NT219- Cryptography

PhD. Ngoc-Tu Nguyen

tunn@uit.edu.vn

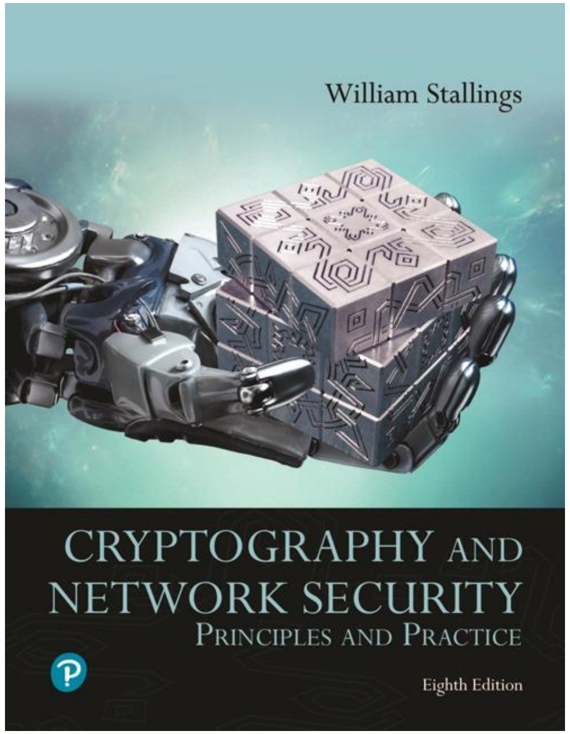
Week2: Cryptanalysis classical cipher systems

Outline

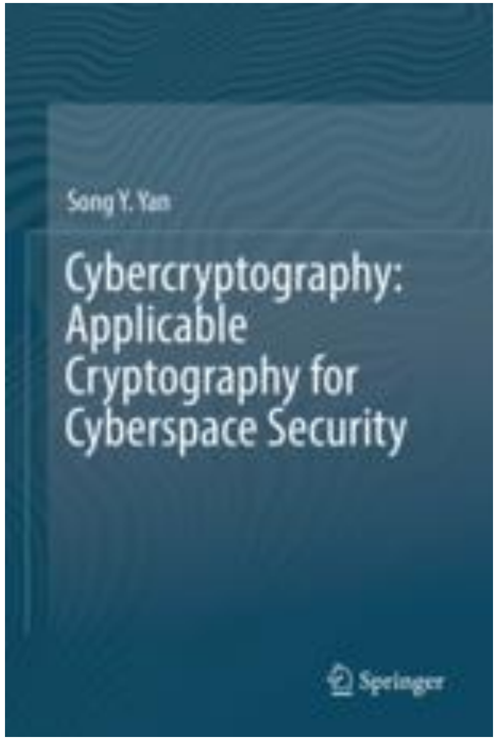
- What is cryptograph?
  - Algorithms
  - Terminologies
  - Application areas
- Cryptanalysis classical cipher systems
  - Algorithms
    - Substitution ciphers
    - Transposition ciphers
  - Cryptanalysis (CTF + tools) ---> assignment

Textbooks and References

- Text books



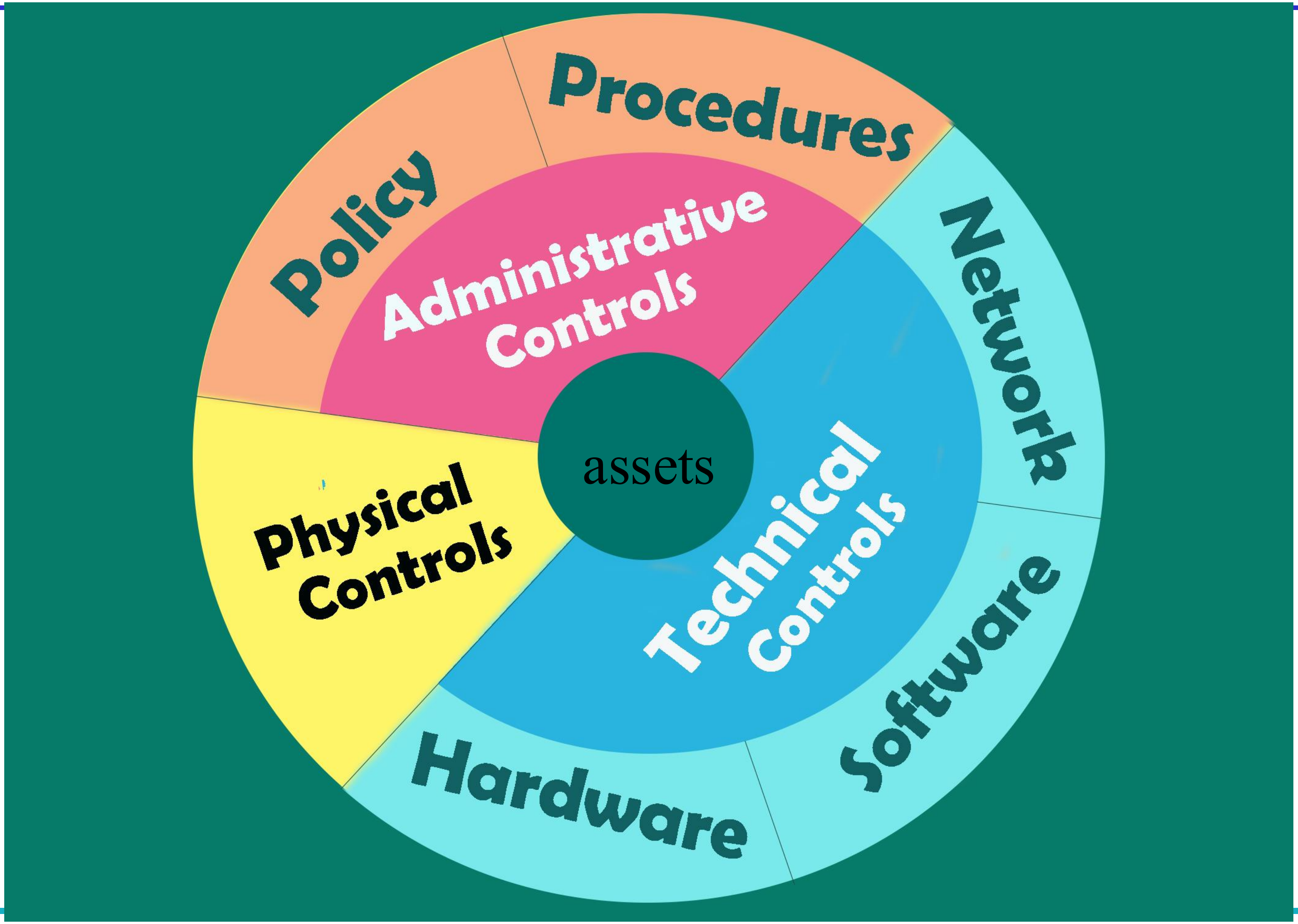
[1] Chapter 1,3



[2] Chapter 1, 4

Motivations

Defense in depth:



Motivations

Technical solutions

Defense in depth:

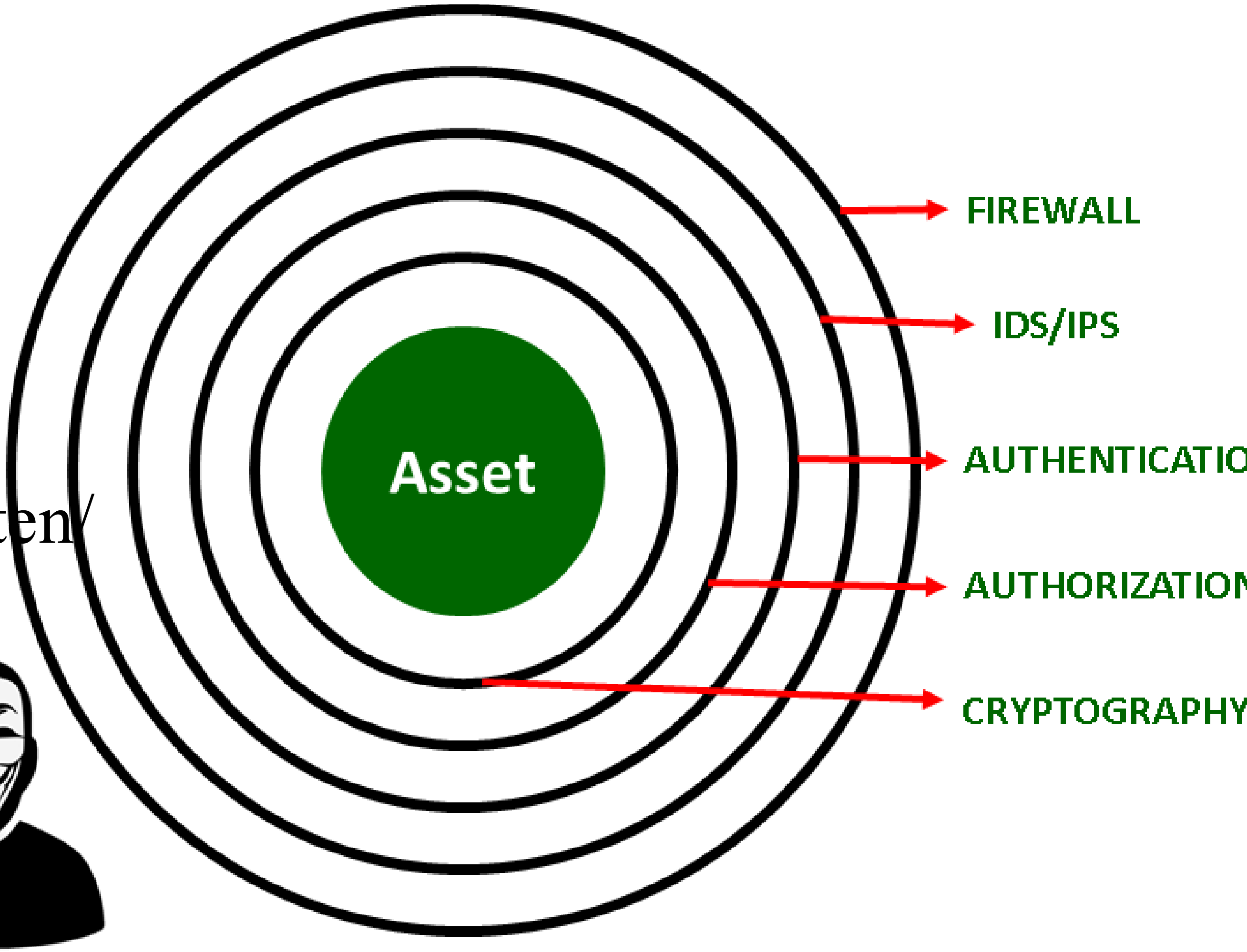
Host or network?

☞ https://owasp.org/www-project-top-ten/

- Internal threats
- External threats
- Partners

adversary

ONION MODEL



Example solutions

What is cryptograph?

- Cryptology= Cryptography + Cryptanalysis

<u>Goals</u>

What?

Cipher systems

- Confidentiality
- Privacy

- Sysmmetric (AES)
- Asymmetric (RSA, ECC, CRYSTALS-KYBER)

- Integrity
- Authentication
- Non-repudiation (Accountability)

Hash functions
Message authentication code (MAC)
Digital signature (digital certificate)

- Availability

How cryptograph works?

1. Cipher systems

"Readable data" → "One-way function" → "Unreadable data"

Plaintext

Ciphertext

Encryption: Easy to compute →

Decryption ←

- Hard if user don't know some special values
- Easy if user know the special values (secret values)

https://codebeautify.org/encrypt-decrypt

https://www.devglan.com/online-tools/aes-encryption-decryption

How cryptograph works?

2. Cryptographic hash functions

"Readable data" — "One-way function" → "Unreadable data"

Plaintext

Digital digest

Easy to compute

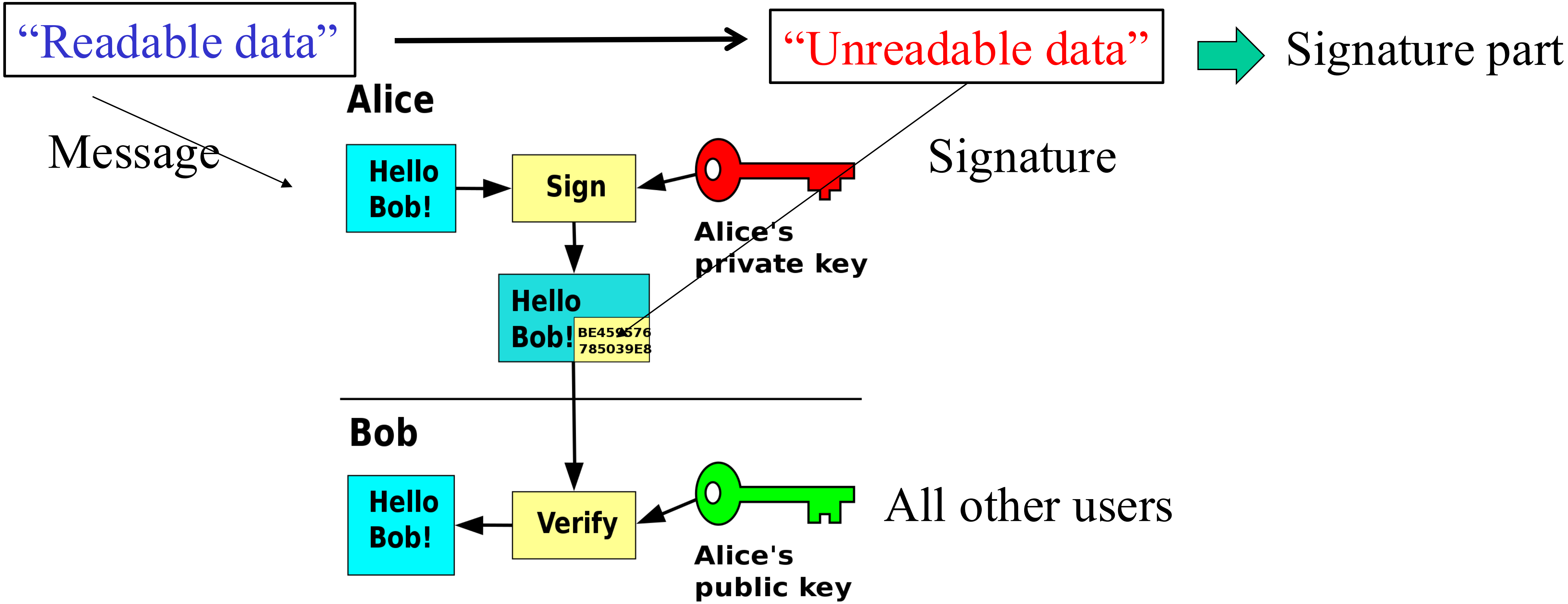
"Very hard" to compute

All sides

https://emn178.github.io/online-tools/sha3_512.html

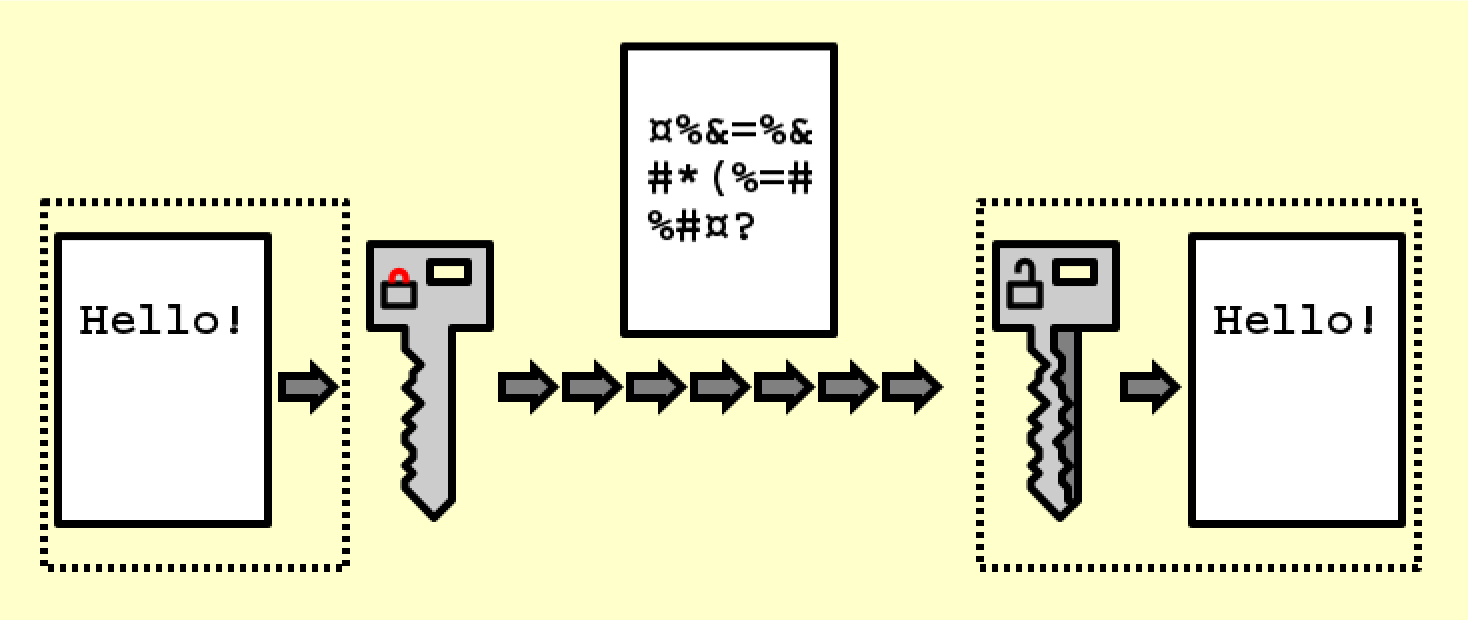
How cryptograph works?

3. Digital signature systems



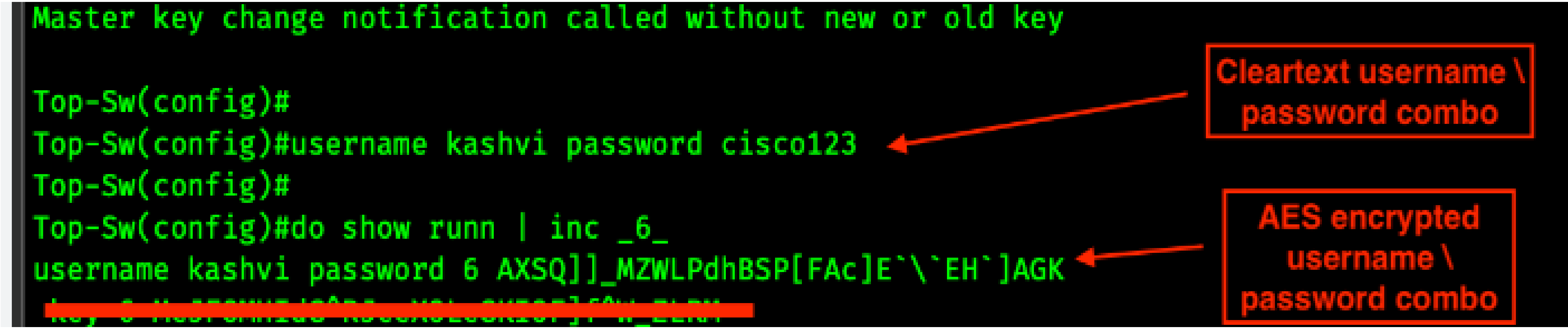
Some application scenarios

- **Cipher systems**
  - Data
    - Transmission state
    - Process state
    - Storage state



Transmission state

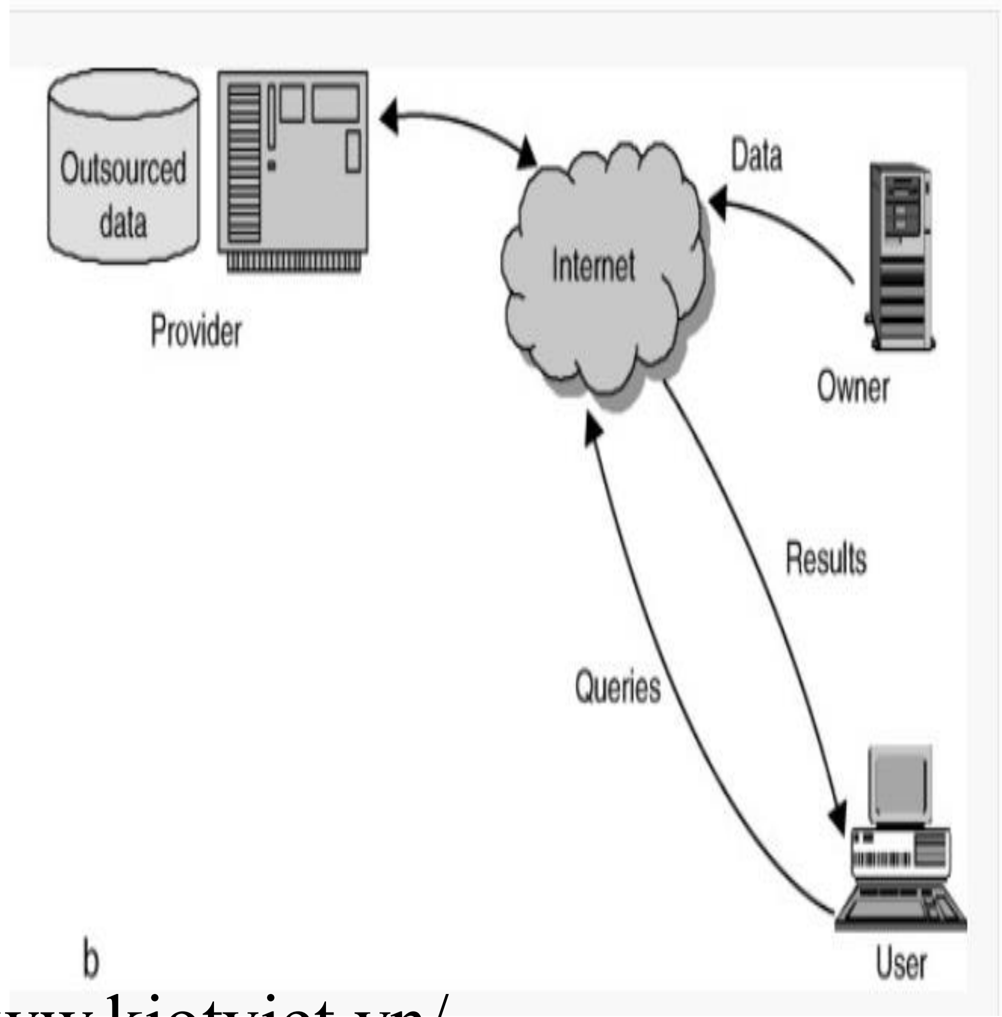
- **Cipher systems**
  - Data
    - Storage state



https://anycript.com/

Some application scenarios

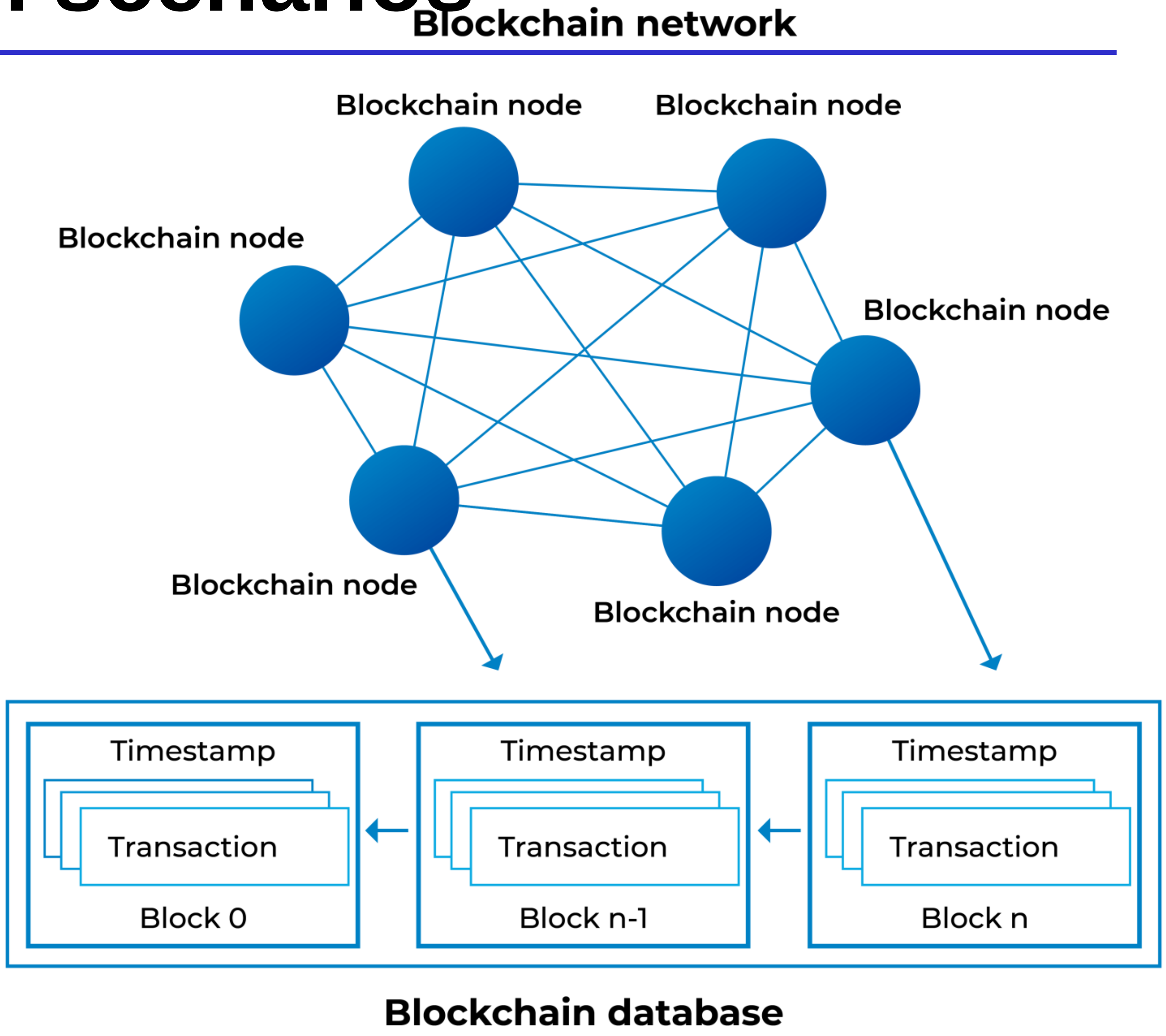
- **Distributed /cloud systems**



https://www.kiotviet.vn/

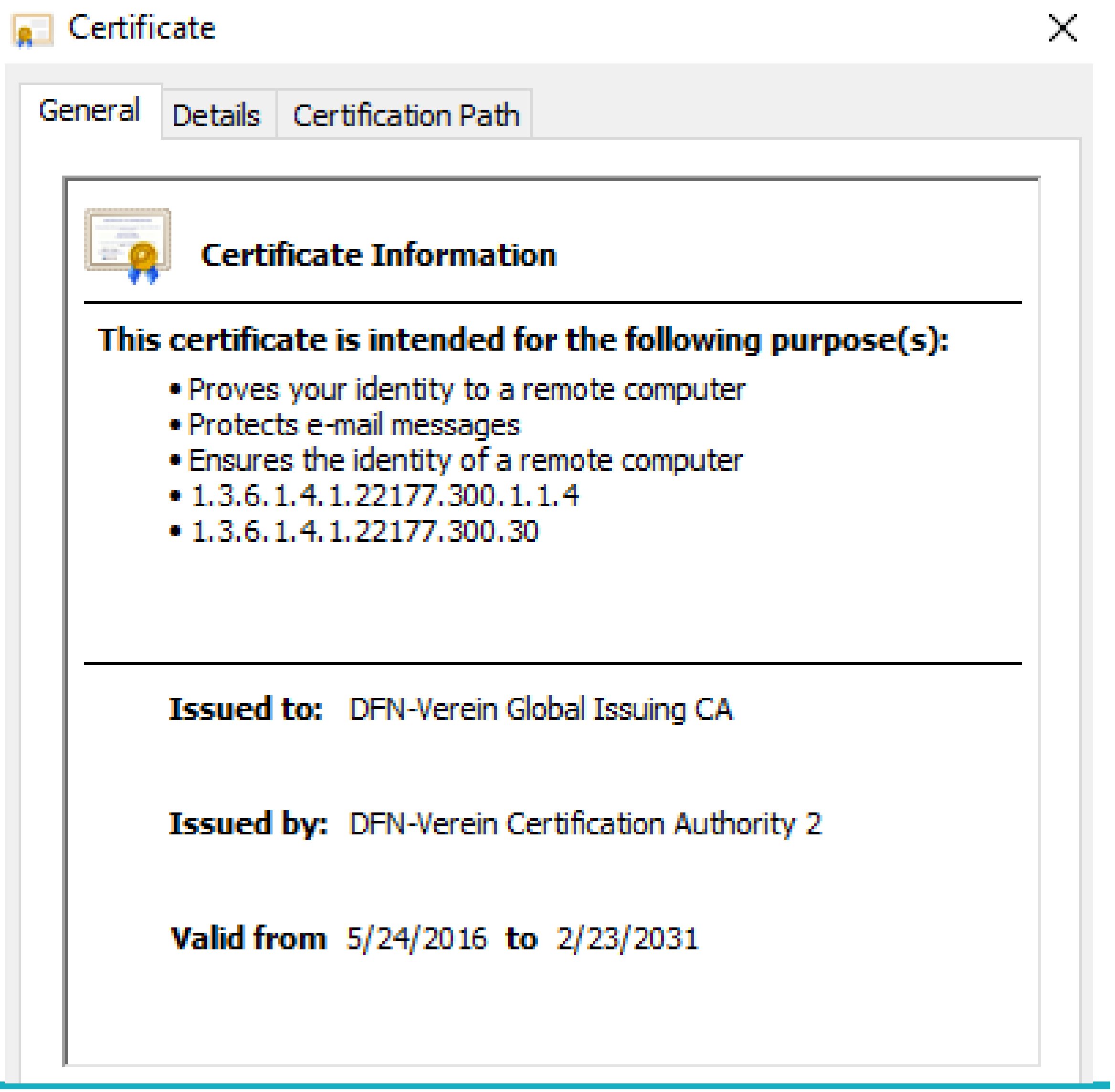
Some application scenarios

- **Blockchain networks**



Some application scenarios

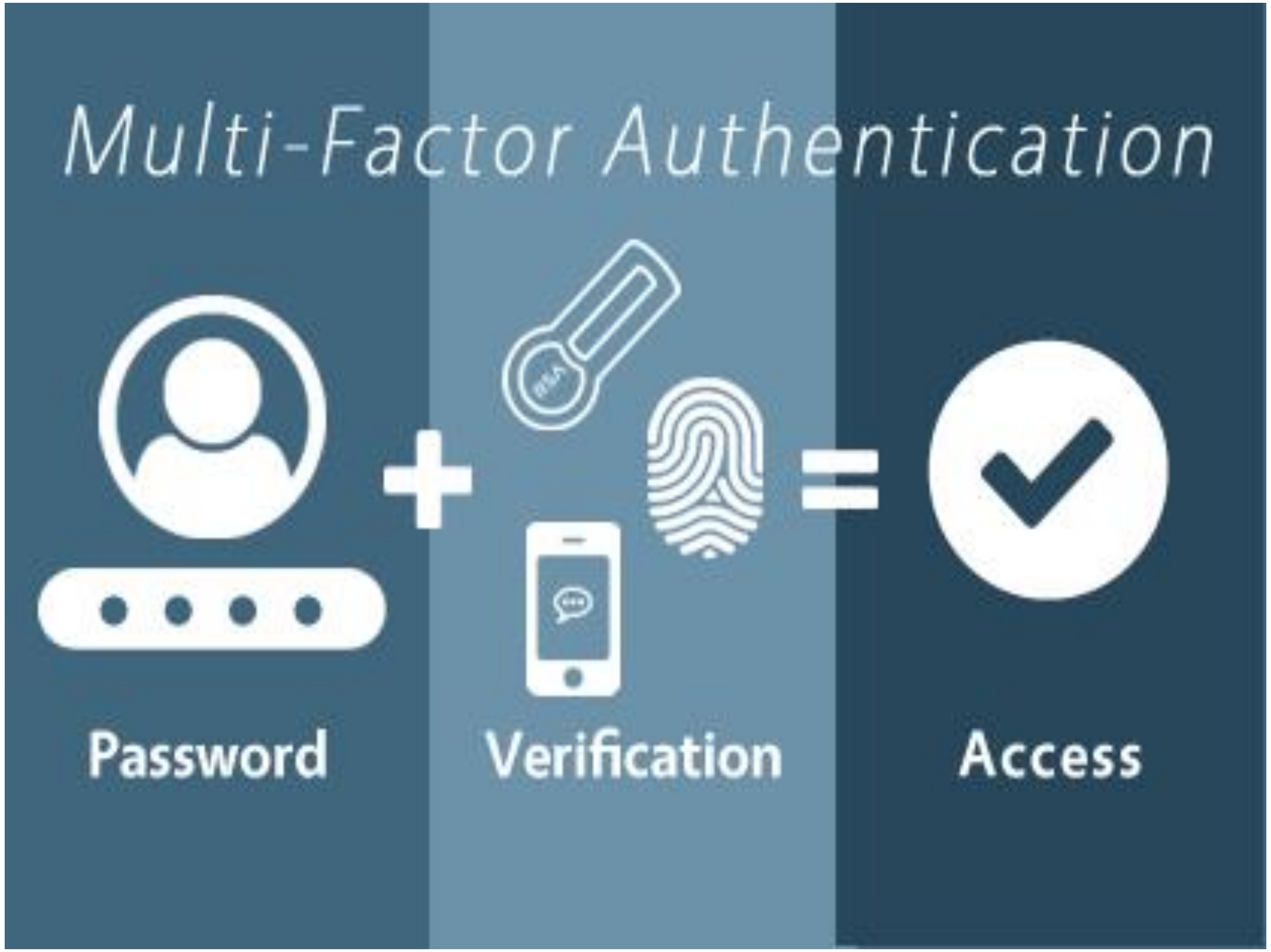
- **Digital signature (Digital certificates)**



https://vpcp.chinhphu.vn/van-ban-ban-hanh/171154.htm

Some application scenarios

- **Authentication**



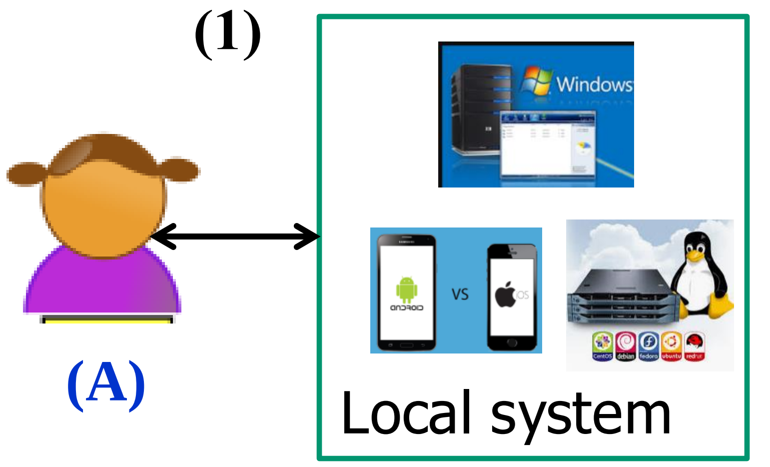
Some application scenarios

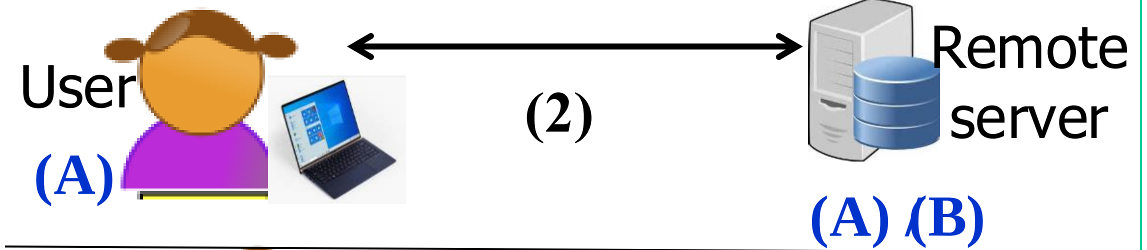
- **Mutual authentication**
  - Identification information?
  - Verification?
  - Exchange authentication factors?

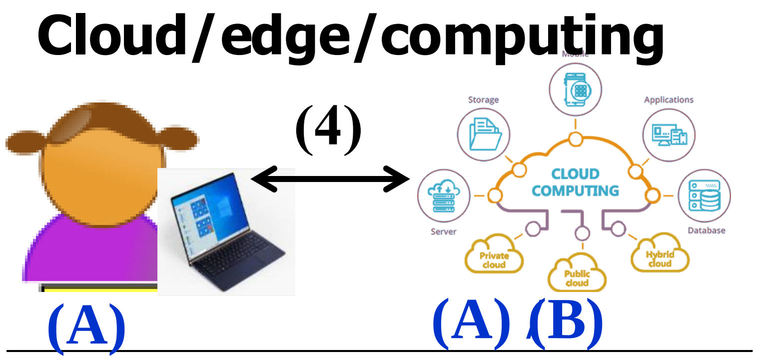
Solutions?

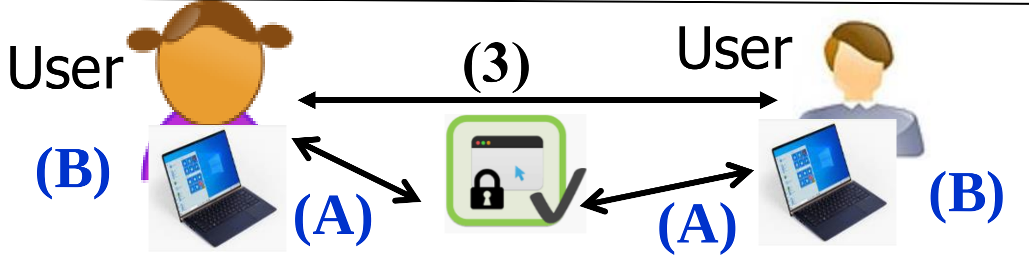
(A) Pre-shared secrets

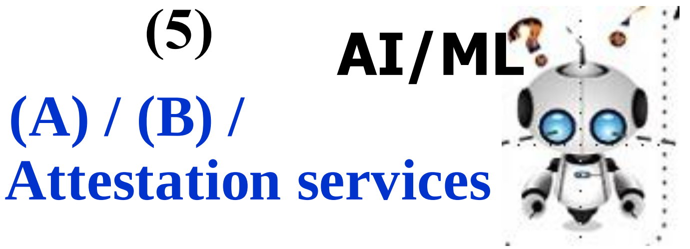
(B) Certificates (PKI)











Some application scenarios

- **Secure internet protocols**

Internet protocol suite

Application layer

BGP · DHCP(v6) · DNS · FTP · HTTP · HTTPS · IMAP · LDAP · MGCP · MQTT · NNTP · NTP · POP · PTP · ONC/RPC · RTP · RTSP · RIP · SIP · SMTP · SNMP · SSH · Telnet · TLS/SSL · XMPP · *more...*

Transport layer

TCP · UDP · DCCP · SCTP · RSVP · *more...*

Internet layer

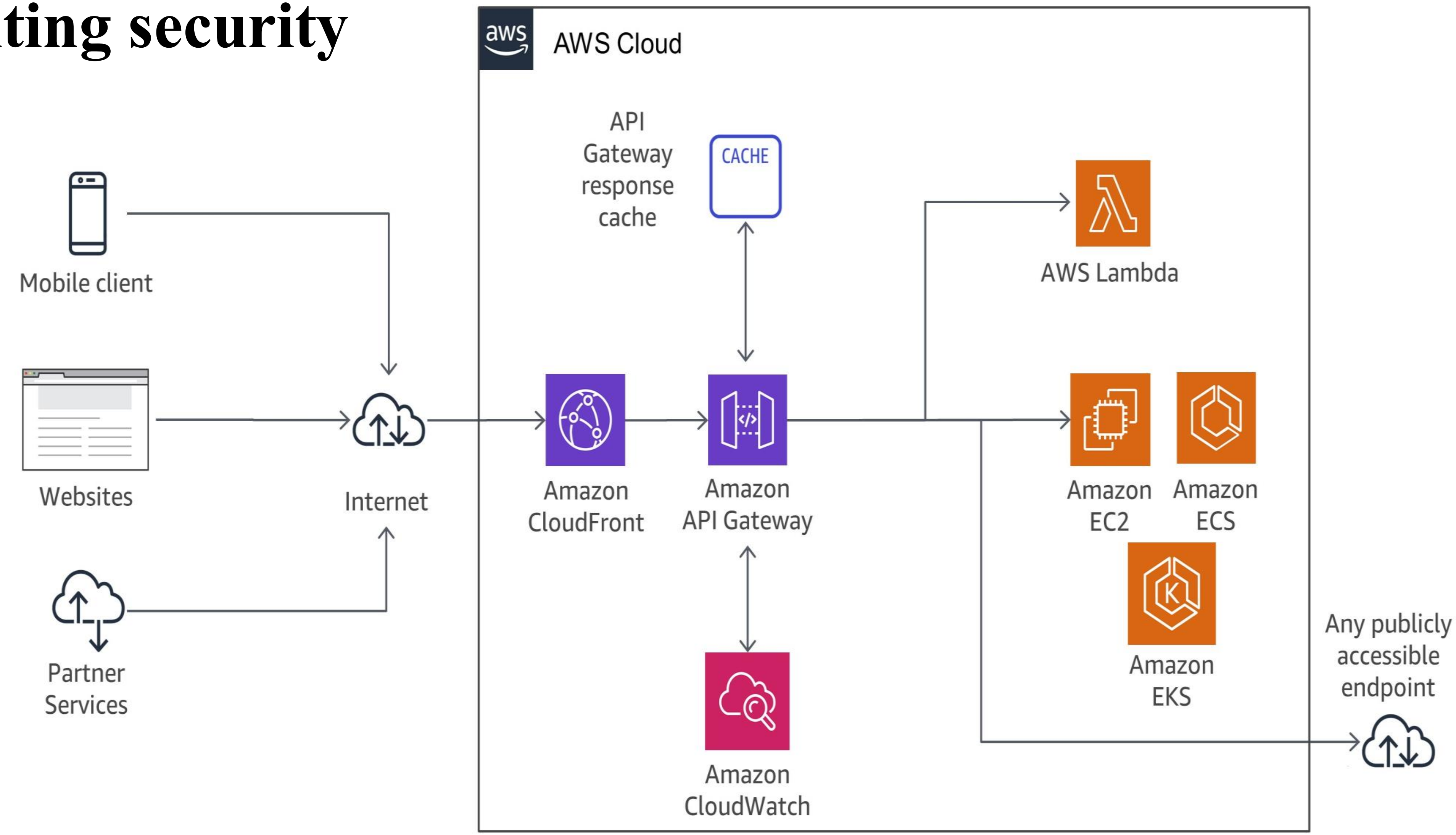
IP (IPv4 · IPv6) · ICMP(v6) · ECN · IGMP · **IPsec** · *more...*

Link layer

ARP · NDP · OSPF · Tunnels (L2TP) · PPP · MAC (Ethernet · Wi-Fi · DSL · ISDN · FDDI) *more...*

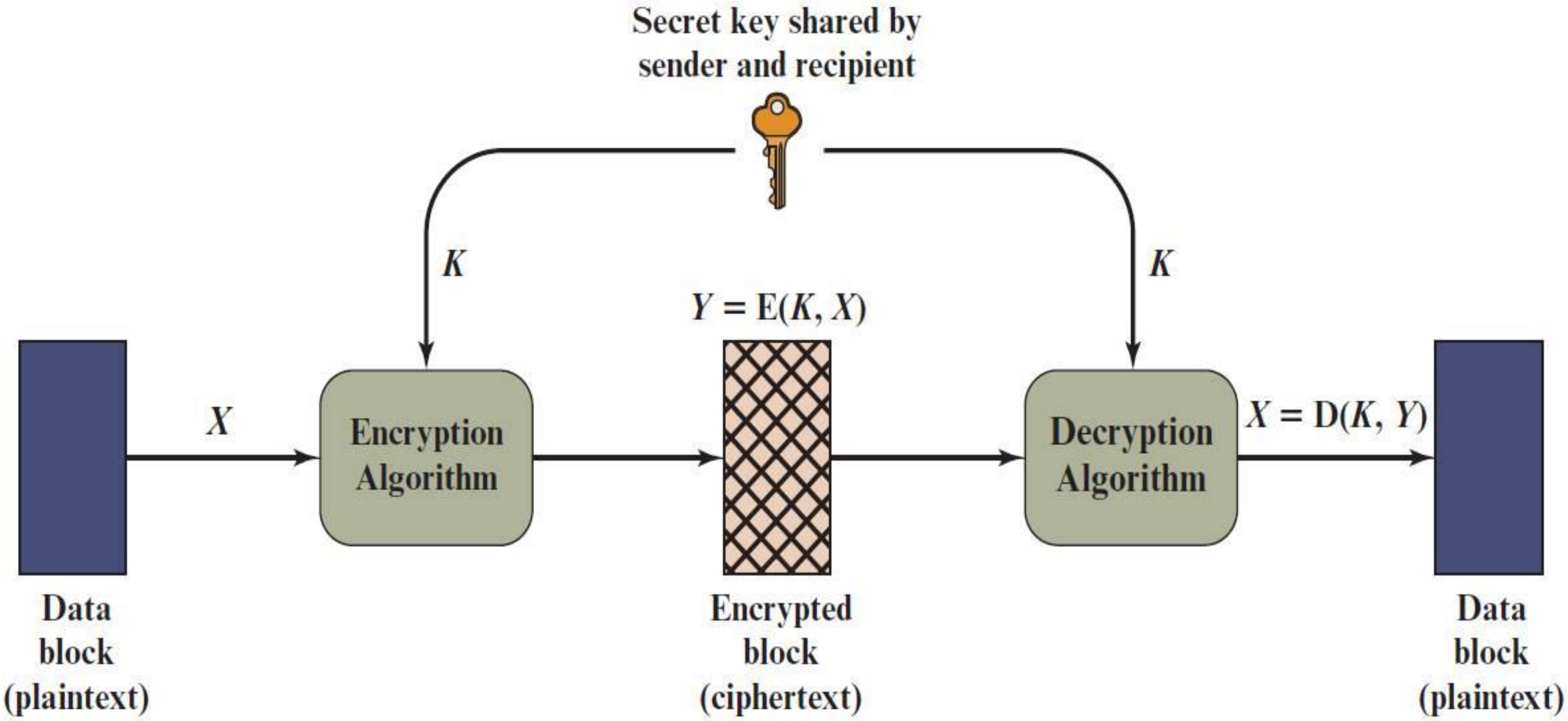
Some application scenarios

Cloud computing security



Classical cipher systems

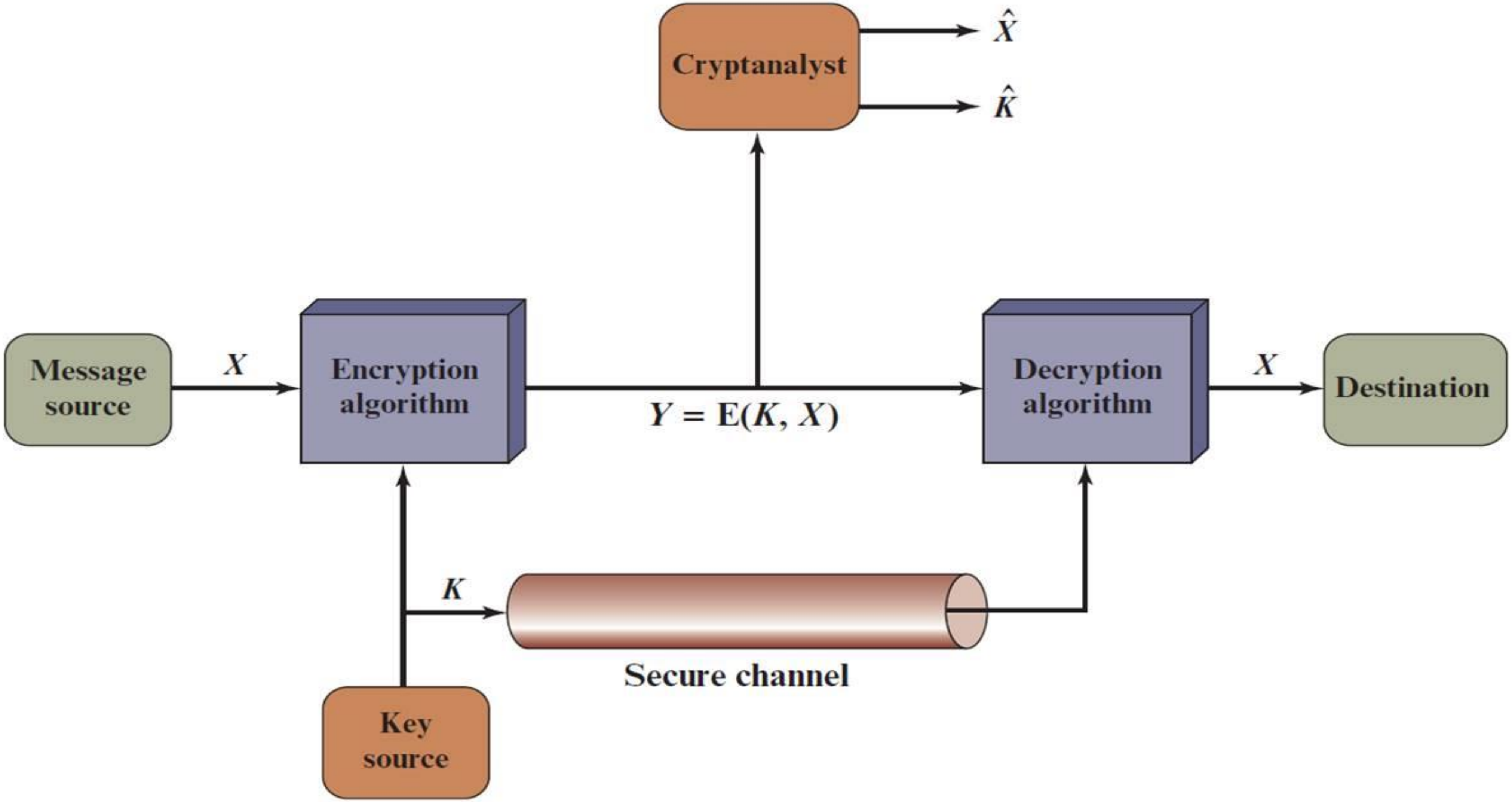
- Symmetric cipher cryptosystems



Classical cipher systems

- There are two requirements for secure use of conventional encryption:
  - A strong encryption algorithm
  - Sender and receiver must have obtained copies of **the secret key** in a secure fashion and must keep the key secure

Classical cipher systems



Clascical Ciphers

1. **Substitution** Techniques
   - Monoalphabetic cipher (1-1)
   - Polyalphabetic cipher (k-k)
2. **Transposition** Techniques

Monoalphabetic cipher

- **Monoalphabetic cipher**:

  1(letter, number) ←→ 1 fixed symbol in ciphertext;

1. **Caesar Cipher**

   A shift cipher that replaces each letter with another letter a fixed number of positions down the alphabet.

2. **ROT13**

   A special case of the Caesar cipher with a shift of 13. Because the alphabet has 26 letters, applying ROT13 twice returns the original text.

3. **Keyword Cipher**

   Constructs a substitution alphabet by first writing a keyword (omitting duplicate letters) and then appending the remaining unused letters of the alphabet in order.

Monoalphabetic cipher

4. **Simple Substitution Cipher**

   Uses a completely randomized permutation of the alphabet to replace each letter.

5. **Atbash Cipher**

   A specific substitution cipher where the alphabet is reversed. A maps to Z, B maps to Y, and so on.

6. **Affine Cipher**

   Uses a mathematical function of the form $E(x)=(ax+b) \bmod m$ (with m being the size of the alphabet) to substitute letters. The constants a and b serve as the keys, with a chosen so that it is coprime with m.

Monoalphabetic cipher

7. Homophonic Substitution Cipher

Instead of a one-to-one mapping, each plaintext letter is replaced by one of several possible ciphertext symbols. The set of possible symbols for each letter remains fixed, but a different symbol is chosen each time the letter appears, which can help mask letter frequencies.

8. Pigpen Cipher

Uses a set of geometric symbols (often based on a grid or "pigpen") to substitute for letters. Each letter corresponds to a specific symbol derived from a simple visual pattern.

Monoalphabetic cipher

(1) Caesar Cipher

- Simplest and earliest known use of a substitution cipher
- Used by Julius Caesar

Key

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	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

plain: MEET ME AFTER THE TOGA PARTY

cipher: JBBQ JB XCQBO QEB QLDX MXOQV

Caesar Cipher Algorithm

- 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

- Algorithm can be expressed as:

$$c = E(3, p) = (p + 3) \bmod (26)$$

- A shift may be of any amount, so that 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

$$p = D(k, C) = (C - k) \bmod 26$$

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

Need large number of keys!

Monoalphabetic cipher

(2) Monoalphabetic substitution

- **Permutation**
  - Of a finite set of elements *S* is an ordered sequence of all the elements of *S* , with each element appearing exactly once

Monoalphabetic cipher

(2) Monoalphabetic substitution

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

EX:

MEET ME AT OUR SPOT

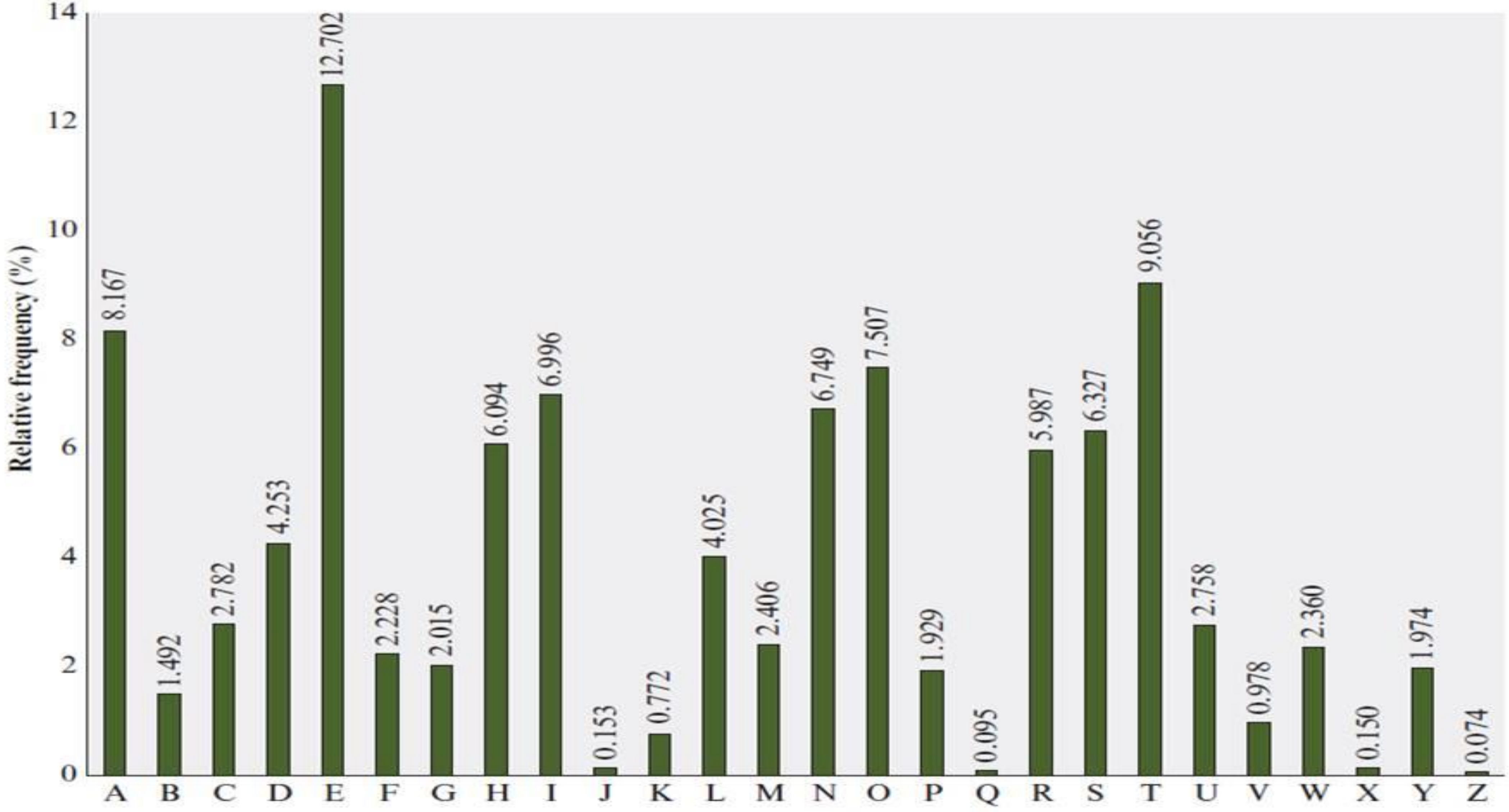
↓

DTTM DT AM GWK LHGM

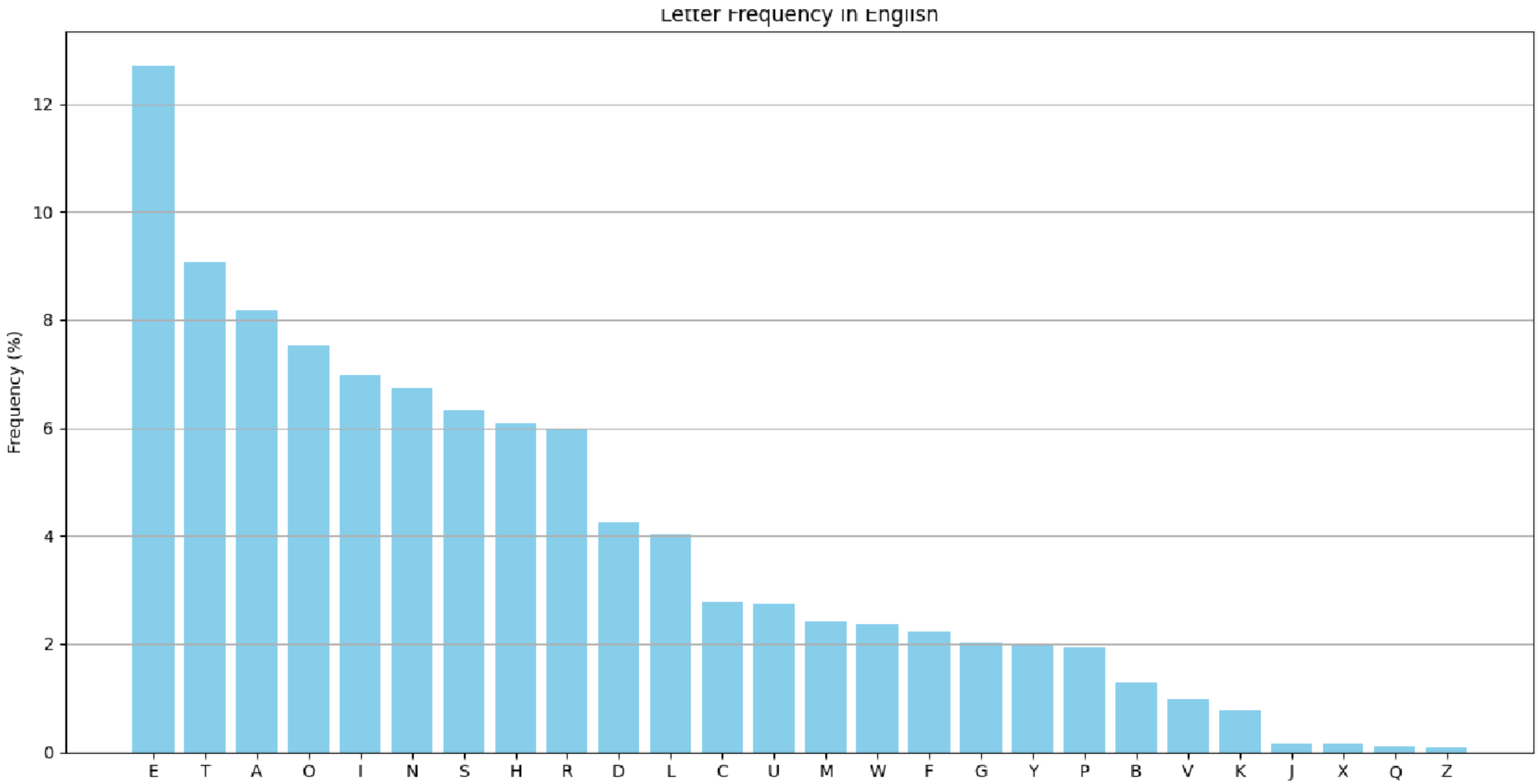
If the “cipher” line can be any permutation of the 26 alphabetic characters, then there are 26! or greater than 4×10^{26} possible keys

This is 10 orders of magnitude greater than the key space for DES

Relative Frequency of Letters in English Text



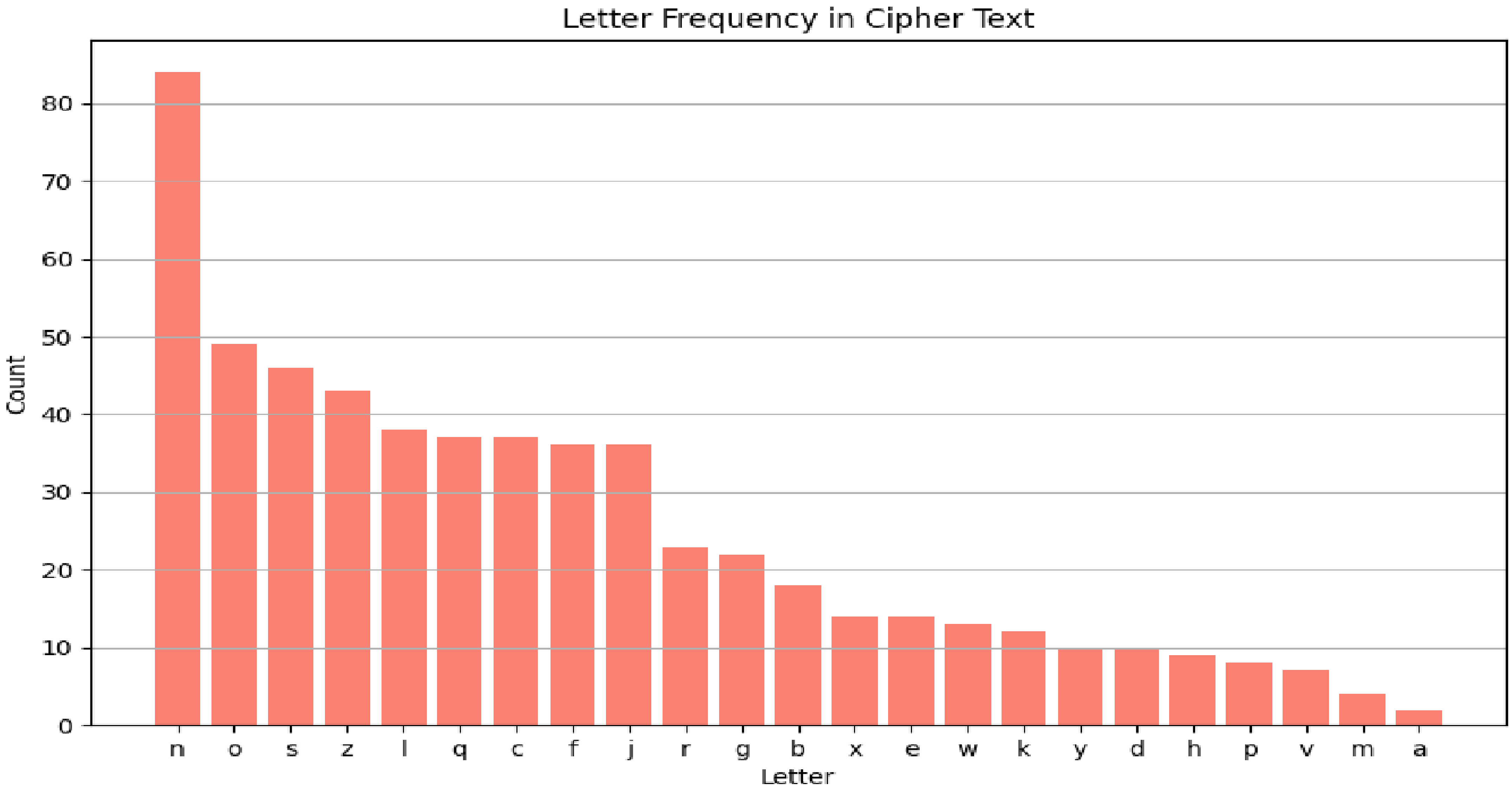
Relative Frequency of Letters in English Text



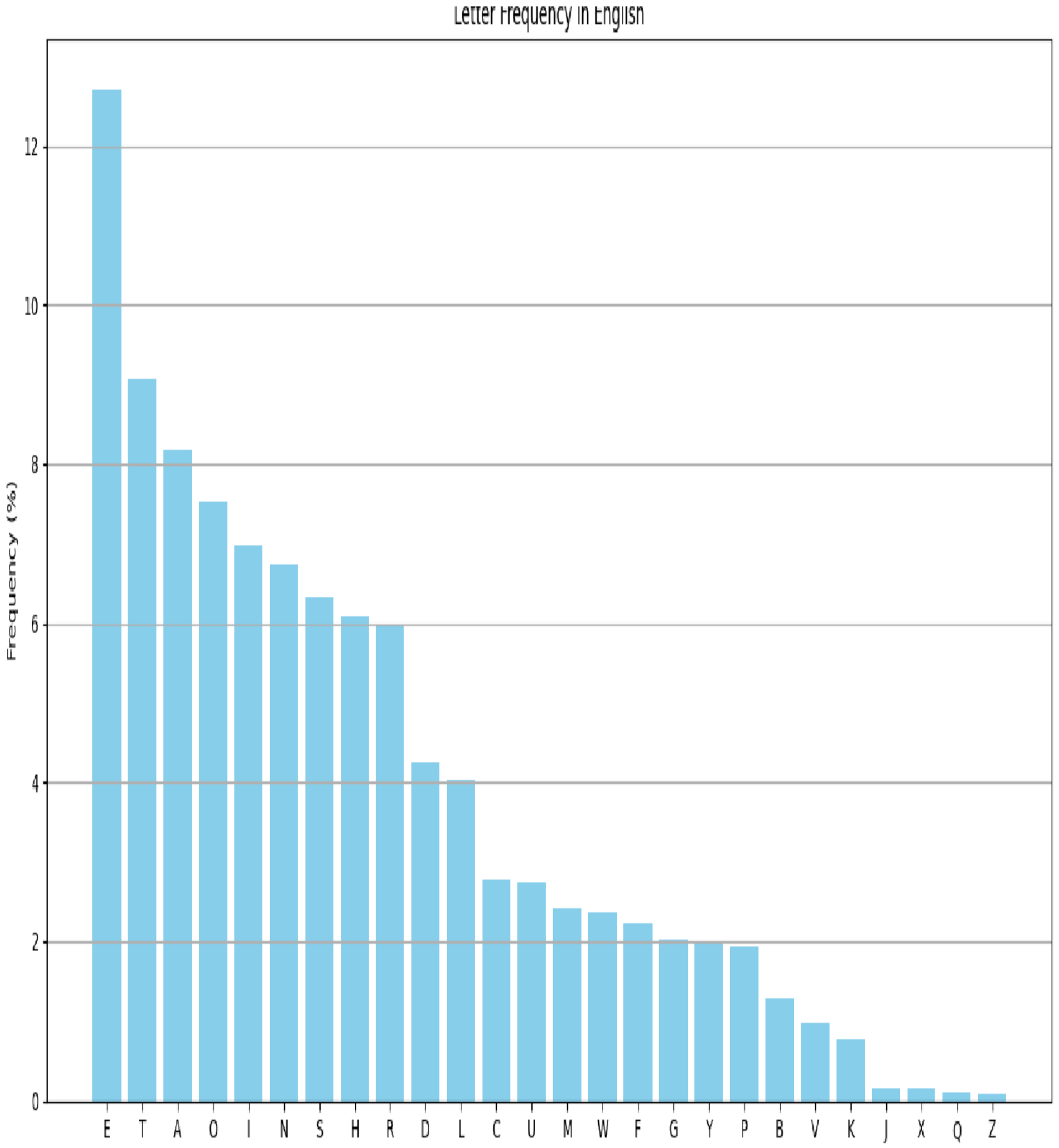
Cryptanalysis on monoalphabetic cipher

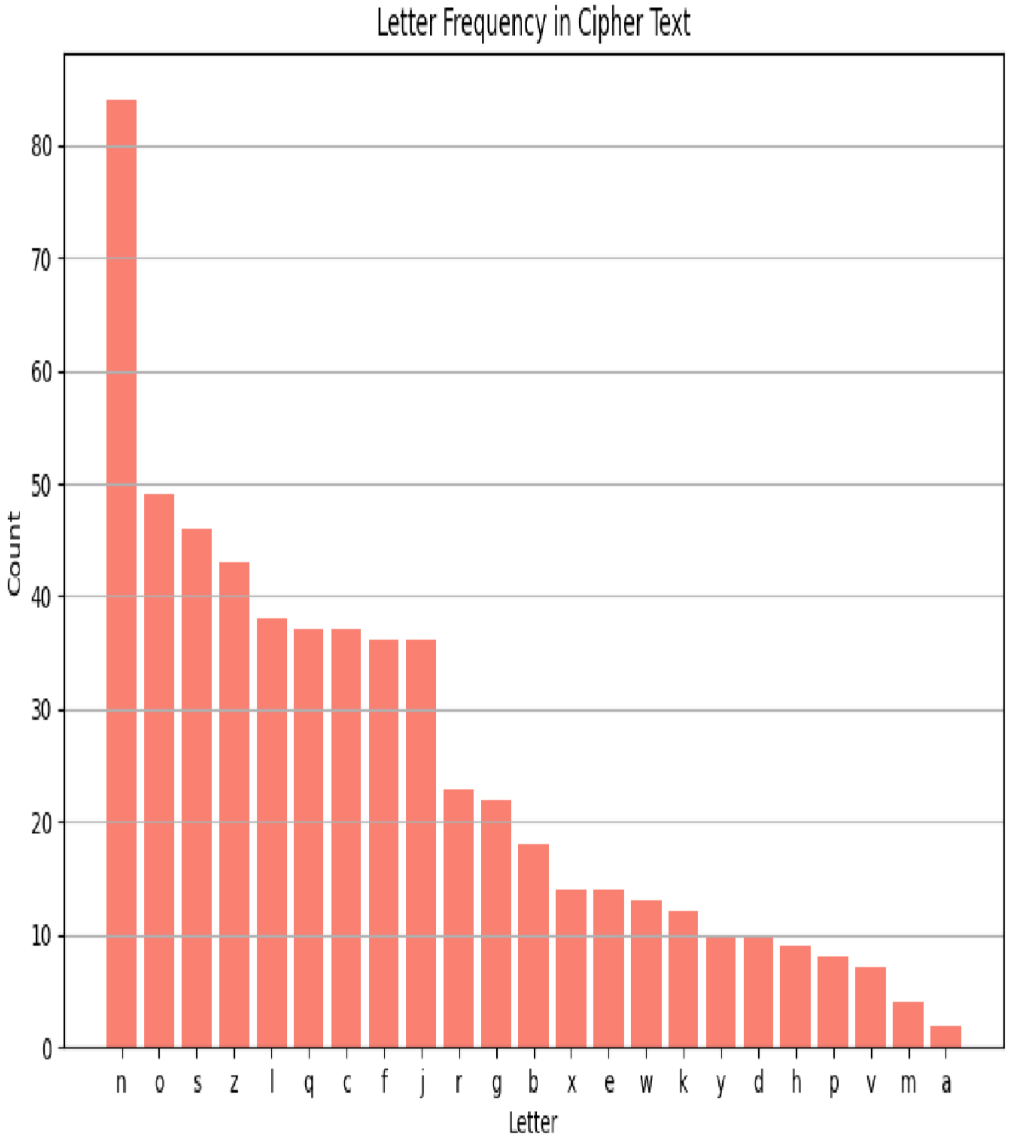
hzsrnqc klyy wqc flo mflwf ol zqdn nsoznj wskn lj xzsrbjnf, wzsxz gqv zqhhnf ol ozn glco zlfnco hnlhrn; nsoznj jnrqosdnc lj fnqj kjsnfbc, wzsxz sc xnjoqsfrv gljn efeceqr. zn rsdnb qrlfn sf zsc zlecn sf cqdsrrn jlw, wzsoznj flfn hnfnojqonb. q csfyrn blgncosx cekksxnb ol cnjdn zsg. zn pjnqmkqconb qfb bsfnb qo ozn xrep, qo zlejc gqozngqosxqrrv ksanb, sf ozn cqgn jllg, qo ozn cqgn oqprn, fndnj oqmsfy zsc gnqrc wsoz loznj gngpnjc, gexz rncc pjsfysfy q yenco wsoz zsg; qfb wnfo zlgn qo naqxorv gsbfsyzo, lfrv ol jnosjn qo lfxn ol pnb. zn fndnj ecnb ozn xlcv xzqgpnjc wzsxz ozn jnkljg hjldsbnc klj soc kqdlejnb gngpnjc. zn hqccnb onf zlejc leo lk ozn ownfov-klej sf cqdsrrn jlw, nsoznj sf crnnhsfy lj gqmsfy zsc olsrno.

Relative Frequency of Letters in Ciphertext



Relative Frequency of Plaintext and Ciphertext





Monoalphabetic Ciphers

- Easy to break because they reflect the frequency data of the original alphabet
- Countermeasure is to provide multiple substitutes (homophones) for a single letter

Polyalphabetic Cipher

- Polyalphabetic Cipher is a substitution cipher in which the cipher alphabet for the plain alphabet may be different at different places during the encryption process;
  - Playfair Cipher (Replace 2 characters by 2 charector)
  - Hill Cipher
  - Vigenere Cipher

Polyalphabetic Cipher

- Polyalphabetic Cipher is a substitution cipher in which the cipher alphabet for the plain alphabet may be different at different places during the encryption process;
  - Playfair Cipher
  - Hill Cipher
  - Vigenere Cipher

Polyalphabetic Ciphers

- Polyalphabetic substitution cipher
  - Improves on the simple monoalphabetic technique by using different monoalphabetic substitutions as one proceeds through the plaintext message

- All these techniques have the following features in common:
  - A set of related monoalphabetic substitution rules is used
  - A key determines which particular rule is chosen for a given transformation

(3) Playfair Cipher

- Best-known multiple-letter encryption cipher (two→ two)
- Treats digrams in the plaintext as single units and translates these units into ciphertext digrams
- Based on the use of a 5 × 5 matrix of letters constructed using a keyword
- Invented by British scientist Sir Charles Wheatstone in 1854
- Used as the standard field system by the British Army in World War I and the U.S. Army and other Allied forces during World War II

Polyalphabetic Cipher

Playfair Key Matrix

- Fill in letters of keyword (minus duplicates) from left to right and from top to bottom, then fill in the remainder of the matrix with the remaining letters in alphabetic order
- 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

Playfair encryption

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

Plaintext: "Hide the gold in the tree stump"

Plaintext diagram:

HI DE TH EG OL DI NT HE TR EX ES TU MP

Ciphertext diagram:

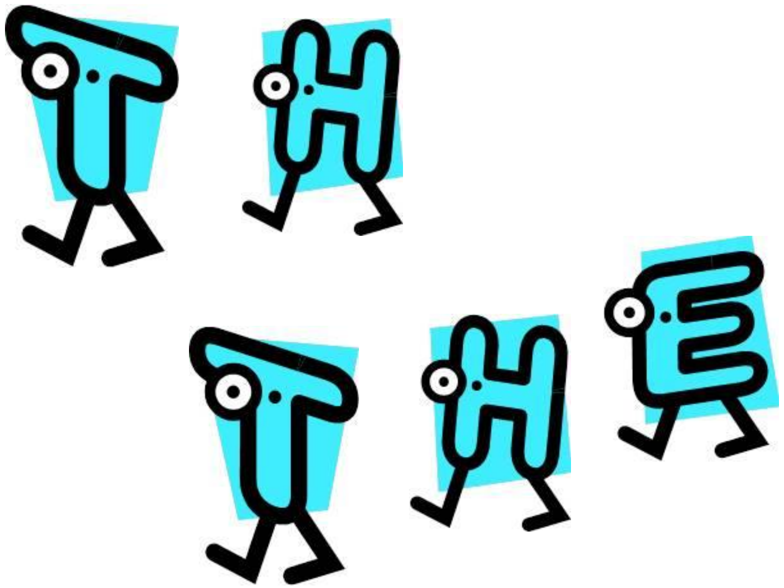
BF CK PD FI .. .. DZ

https://en.wikipedia.org/wiki/Playfair_cipher

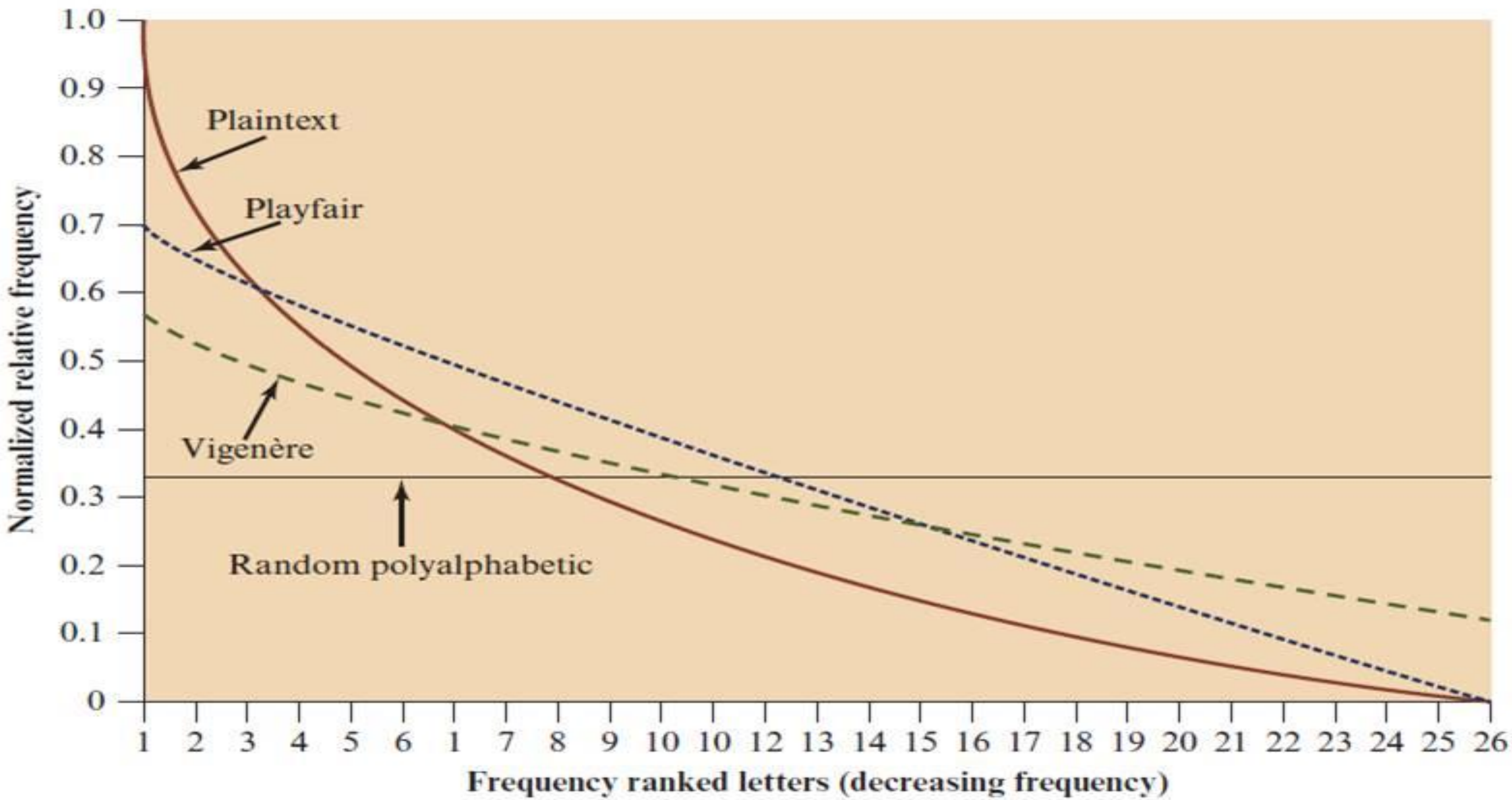
Polyalphabetic Cipher

Cryptoanalys Playfair cipher

- Digram
  - Two-letter combination
  - Most common is *th*
- Trigram
  - Three-letter combination
  - Most frequent is *the*



Relative Frequency of Occurrence of Letters



Polyalphabetic Cipher

(4) Hill Cipher

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
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

- Developed by the mathematician Lester Hill in 1929
- Strength is that it completely hides single-letter frequencies
  - The use of a larger matrix hides more frequency information
  - A 3 x 3 Hill cipher hides not only single-letter but also two-letter frequency information
- Strong against a ciphertext-only attack but easily broken with a known plaintext attack

$$C = K.P \bmod 26$$

$$\begin{pmatrix} k_{1,1} & k_{1,2} & k_{1,3} \\ k_{2,1} & k_{2,2} & k_{2,3} \\ k_{3,1} & k_{3,2} & k_{3,3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \bmod 26$$

(5) Vigenère Cipher

- Best known and one of the simplest polyalphabetic substitution ciphers
- In this scheme the set of related monoalphabetic substitution rules consists of the 26 Caesar ciphers with shifts of 0 through 25
- Each cipher is denoted by a key letter which is the ciphertext letter that substitutes for the plaintext letter a

https://en.wikipedia.org/wiki/Vigen%C3%A8re_cipher

Example of Vigenère Cipher

- To encrypt a message, a key is needed that is as long as the message
- Usually, the key is a repeating keyword
- For example, if the keyword is *deceptive*, the message "we are discovered save yourself" is encrypted as:

plaintext: wearediscoveredsaveyourself

key: deceptivedeceptivedeceptive

ciphertext: ??

Vigenère Cipher

- Vigenère matrix

	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

Vigenère Autokey System

- Example:

  key: deceptivewearediscoveredsav

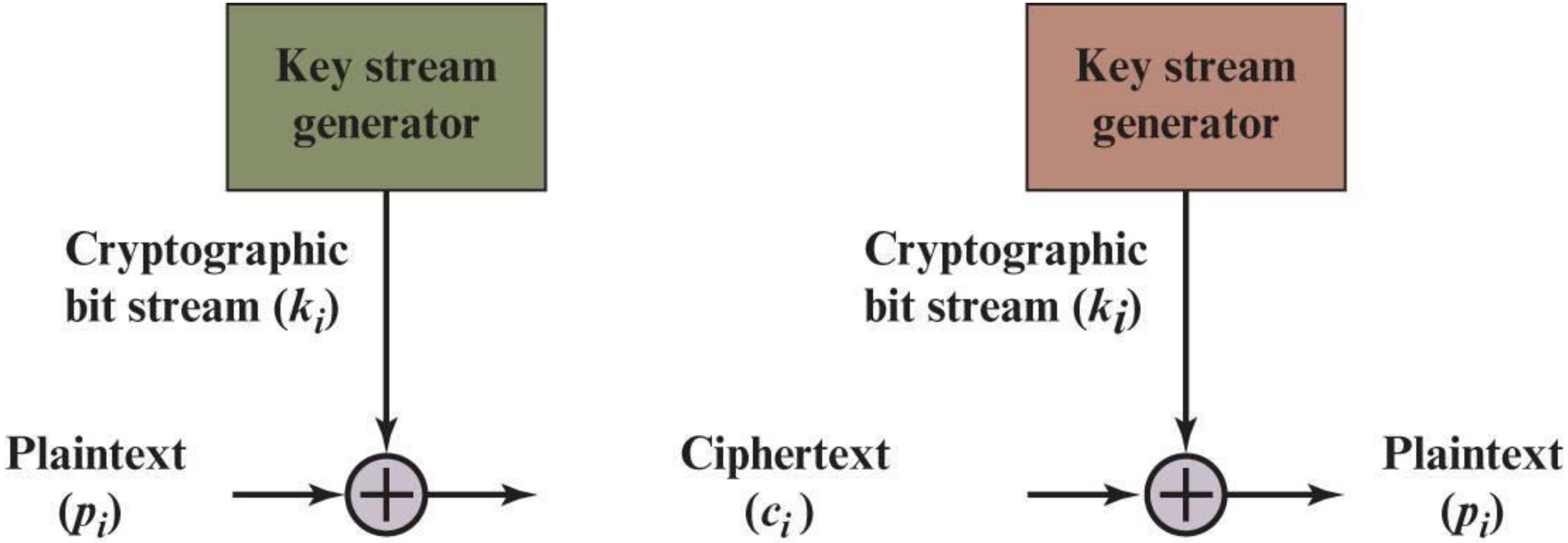
  plaintext: wearediscoveredsaveyourself

  ciphertext: ZICVTWQNGKZEIIGASXSTSLVVWLA

- Even this scheme is vulnerable to cryptanalysis
  - Because the key and the plaintext share the same frequency distribution of letters, a statistical technique can be applied

Stream ciper

(6) Vernam Cipher



https://en.wikipedia.org/wiki/Gilbert_Vernam

One-Time Pad

- Improvement to Vernam cipher proposed by an Army Signal Corp officer, Joseph Mauborgne
- Use a **random key that is as long as the message** so that the key need not be repeated
- Key is used to encrypt and decrypt a single message and then is discarded
- Each new message requires a new key of the same length as the new message

- **Scheme is unbreakable**
  - Produces random output that bears no statistical relationship to the plaintext
  - Because the ciphertext contains no information whatsoever about the plaintext, there is simply no way to break the code

Difficulties

- The one-time pad offers complete security but, in practice, has two fundamental difficulties:
  - There is the practical problem of making large quantities of random keys
    - Any heavily used system might require millions of random characters on a regular basis
  - Mammoth key distribution problem
    - For every message to be sent, a key of equal length is needed by both sender and receiver
- Because of these difficulties, the one-time pad is of limited utility
  - Useful primarily for low-bandwidth channels requiring very high security
- The one-time pad is the only cryptosystem that exhibits *perfect secrecy* (see Appendix F)

Transposition ciphers

Goals: scrambles the positions of characters

(1) Rail fence cipher

(2) Columnar Transposition Cipher

https://en.wikipedia.org/wiki/Transposition_cipher

Transposition cipher

(1) Rail fence cipher

- Simplest transposition cipher
- Plaintext is written down as a sequence of diagonals and then read off as a sequence of rows
- To encipher the message "meet me after the toga party" with a rail fence of depth 2, we would write:

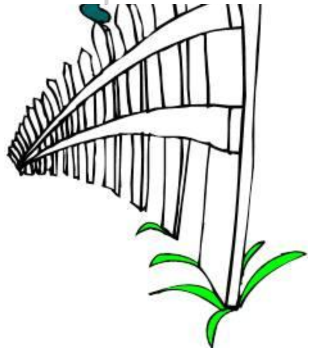
m	e	m	a	t	r	h	t	g	p	r	y
	e	t	e	f	e	t	e	o	a	a	t

Ciphertext

Encrypted message is:

MEMATRHTGPRYETEFETEOAAT

https://en.wikipedia.org/wiki/Rail_fence_cipher



Columnar Transposition Cipher

- Is a more complex transposition
- Write the message in a rectangle, row by row, and read the message off, column by column, but permute the order of the columns
  - The order of the columns then becomes the key to the algorithm

Key:	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

Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ

Summary

- Present an overview of the main concepts of symmetric cryptography
- Explain the difference between cryptanalysis and brute-force attack
- Understand the operation of a monoalphabetic substitution cipher
- Understand the operation of a polyalphabetic cipher
- Present an overview of the Hill cipher

Modern cipher systems