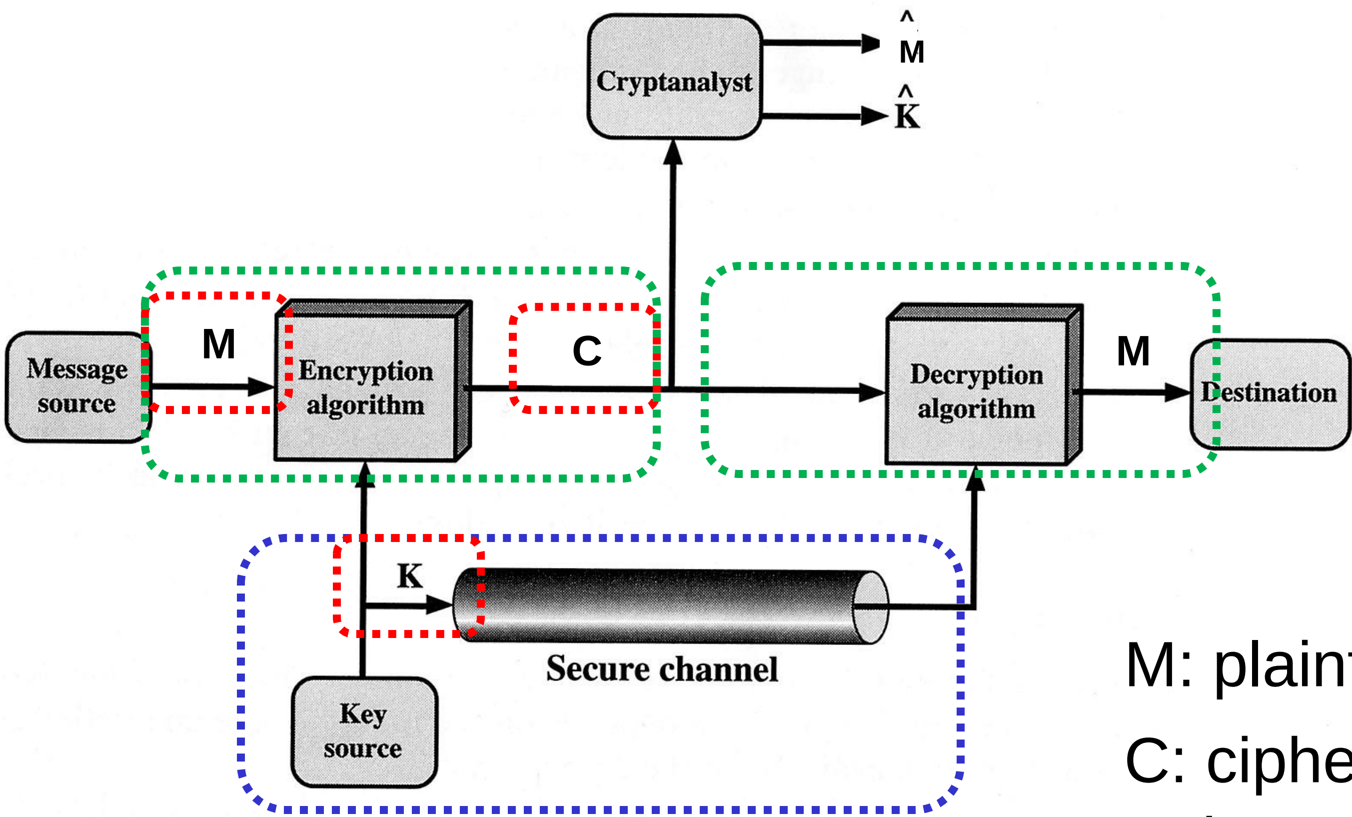

Shared key cipher
Cryptanalysis
Attack(s)
Sender
plantext
Encryption algorithm
ciphertext
Decryption algorithm
plantext
receiver
public key
public key directory
private key
Public key cipher

Basic Terminology



- **Cryptosystem**:
  - A system includes 3 algorithms:
    key generation + Encryption + Decryption
- Plaintext (**M**):
  - the original message, e.g., love
- Ciphertext (**C**):
  - the coded (or encrypted) message, e.g., MPWF
- **Cipher** (Encryption/Decryption algorithm):
  - algorithm for transforming from plaintext (ciphertext) to ciphertext (plaintext), e.g., Caesar, DES, AES.
- Key(s):
  - info. used in cipher known only to sender/receiver in symmetric cipher.
- Encrypt (encipher): **E(M)**
  - converting plaintext to ciphertext; e.g., love -> MPWF
- Decrypt (decipher): **D(C)**
  - recovering ciphertext from plaintext; e.g., MPWF->love

A symmetric key based Cryptosystem



This (key exchange/ distribution) will be covered later

M: plaintext

C: ciphertext

K: key

E: encrypt $C=E_K(M)$

D: decrypt $M=D_K(C)$

Notations in a Cryptosystem

Symmetric

M: plaintext

C: ciphertext

K: key

E: encrypt $C=E_K(M)$

D: decrypt $M=D_K(C)$

Asymmetric

Ke: encryption key

Kd: decryption key

E: encrypt C=E(M, Ke)

D: decrypt M=D(C, Kd)

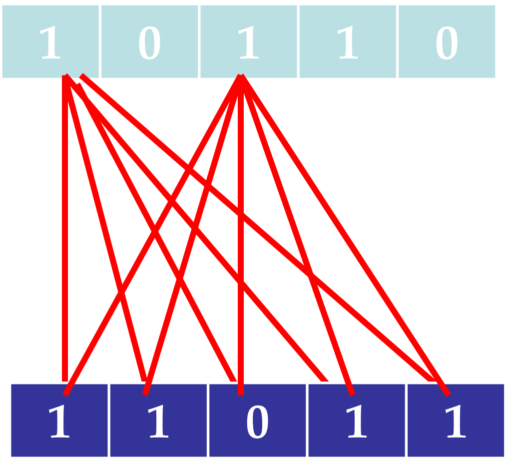
<M>K indicates that “M" is encrypted with the key “K".

<M1| M2>K indicates that “M1" and “M2” are encrypted with the key “K“, Where M1 | M2 (or M1 || M2) is the concatenation of M1 with M2.

Principle of ciphers

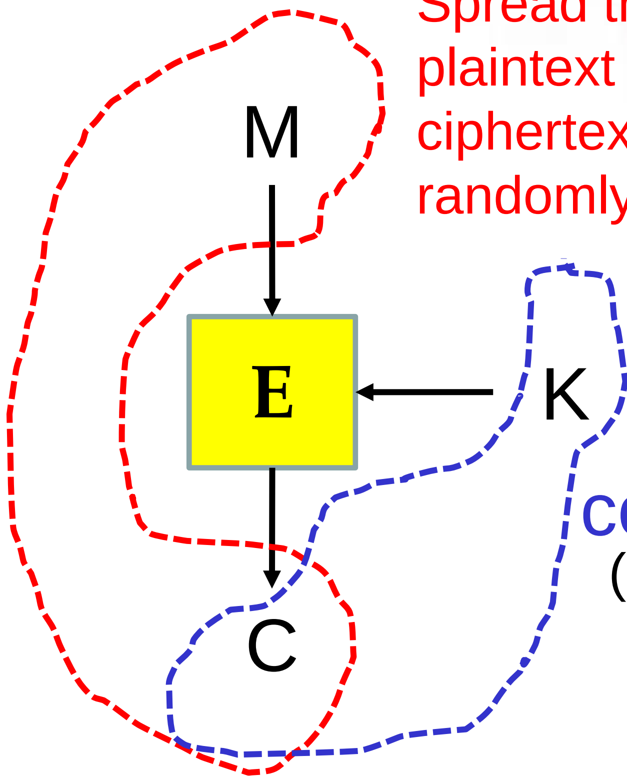
- Use both substitution and transposition
  - Proved by C. E. Shannon
  - Using information theory in 1945
  - *Product* ciphers
    - combines two or more transformations in a manner intending that the resulting cipher is more secure.

- Formulate the principles of "confusion" (standing for substitution) and "diffusion" (standing for transposition)



diffusion (transposition/permutation)

Spread the information from single plaintext letters over the entire ciphertext, making the plaintext randomly distributed in the cyphertext.



confusion (substitution)

Make relationship between ciphertext and key as complex as possible, and make it harder to find the key

- **Substitution**
  - Change the character(s)
  - Plaintext: come here at once
  - Ciphertext: 
- **Transposition (Permutation)**
  - Change the order of character
  - Plaintext: come here at once

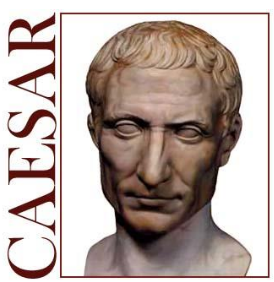
  - Cipheretxt: ocemeher t a noec

Classical (Historical) ciphers

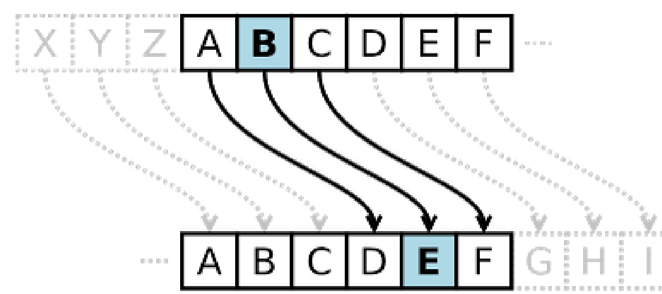
- Substitution ciphers
  - Monoalphabetic
    - The same plaintext letters are always replaced by the same ciphertext letters
  - Polyalphabetic
    - using multiple substitution alphabets
- Transposition ciphers

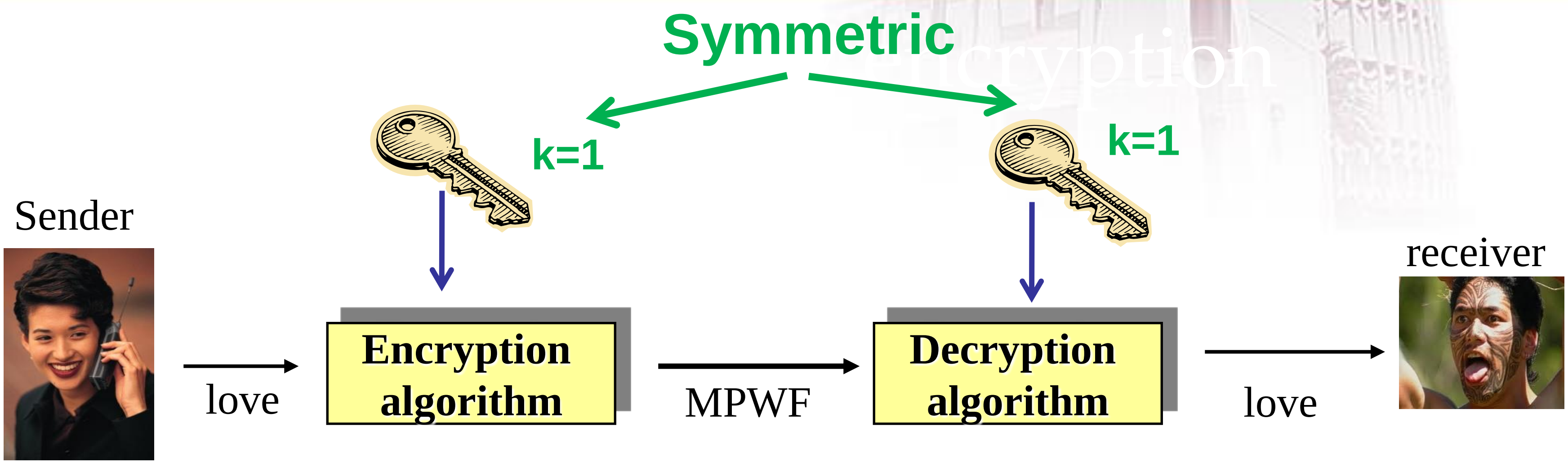
Caesar Cipher

- earliest known **substitution** cipher
  - Also known as shift cipher
- named after Julius Caesar
  - who used it in his private correspondence
- replaces each letter by the ith letter on



key = i





Q: If key (k) is 5? What will be the ciphertext of the plantext “love”?

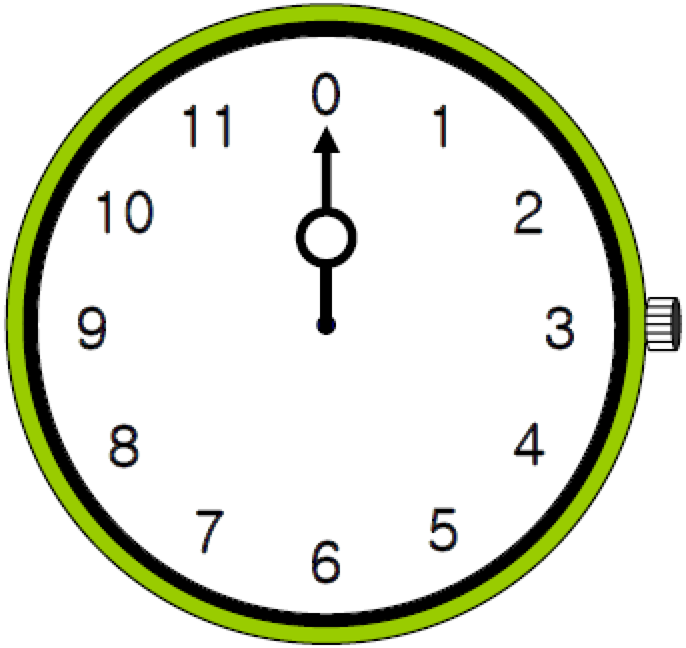
A: QTAJ 

$$C \equiv E(M) \equiv M + k \mod 26$$

a **modulo** *n* (abbreviated as *a* mod *n*)

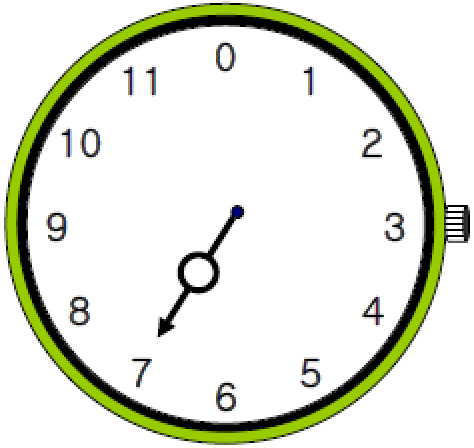
Mod. operation practice using a clock

- What is this?
  - An analog clock



Mod. operation practice

- Q: if we move the hour[short] hand by 6 hours?
  consider it's a 12-hour clock

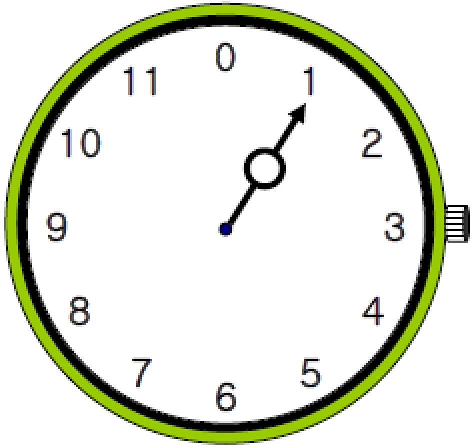


- 13? No.

- It indicates…

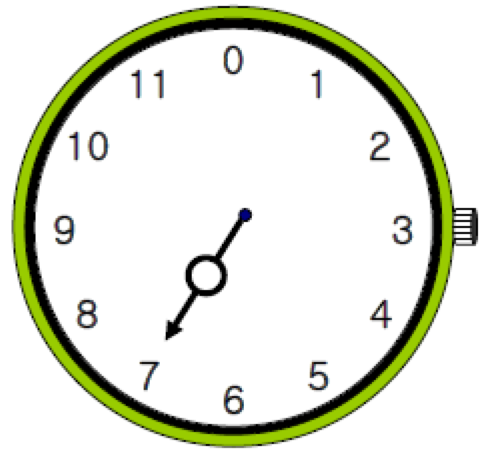
$$7+6 \bmod 12 \equiv 1$$

Or $$7+6 \equiv 1 \pmod{12}$$



Mod. operation practice

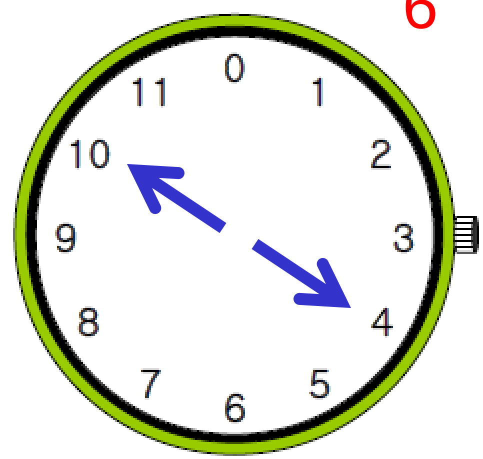
- When the hour[short] hand points out 7,
- how can we make it indicates 0, not using a negative number?



Not negative number: (-7)

$$(7 + \square) \bmod 12 = 0$$

6

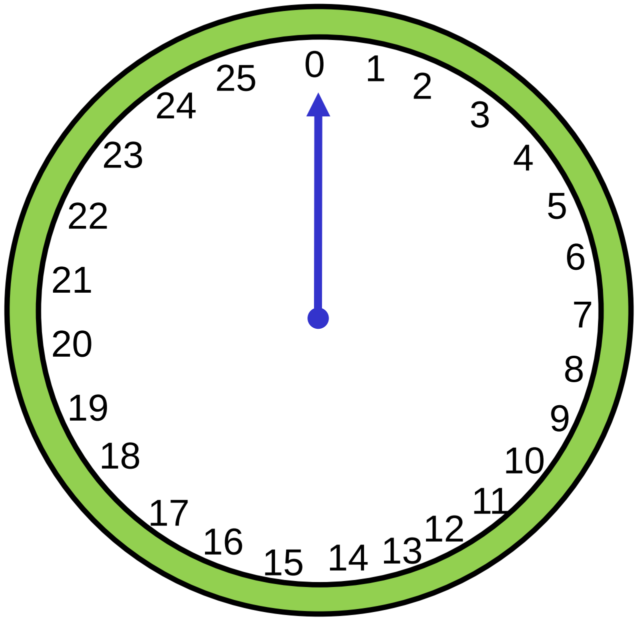


- It is 5.
- **Q ? (4-6) mod 12?**

New clock ??

mathematically give each letter a number

letter	a	b	c	d	e	…	w	x	y	z
number	0	1	2	3	4	…	22	23	24	25



New clock

Source: funny-clock-a.qehqlw.webhop.net / picsobsession.blogspot.com

letters	a	b	c	d	e	…	w	x	y	z
	0	1	2	3	4	…	22	23	24	25

Ex.1) key=3, E(c) ?

Sol ≡ c + k mod 26
≡ 2 + 3 mod 26
≡ 5 ➔ F

Ex.2) key=3, E(x)?

Sol ≡ x + k mod 26
≡ 23 + 3 mod 26
≡ 26 mod 26 ≡ 0 ➔ A

Q) key=3, E(y) ?

≡ 24 + 3 mod 26
≡ 27 mod 26 ≡ 1 ➔ B

P ≡ D(C) ≡ (n-k) mod 26

Ex.3) key=3, D(F)? Sol.) (5-3) mod 26 ≡ 2 ➔ c

Ex.4) key=3, D(A) ? **Sol.) (0-3) mod 26 ≡ (-3) mod 26 ≡ 23 ➔ x**

We will use this *mod* operation again when we learn the principle of **RSA** (most well-known public key encryption algorithm)

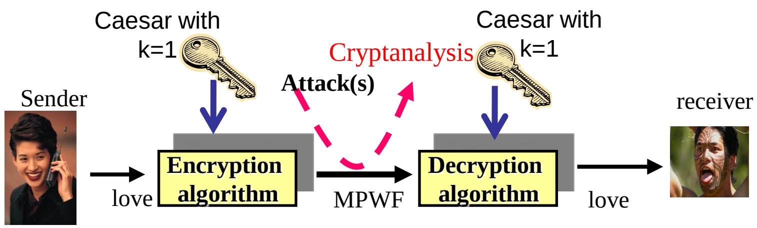
RSA (Ron Rivest, Adi Shamir and Leonard Adleman)

One more exercise ?

- Cipher (Caesar) is known
- Key is hidden
- Q: What is then the plaintext for the following ciphertext?

SRHECP

uofcan



- given ciphertext, just try all shifts of letters (k=1, 2, 3, …)
- You did break the cipher using a brute-fore attack!

Q: possible key size?

Caesar cipher

- Why Caesar cipher is too week?
  - Possible key size is? 25

P ≡ D(C) ≡ (n-k) mod 26

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

25 keys

plaintext

Caesar cipher

break ciphertext "GCUA VQ DTGCM"

Answer : “easy to break”

Q: any ideas to improve this cipher?

A: use a Mono alphabetic Cipher which uses an arbitrary substitution

We could shuffle (jumble) the letters **arbitrarily**; each plaintext letter maps to a different random ciphertext letter

Mono alphabetic Cipher

1) Use a key phrase (or keyword): e.g., ASINTOR.

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

2) Use other alphabet after ‘R’ except them used in the key phrase

3) Use remaining alphabets in the alphabetical order

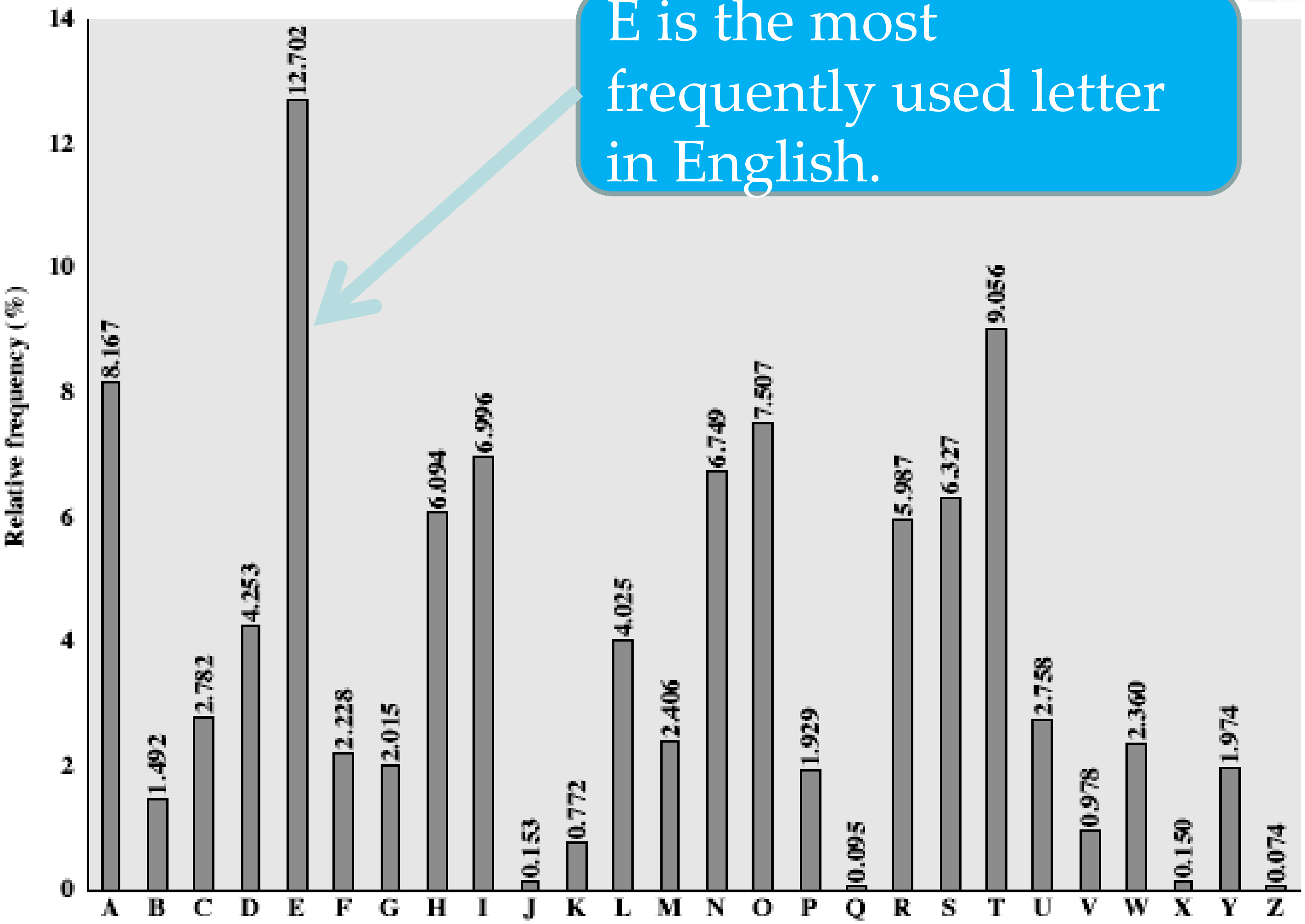
Q: how many cases?

A: $26! = 26 \times 25 \times 24 \ldots \times 2 \times 1 \approx 4 \times 10^{26}$

Q: Can we perform a cryptanalysis?

Mono alphabetic Cipher

E is the most frequently used letter in English.



The most common letters in English are ..

1. **E**
2. T
3. A
4. O
5. I
6. N
7. S
8. H
9. R

An example

- Ciphertext using Mono alphabetic Cipher

SLAAP UNAOL NVVKZ ABMMP UDOPS LRLLW PUNAO LIHKZ ABMMV BAPZA OLJYP APJHS WYVIS LTAOH AHUFJ VTWBA LYMPY LDHSS ULLKZ AVZVS CLPAO HZAVH JAHZN HALRL LWLYP AOHZA VMPNB YLVBA DOPJO IPAZH YLOHY TMBSH UKKLU FAOLT LUAYF PAOHZ AVKVA OPZDP AOVBA BUYLH ZVUHI SFKLS HFPUN AYHMM PJ

- Frequency computation result

A: 29, B: 9, C: 1, D: 4, E: 0, F: 5, G: 0, H: 18, I: 4, J: 6, K: 7, L: 24, M: 9, N: 6, O: 14, P: 18, Q: 0, R: 2, S: 10, T: 4, U: 11, V: 14, W: 4, X: 0, Y: 11, Z: 12

1. **E**
2. T
3. A

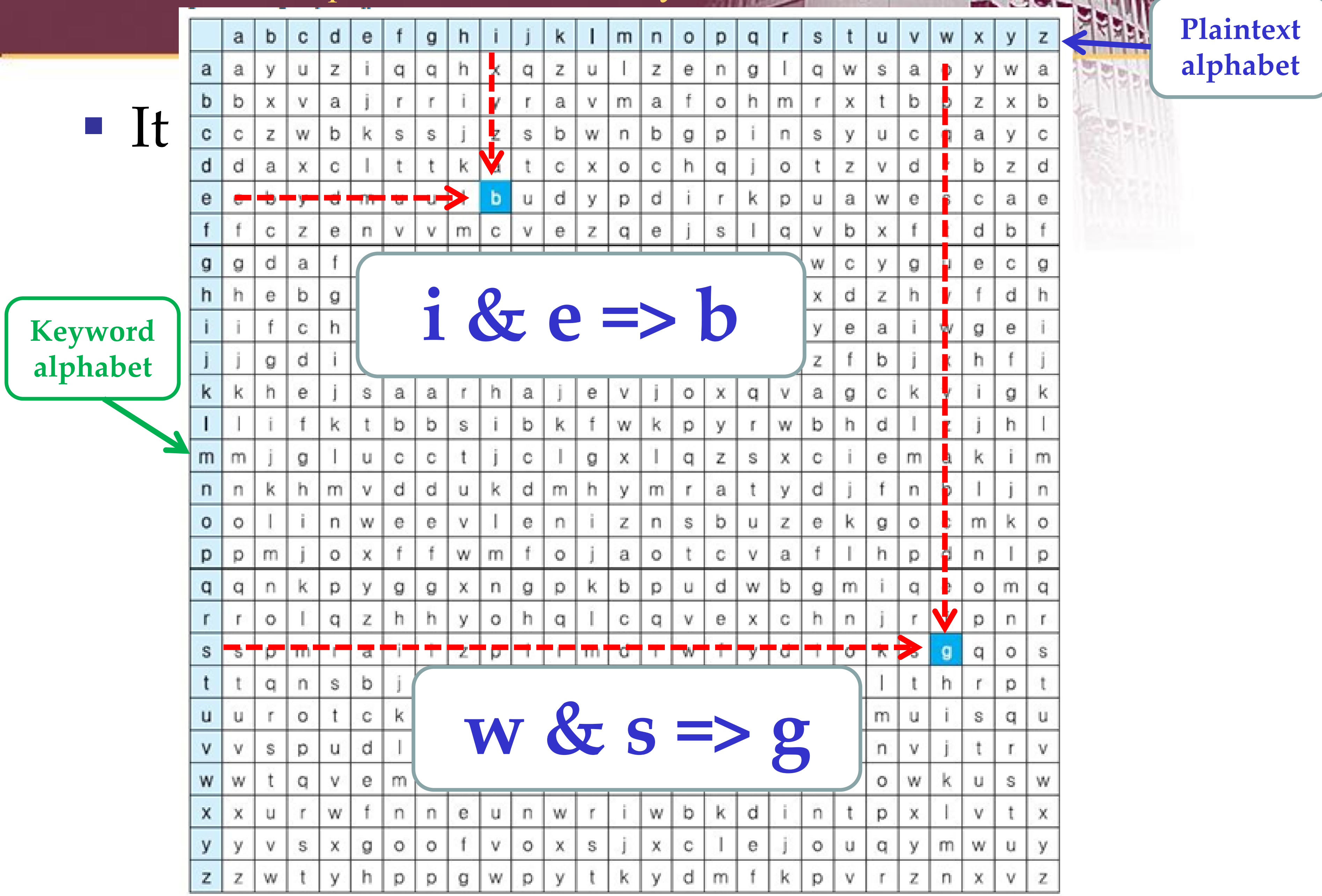
…

- **Q: Strong candidate** key(s)?
  - **‘E’-> ‘A’ (22)**
  - **‘E’-> ‘L’ (7)**

- Q: Can we hide the frequency of alphabet?

- Any ideas?
  - Using **a series of different Caesar ciphers in sequence with different shift values**.

- It



Vigenère cipher

Plaintext	w	i	s	h	t	o	b	e	f	r	e	e	f	r	o	m	m	y	s	e	l	f
keyword	s	e	c	r	e	t	i	s	b	e	a	u	t	i	f	u	l	s	e	c	r	e
ciphertext	g	b	s	y	a	x	f	a	r	p	i	c	j	t	x	f	w	o	u	k	l	u

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

w & s => g
i & e => b

Polygram Substitution
(polyalphabetic substitution.)

Using Vigenère cipher, E can be enciphered as different ciphertext letters at different points in the message, thus defeating simple frequency analysis

Vigenère cipher (cont'd)

- Cryptanalysis? The primary weakness of the Vigenère cipher is *the repeating nature of its key*.
- Kasiski test:
  - takes advantage of the fact that *repeated words* may, by chance, sometimes *be encrypted using the same key letters*, leading to *repeated groups in the ciphertext*.
  - For example, consider the following encryption using the keyword ABCD:
  - Plaintext: ***CRYPTO***ISSHORTFOR***CRYPTO***GRAPHY
  - Key: ABCDABCDABCDABCDABCDABCDABCD
  - Ciphertext: ***CSASTP***KVSIQUTGQU***CSASTP***IUAQJB
- There are more substitution ciphers though, we move on..

Classical (Historical) ciphers

- Substitution ciphers
  - Monoalphabetic
  - Polyalphabetic
- Transposition ciphers

Transposition (**permutation**) Cipher

- these hide the message by rearranging the letter order
- without altering the actual letters used
- can recognize these since have *the same frequency distribution* as the original text

Transposition Cipher

- Rail fence cipher (a.k.a., a zigzag cipher)
  - write message letters out downwards and diagonally over a number of rows or rails
  - then read off cipher row by row
  - Plaintext: meet me after the toga party

M	E	M	A	T	R	H	T	G	P	R	Y
E	T	E	F	E	T	E	O	A	A	T	

  - ciphertext: MEMATRHTGPRYETEFETEOAAT

Rail fence cipher (cont'd)

p_1	p_2	p_3	p_4
p_5	p_6	p_7	p_8
p_9	p_{10}	p_{11}	p_{12}
p_{13}	p_{14}	p_{15}	p_{16}

- $P=p_1p_2p_3\cdots p_{16}$
- $C=p_1p_5p_9p_{13}\cdots p_{12}p_{16}$
- Ex. Plaintext:
  - LASTNITEWASHEAVENPLEASEMARRYME
- Ciphertext:
  - LTELAAEAERSWVARTAESYNSNEMIHPME
- Ciphertext
  - LEVSMTAAYIEERNHLRTSPASANMAWEEE

LASTNI
TEWASH
EAVENP
LEASEM
ARRYME

LTELAA
EAERSW
VARTAE
SYNSNE
MIHPME

This process is iterated, it is hard to guess the plaintext

Rounds in encryption

Plaintext	u	n	i	v	e	r	s	i	t	y	o	f	c	a	n	t	e	r	b	u	r	y

key	4	5	3	2	8	7	6	1	
Plaintext′	u	n	i	v	e	r	s	i	
	t	y	o	f	c	a	n	t	
	e	r	b	u	r	y	0	0	padding

Cipher text	i	t	0	v	f	u	i	o	b	…

Now think about …

- Q: Why does the ciphers introduced so far not secure?
- A: because of language characteristics / **weak key size**
- Q: Any ideas to improve them (you already know the answer)?
- A: Use both substitution and transposition *together*

From classical to modern ciphers

- Consider using 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
- Q: What is this type of ciphers called?
  - A: **product** ciphers
- This is the ***bridge*** from classical to modern ciphers

- Q: What is most well-known and widely used modern cipher(s)?
- A: **DES** (Data Encryption Standard),
  AES (Advanced Encryption Standard),…