Classical Encryption Techniques

Dr. Ruba Al Omari

EECS 3481 – Applied Cryptography

Topics

- Symmetric Cipher Model
- Substitution Techniques
- Transposition Techniques

Today's lecture slides were prepared from "Cryptography and Network Security", 8/e, by William Stallings, Chapter 3 – "Classical Encryption Techniques".

Topics

- Symmetric Cipher Model
- Substitution Techniques
- Transposition Techniques

Symmetric Cipher Model

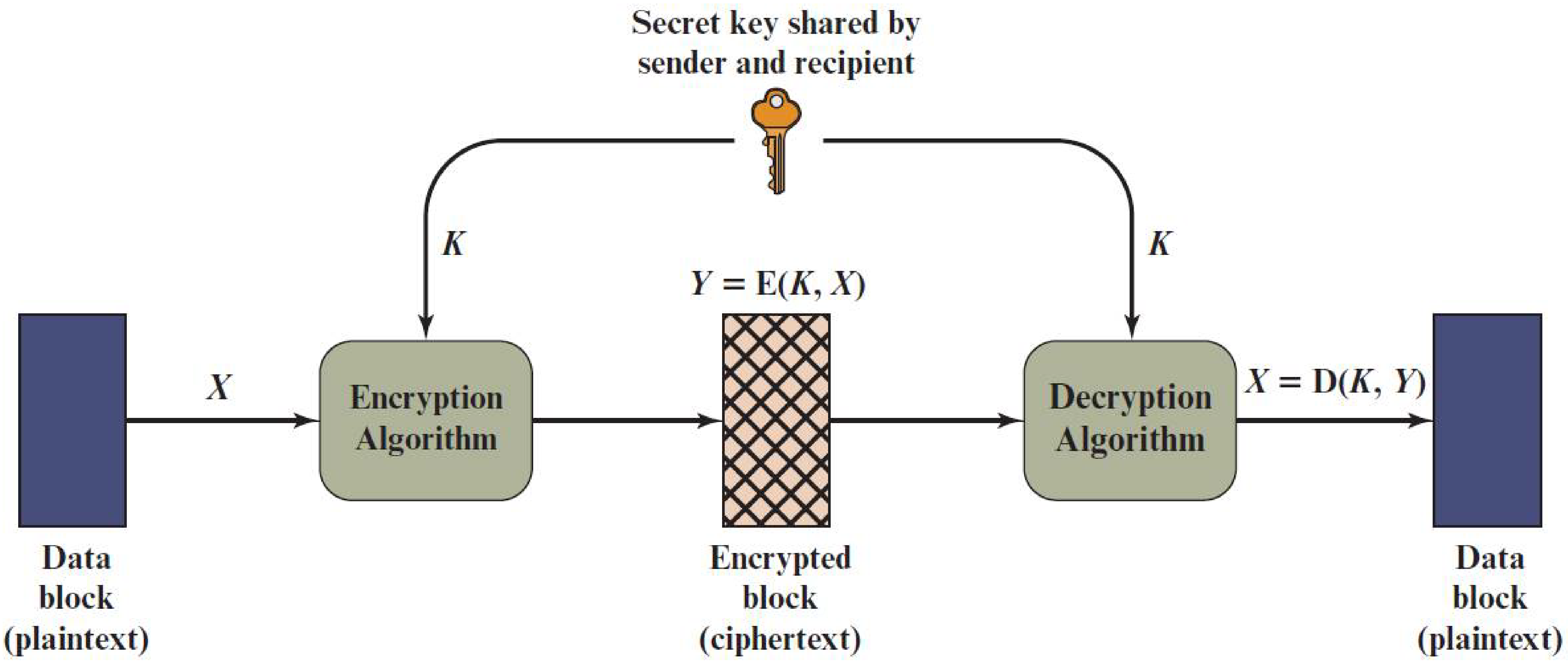


Figure Source: Cryptography and Network Security, 8th Edition, by William Stallings

Symmetric Cipher Model

- There are two requirements for secure use of conventional encryption:

1. A strong encryption algorithm: The opponent should be unable to decrypt ciphertext or discover the key even if they have a number of ciphertexts/plaintexts pairs.

2. Sender and receiver must have obtained copies of the secret key in a secure fashion and must keep the secret key secure.
   - We do not need to keep the algorithm secret; we need to keep only the key secret.

Symmetric Cryptosystem

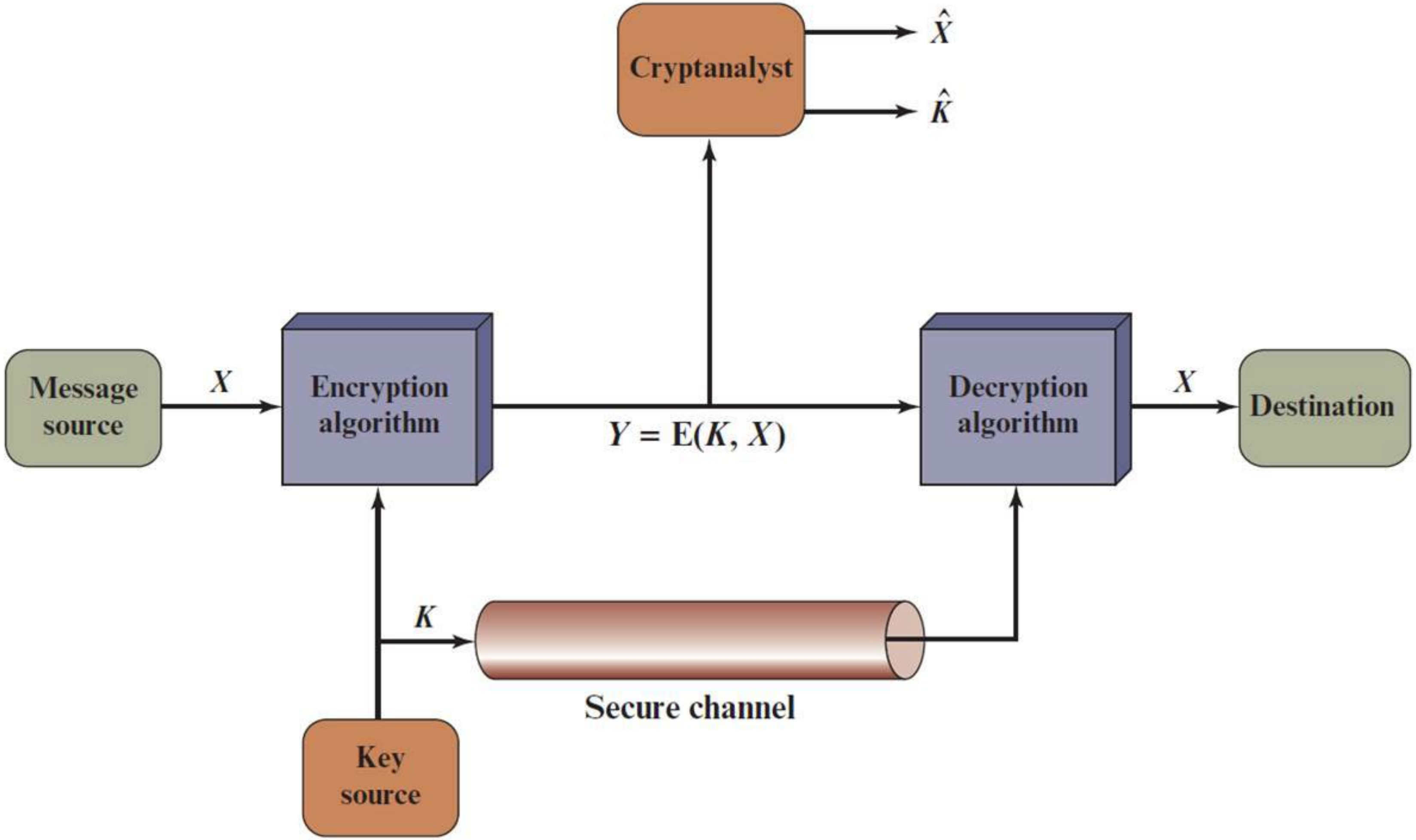
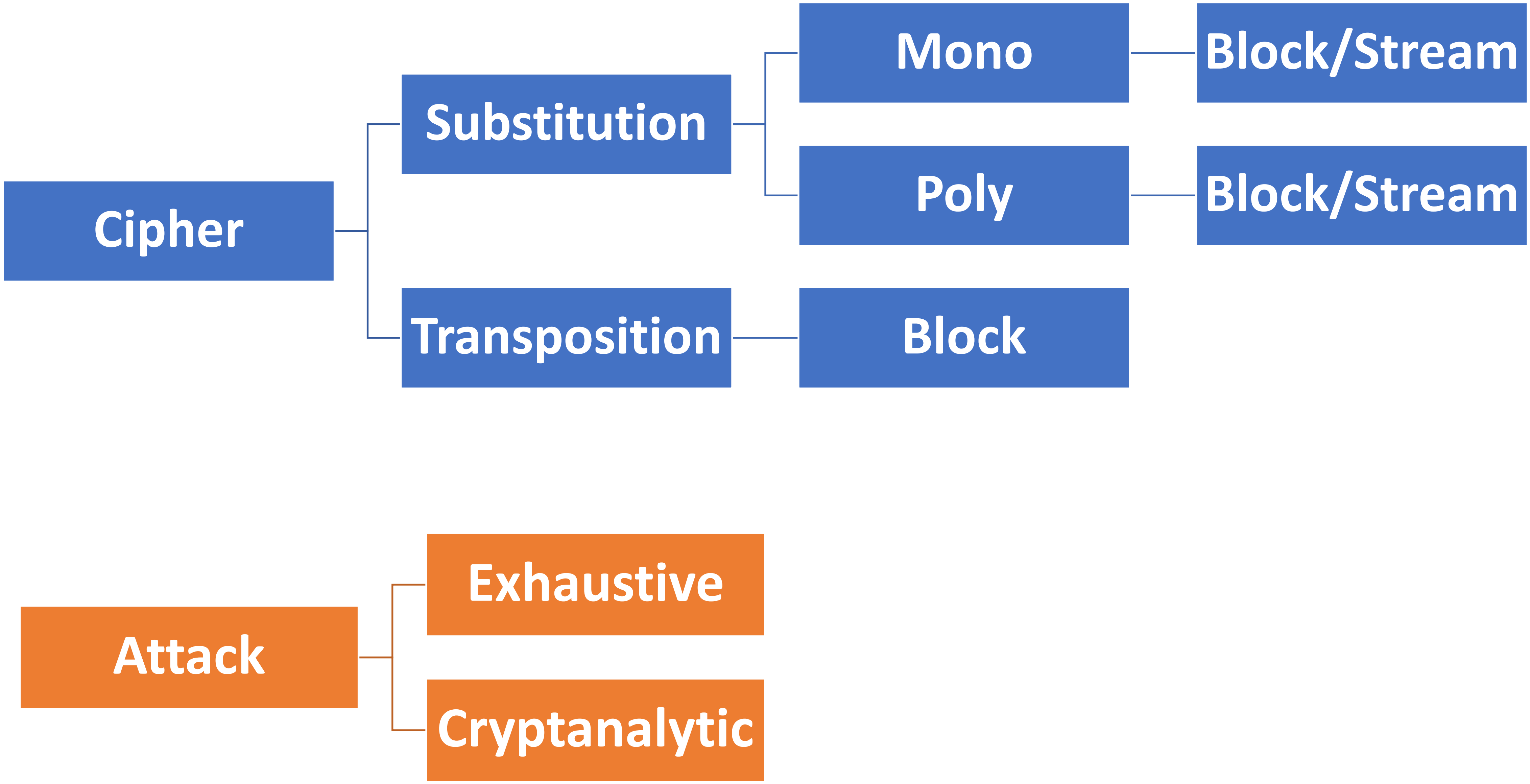


Figure Source: Cryptography and Network Security, 8th Edition, by William Stallings


Cipher
Substitution
Transposition
Mono
Poly
Block
Block/Stream
Block/Stream
Attack
Exhaustive
Cryptanalytic

Cipher Techniques

Cipher

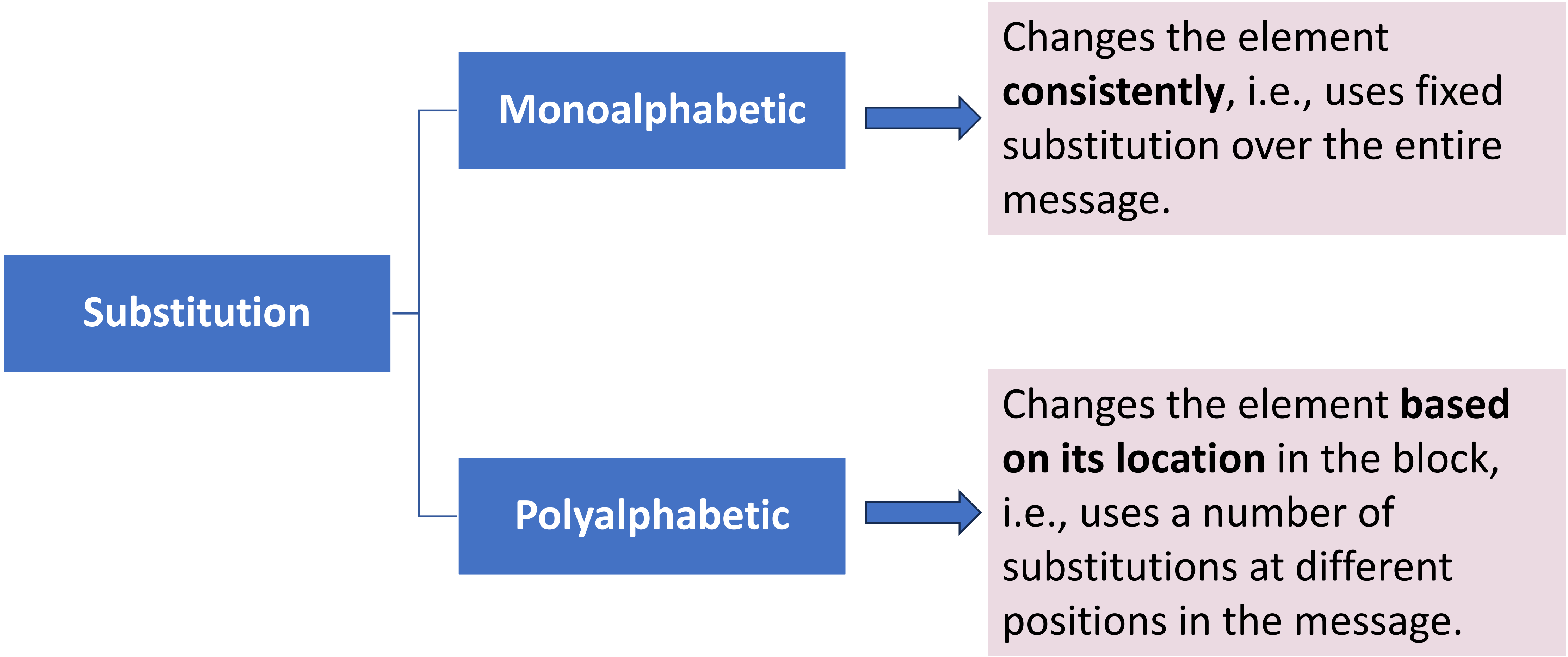
Substitution

Changes the element: the letters of plaintext are replaced by other letters or by numbers or symbols.

Transposition

Keeps the element but **permutes its position**.

Substitution Techniques



Block vs. Stream

Cipher

Block

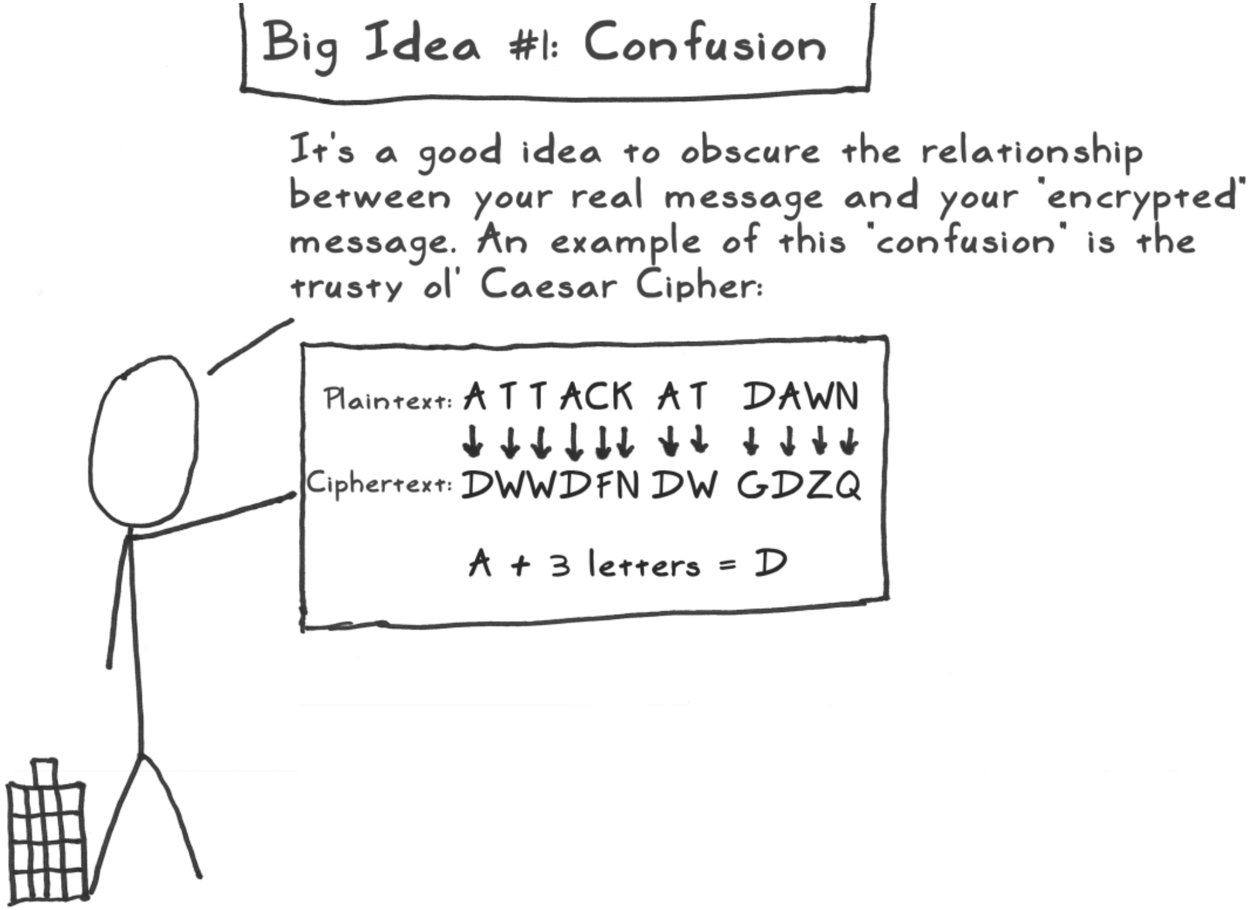
Simple: Encrypt **fixed-size blocks** at a time (typically 64 or 128 bits).

Stream

More complex: Encrypt **bit by bit** or byte by byte at a time.
Suitable for real-time communication or data transmission where data length is unknown or variable.

Confusion

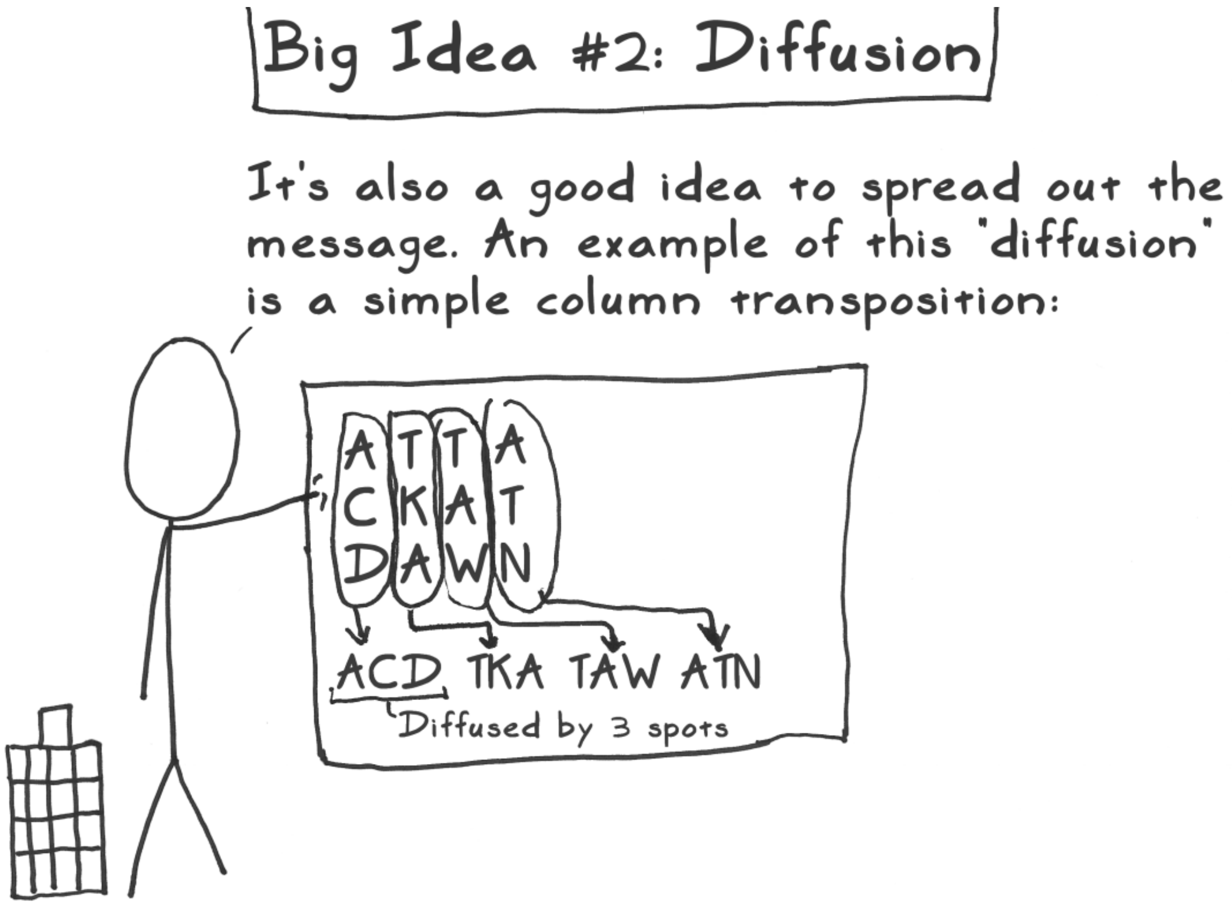
- **Confusion** **hides patterns** between plaintext and ciphertext, making it hard for attackers to guess the key or gain information about the plaintext from the ciphertext.



Source: https://www.moserware.com/assets/stick-figure-guide-to-advanced/aes_act_2_scene_02_confusion_1100.png

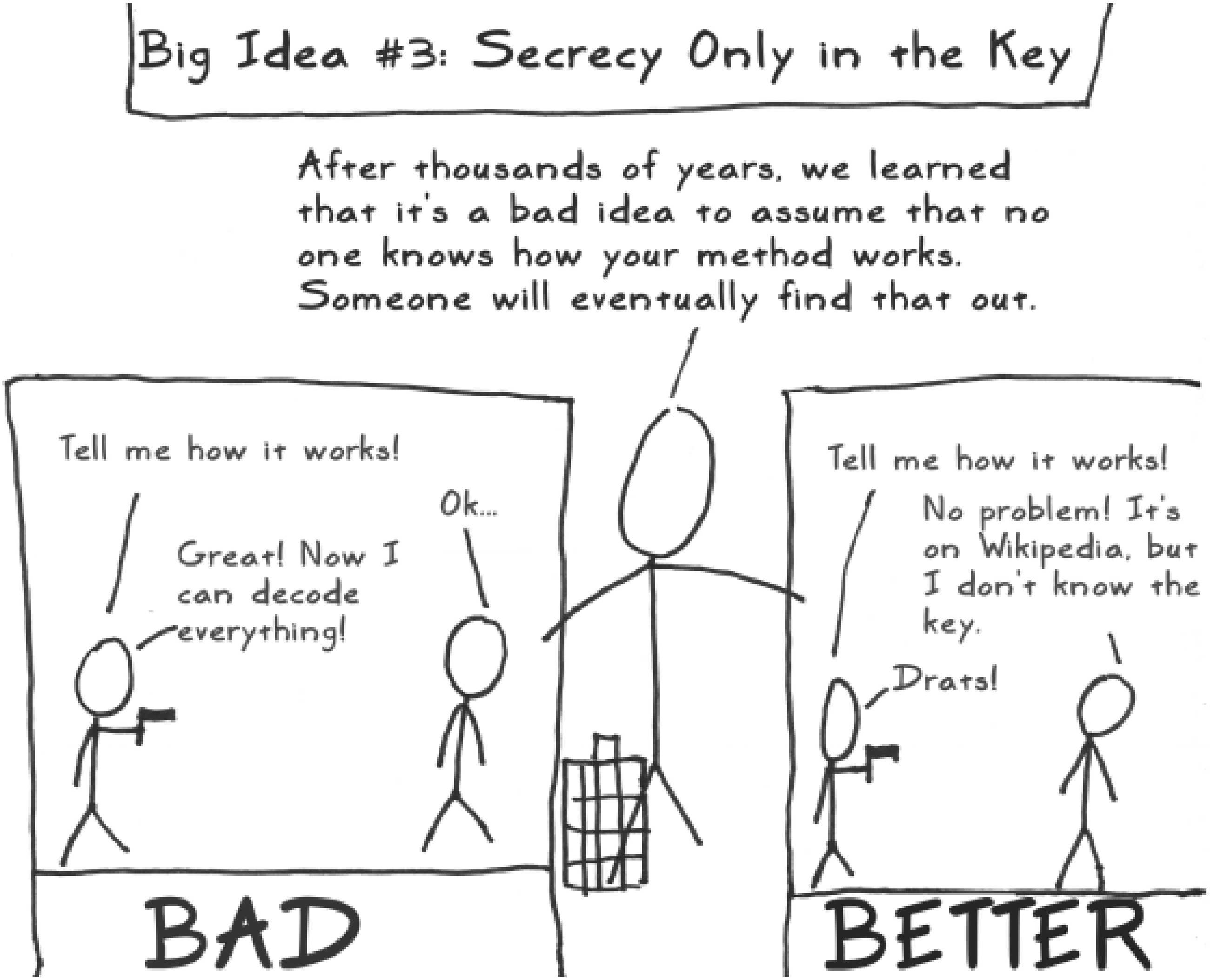
Diffusion

- **Diffusion** **spreads the influence** of individual plaintext elements across the entire ciphertext, through rearranging the bits or bytes of the plaintext in a complex and systematic manner.



Source: https://www.moserware.com/assets/stick-figure-guide-to-advanced/aes_act_2_scene_03_diffusion_1100.png

Secrecy Only in the Key



Source: http://www.moserware.com/assets/stick-figure-guide-to-advanced/aes_act_2_scene_04_key_secrecy_1100.png

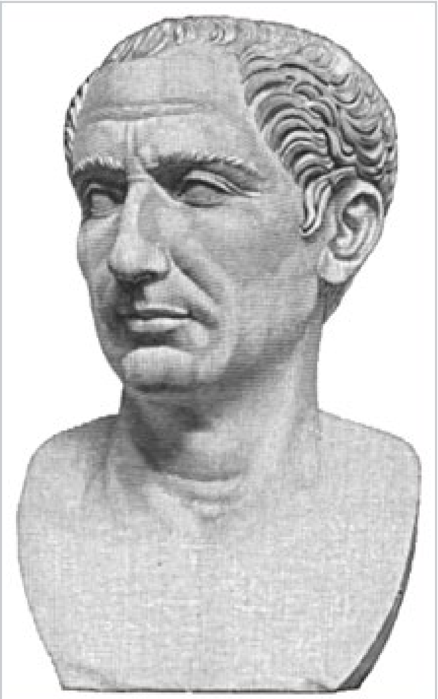
Topics

- Symmetric Cipher Model
- Substitution Techniques
- Transposition Techniques

Classical Ciphers

- Caesar Cipher
- Monoalphabetic Ciphers
- Affine Caesar
- Playfair Cipher
- Hill Cipher
- Polyalphabetic Ciphers
    - Vigenère Cipher
    - Vernam Cipher
- One-Time Pad

Caesar Cipher



The Caesar cipher is named for Julius Caesar, who used an alphabet where decrypting would shift three letters to the left.

Source: https://en.wikipedia.org/wiki/Caesar_cipher

Caesar Cipher

- *Symmetric, Substitution, Mono-Alphabetic*
- Involves replacing each letter of the alphabet with the letter standing three places further down the alphabet.

Plaintext

The Key is 3

meet me after the party

PHHW PH DIWHU WKH SDUWB

Ciphertext

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

Caesar Encryption

- Can define transformation by listing all possibilities:

```
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 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
0	1	2	3	4	5	6	7	8	9	10	11	12

n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

Caesar Encryption Algorithm

1. Read the plaintext file into an array of bytes pt
2. Clean pt keeping only letters
3. Shift: $ct[i] \;=\; [\,pt[i] \;+\; key\,]\,\%\,26$
4. Write the ciphertext array ct to a file.

`The key of this code shift is: three`
`THE KEY OF THIS CODE SHIFT IS THREE`
`THEKEYOFTHISCODESHIFTISTHREE`
`WKHNHBRIWKLVFRGHVKLIWLVWKUHH`

Caesar Encryption Algorithm

- Algorithm can be expressed as:

$$C = E(3, p) = (p + 3)\ mod\ (26)$$

- A shift may be of any amount, so that the general Caesar algorithm is:

$$C = E(k, p) = (p + k)\ mod\ 26$$

- Where k takes on a value in the range **1** to **25**

Caesar Decryption Algorithm

1. Read the ciphertext file into an array of bytes ct
2. Un-Shift: $pt[i] = [\,(ct[i] - key\,]\ \boldsymbol{mod}\ 26$
3. Write the ciphertext array pt to a file.

- The decryption algorithm is simply:

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

Caesar Exhaustive Attack

- The encryption and decryption algorithms are known.
- Try every possible key in the key space.
- How big is the key space?
  - There are only 25 keys to try.
- But how do you recognize success?
  - The language of the plaintext is known and easily recognizable.

Caesar Exhaustive Attack

KEY	PHHW PH DIWHU WKH WRJD SDUWB
1	oggv og chvgt vjg vqic rctva
2	nffu nf bqufs 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

- Can you enlarge the key space?
  - Yes, can make it 26! ($\approx 10^{26} \approx 2^{88}$) $\Rightarrow$ monoalphabetic ciphers.

Monoalphabetic Cipher

Can you enlarge the key space?

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

- If $S = \{a, b, c\}$, how many permutations are there? What are they?
  - 6
  - abc, acb, bac, bca, cab, cba

- In general, there are $\boldsymbol{n}!$ permutations of a set of $\boldsymbol{n}$ elements.
  - 1st element can be chosen in one of $\boldsymbol{n}$ ways, the 2nd in $\boldsymbol{n-1}$ ways, the 3rd in $\boldsymbol{n-2}$ ways, etc...

Monoalphabetic Cipher

```
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 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
```

- If the “cipher” line can be any permutation of the 26 alphabetic characters, then there are 26! or greater than 4 x 10^{26} **possible keys** (403,291,461,126,605,635,584,000,000)

- *Monoalphabetic substitution* cipher: A **single** cipher alphabet (mapping from plain alphabet to cipher alphabet) is used per message.

Key Terminology /1 - Keyspace

- **Key space** $\boldsymbol{K}$ (or **keyspace**), the **set** of possible keys. Examples:
  - What is the keyspace for **Caesar?**
    - **Caesar** $\boldsymbol{K}$ is the **set** of all permutations of the alphabet, one substitution for each letter (based on the **shift** value).
  - What is the keyspace for **Monoalphabetic?**
    - **Monoalphabetic** $\boldsymbol{K}$ is the **set** of all permutations of the alphabet, with **arbitrary** substitution for each letter.

Key Terminology /2 – Keyspace Size

- **Key space size** $||K||$, the **number** of possible keys or elements (an integer) in the K set. Examples:
  - What is the keyspace size for **Caesar?**
    - **Caesar** $||K|| = 25$
  - What is the keyspace for **Monoalphabetic?**
    - **Monoalphabetic** $||K|| = 26! = 403{,}291{,}461{,}126{,}605{,}635{,}584{,}000{,}000$

Key Terminology /3 – Key Length

- In modern ciphers, we work in bits and the **key length** is determined by the number of bits of the key (e.g., AES with a 128-bit key).

- Each bit of the key can take the values **0** or **1**, independently.

- The number of possible keys for $\boldsymbol{n}$-bit key is $\mathbf{2^n}$.

Key Terminology /4 – Key Length

- **Key length** (or key size) n in bit, the base 2-logarithm of the keyspace size:

$$K \text{ has } \|K\| = 2^n \text{ keys and}$$

$$n = \log_2(\|K\|)$$

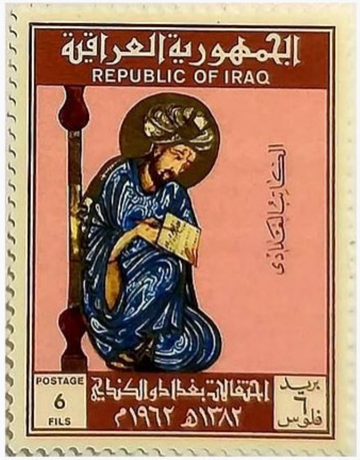
Where n is the key length in bit, and K is the keyspace.

- What is the key length for **Caesar?**
  - **Caesar**, $n = \log_2(25) = \frac{\log(25)}{\log(2)} = 4.6\ bit$
- What is the key length for **Monoalphabetic?**
  - **Monoalphabetic**, $n = \log_2(26!) = \frac{\log(26!)}{\log(2)} = 88\ bit$

Monoalphabetic Cryptanalytic Attack

- Plaintext has certain patterns (regularities)
  - A **Crib** such as: Date, From, GET, Dear ...
  - Language **Statistics** such as N-Gram Frequencies.
- Do they die hard (survive the encryption)?
  - Compute the letter frequencies in ciphertext;
  - The largest is probably the shifted ‘E’ (or ‘T’);
  - Subtract to find the key.

al-Kindi



al-Kindi on Iraqi stamp from 1962

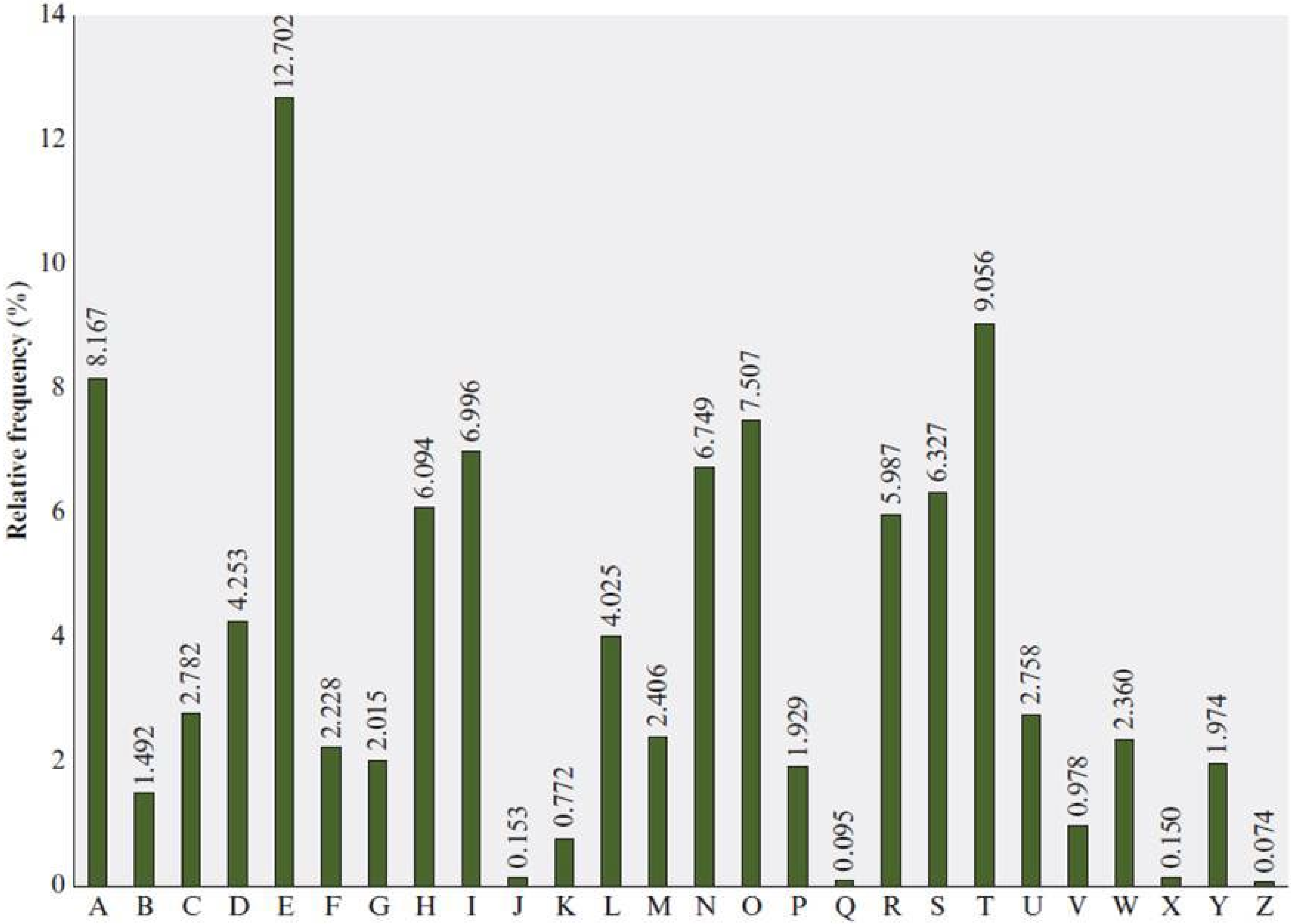
Born c. 801 Kufa, Abbasid Caliphate (now in Iraq)

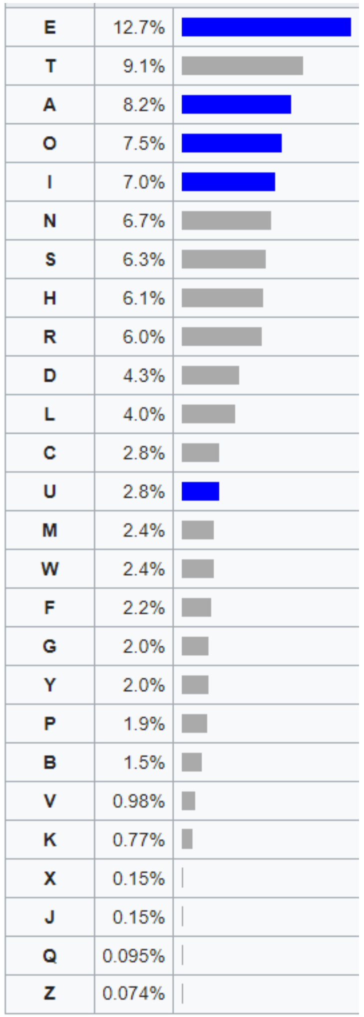
Died c. 873 (aged approximately 72) Baghdad, Abbasid Caliphate (now in Iraq)

The first page of al-Kindi's manuscript "On Deciphering Cryptographic Messages", containing the oldest known description of cryptanalysis by frequency analysis.

Source: https://en.wikipedia.org/wiki/Al-Kindi

Relative Frequency of Letters in English Text





Left Figure Source: Cryptography and Network Security, 8th Edition, by William Stallings
Right figure Source: https://en.wikipedia.org/wiki/Letter_frequency

Monoalphabetic Cryptanalytic Attack Example

Cipher Text

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

Letters Frequency

P 13.33	H 5.83	F 3.33	B 1.67	C 0.00
Z 11.67	D 5.00	W 3.33	G 1.67	K 0.00
S 8.33	E 5.00	Q 2.50	Y 1.67	L 0.00
U 8.33	V 4.17	T 2.50	I 0.83	N 0.00
O 7.50	X 4.17	A 1.67	J 0.83	R 0.00
M 6.67				

Monoalphabetic Cryptanalytic Attack Example

P 13.33	H 5.83	F 3.33	B 1.67	C 0.00
Z 11.67	D 5.00	W 3.33	G 1.67	K 0.00
S 8.33	E 5.00	Q 2.50	Y 1.67	L 0.00
U 8.33	V 4.17	T 2.50	I 0.83	N 0.00
O 7.50	X 4.17	A 1.67	J 0.83	R 0.00
M 6.67				

E	12.7%
T	9.1%
A	8.2%
O	7.5%
I	7.0%
N	6.7%
S	6.3%
H	6.1%
R	6.0%
D	4.3%
L	4.0%
C	2.8%
U	2.8%
M	2.4%
W	2.4%
F	2.2%
G	2.0%
Y	2.0%
P	1.9%
B	1.5%
V	0.98%
K	0.77%
X	0.15%
J	0.15%
Q	0.095%
Z	0.074%

Monoalphabetic Cryptanalytic Attack Example

P 13.33	H 5.83	F 3.33	B 1.67	C 0.00
Z 11.67	D 5.00	W 3.33	G 1.67	K 0.00
S 8.33	E 5.00	Q 2.50	Y 1.67	L 0.00
U 8.33	V 4.17	T 2.50	I 0.83	N 0.00
O 7.50	X 4.17	A 1.67	J 0.83	R 0.00
M 6.67				

E	12.7%
T	9.1%
A	8.2%
O	7.5%
I	7.0%
N	6.7%
S	6.3%
H	6.1%
R	6.0%
D	4.3%
L	4.0%
C	2.8%
U	2.8%
M	2.4%
W	2.4%
F	2.2%
G	2.0%
Y	2.0%
P	1.9%
B	1.5%
V	0.98%
K	0.77%
X	0.15%
J	0.15%
Q	0.095%
Z	0.074%

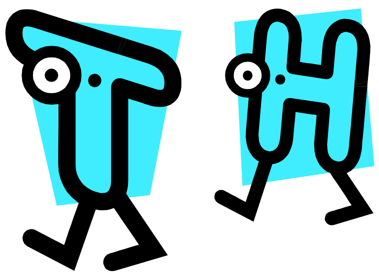
Monoalphabetic Cryptanalytic Attack Example

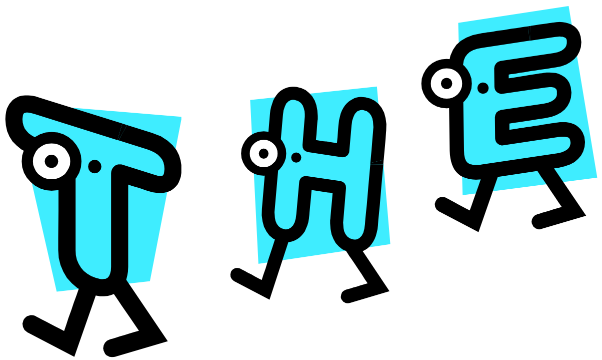
P 13.33	H 5.83	F 3.33	B 1.67	C 0.00
Z 11.67	D 5.00	W 3.33	G 1.67	K 0.00
S 8.33	E 5.00	Q 2.50	Y 1.67	L 0.00
U 8.33	V 4.17	T 2.50	I 0.83	N 0.00
O 7.50	X 4.17	A 1.67	J 0.83	R 0.00
M 6.67				

E	12.7%
T	9.1%
A	8.2%
O	7.5%
I	7.0%
N	6.7%
S	6.3%
H	6.1%
R	6.0%
D	4.3%
L	4.0%
C	2.8%
U	2.8%
M	2.4%
W	2.4%
F	2.2%
G	2.0%
Y	2.0%
P	1.9%
B	1.5%
V	0.98%
K	0.77%
X	0.15%
J	0.15%
Q	0.095%
Z	0.074%

Frequency of Other Letters in English Text

- Monogram:
  - **E** (13%), **T** (9%), **A** (8%); O, N, R, I, S H (6%), L (4%); F, C, M, U, G, Y, P, W (3%); B, V, K (1%)
- Digram (or bigram)
  - Two-letter combination
  - Most common is **TH**, **HE**, **IN**, ER, AN, RE, ...
- Same-letter Digram
  - **LL**, **EE**, **SS**, OO, TT, FF, ...
- Trigram
  - Three-letter combination
  - Most frequent is **THE**, **AND**, **ING**, ENT, ION, HER, ...





Monoalphabetic Cryptanalytic Attack Example

- In our ciphertext, the most common digram is ZW, which appears three times.
- So we make the correspondence of Z with t and W with h.

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

t a e e te a that e e a a

VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

e t ta t ha e ee a e th t a

EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

e e e tat e the t

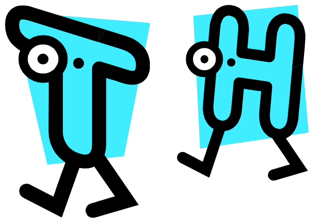
Monoalphabetic Cryptanalytic Attack Example

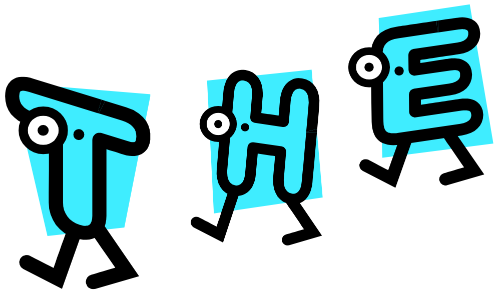
```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

```
it was disclosed yesterday that several informal but

direct contacts have been made with political

representatives of the viet cong in moscow
```

Obliterating Patterns

- Monoalphabetic ciphers are **easy to break** because they reflect the frequency data of the original alphabet.
- To defeat the cryptanalyst, we must prevent PT's patterns from appearing in CT; i.e. make CT as random as possible—maximize its entropy.
- How about these attempts:
  - Compose two ciphers (Affine)
  - Encrypt multiple letters of plaintext
    - Encrypt in blocks (Playfair, Hill)
  - Use multiple cipher alphabets
    - Different mappings for same PT letter (Vigenère, Vernam)





Affine Cipher

Affine Cipher

- A symmetric product cipher
  $c \equiv \alpha p + \beta \pmod{26}$ where $\alpha \epsilon [1, 25]$ and $\beta \epsilon [0, 25]$
- Encryption Example
  $Key = (\alpha, \beta) = (3,5)$, if **P**="CS", what is **C**?
  - **P**="CS" leads to **C**="LH"
- Decryption function
  $$p \equiv (c - \beta) / \alpha \pmod{26}$$

A	0
B	1
C	2
D	3
E	4
F	5
G	6
H	7
I	8
J	9
K	10
L	11
M	12
N	13
O	14
P	15
Q	16
R	17
S	18
T	19
U	20
V	21
W	22
X	23
Y	24
Z	25

Affine Cipher

- Decryption Example
  For encr$yption\ key$ $(3{,}5)$, if **C**="EM", what is **P**=?

$$c \equiv \alpha\, p + \beta \ \ (mod\ 26)$$
$$c \equiv 3\, p + 5 \ \ (mod\ 26)$$
$$c - 5 \equiv 3\, p \ \ (mod\ 26)$$

- We don't want fractions, we want to replace $3\ (mod\ 26)$ by 1

$$9(c-5) \equiv 9.\,3\ p \equiv p\ (mod\ 26)$$

- **C**="EM", leads to **P**="RL"

A	0
B	1
C	2
D	3
E	4
F	5
G	6
H	7
I	8
J	9
K	10
L	11
M	12
N	13
O	14
P	15
Q	16
R	17
S	18
T	19
U	20
V	21
W	22
X	23
Y	24
Z	25

Affine Cryptanalytic Attacks

- Are there any limitations on the value of $\boldsymbol{\beta}$ in $\boldsymbol{c \equiv \alpha\, p + \beta \pmod{26}}$?
  - No

Affine Cryptanalytic Attacks

- Determine which values of α are not allowed.
  - 2, 4, 6, 8, 10, 12, 13, 14, 16, 18, 20, 22, 24. Why?
    - α and 26 should be relatively prime (i.e., gcd(α,26)=1), Why?
    - To enable us to find a value to multiply $(c - \beta) \equiv p\alpha \ (mod\ 26)$ equation with that will result in an $\alpha\ (mod\ 26) = 1$, so we can find p
    - We call that value the Modular Multiplicative Inverse.

  - Any value of α larger than 25 is equivalent to α mod 26.

Affine Cryptanalytic Attacks

- Known Ciphertext … frequency based.

- Example: A ciphertext has been generated with an affine cipher. The most frequent letter of the ciphertext is "**B**," and the second most frequent letter of the ciphertext is "**U**." Break this code.

  1. Assume that the most frequent plaintext letter is **e** and the second most frequent letter is **t**.
  2. **e = 4**; **B = 1**; **t =19**; **U = 20**.
  3. **1** = (**4***a* + *b*) mod 26
  4. **20** = (**19***a* + *b*) mod 26
  5. 19 = 15*a* mod 26. By trial and error, we solve: ***a* = 3**.
  6. Then 1 = (12 + *b*) mod 26. By observation, ***b* = 15**.

A	0
B	1
C	2
D	3
E	4
F	5
G	6
H	7
I	8
J	9
K	10
L	11
M	12
N	13
O	14
P	15
Q	16
R	17
S	18
T	19
U	20
V	21
W	22
X	23
Y	24
Z	25

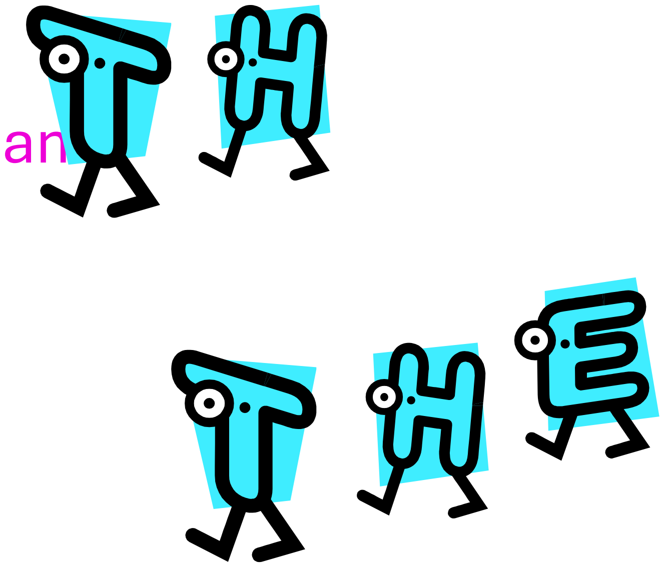
Affine Cryptanalytic Attacks

- Known Ciphertext … frequency based.
- Known Plaintext Attack … how many pairs?
  - 12 x 26 = 312
  - Why 12? And why 26?
    - Allowable α and β, what are these values?
- What if we pick α that doesn't have an inverse?

Playfair Cipher

Recall - Obliterating Patterns

- To defeat the cryptanalyst, we must prevent PT's patterns from appearing in CT; i.e. make CT as random as possible—maximize its entropy.
- How about these attempts:
  - Compose two ciphers (Affine)
  - Encrypt multiple letters of plaintext
    - Encrypt in blocks (Playfair, Hill)
  - Use multiple cipher alphabets
    - Different mappings for same PT letter (Vigenère, Vernam



Playfair Cipher

- 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.

The Playfair system was invented by Charles Wheatstone, who first described it in 1854.

Lord Playfair, who heavily promoted its use.

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

Playfair Key Matrix

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

- What is the keyword in the matrix below?

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 Rules /1

- Plaintext is encrypted two letters at a time, according to the following rules:
  - Each plaintext letter in a pair is replaced by the letter that lies in its own **row** and the **column** occupied by the other plaintext letter. (**hs** ➔ **BP**, **ea** ➔ **IM** (or **JM**)).

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 Rules /2

- **Repeating** plaintext letters that are in the same pair are separated with a filler letter, such as x. (ba**ll**oon → ba **lx lo** on).

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 Rules /3

- Two plaintext letters that fall in the **same row** of the matrix are each replaced by the letter **to the right**, with the 1st element of the row circularly following the last. (**ar** ➔ **RM**).

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

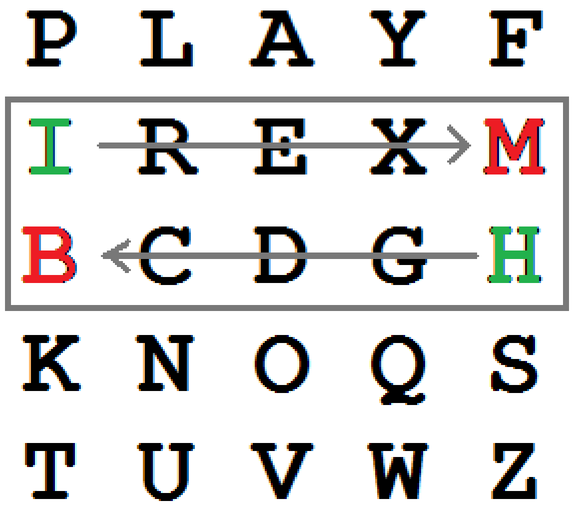
Playfair Rules /4

- Two plaintext letters that fall in the **same column** are each replaced by the letter **beneath**, with the top element of the column circularly following the last. (**mu** → **CM**).

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 Cipher

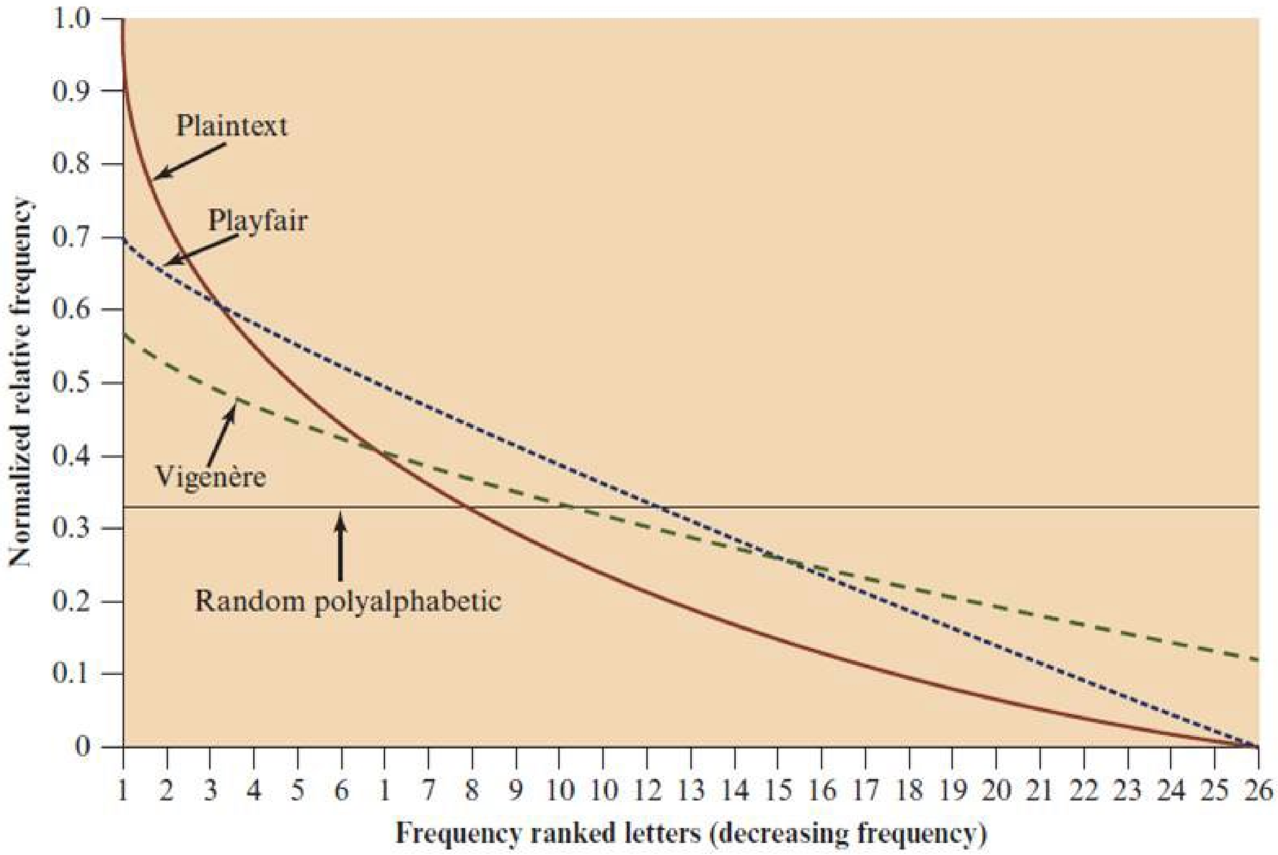
- The Playfair cipher is a **great advance over simple monoalphabetic** ciphers.
  - There are only 26 letters, but there are 26 * 26 = 676 digrams, so identification of individual digrams is more difficult.
- The Playfair cipher is **relatively easy to break**, because it still leaves much of the structure of the plaintext language intact.



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

Relative Frequency of Occurrence of Letters

- Playfair cipher has a flatter distribution than does plaintext, but nevertheless, it reveals plenty of structure for a cryptanalyst to work with.

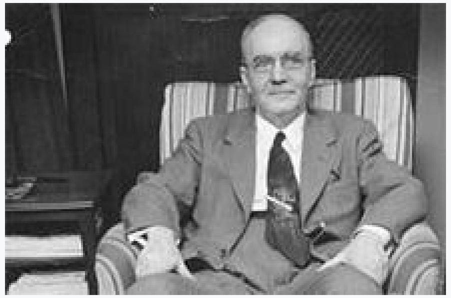


Hill Cipher

Hill Cipher

- Developed by the mathematician Lester Hill in 1929.
- **Strength**: 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.

Dr.
Lester S. Hill



Lester S. Hill on May 16, 1956

Born	Lester Sanders Hill[1] January 18, 1891 New York City
Died	January 9, 1961 (aged 69)[2][3] Bronxville, New York[2]
Nationality	American
Occupation(s)	mathematician and cryptographer
Known for	the Hill cipher (1929)
Notable work	Cryptography in an Algebraic Alphabet (1929)[4]

Multiplicative Inverse

$$\boldsymbol{A * A^{-1} = 1} \text{ or } \boldsymbol{A * \frac{1}{A} = 1}$$

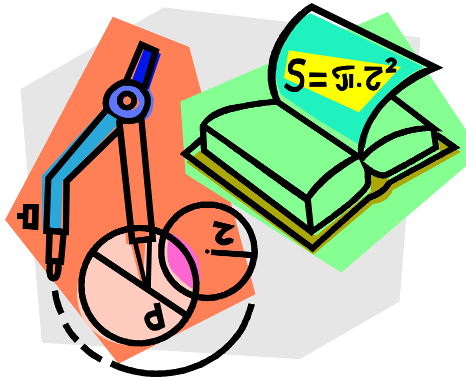
- Examples:
  - $7 * 7^{-1} = 1$
  - $2 * 2^{-1} = 1$
  - $12 * 12^{-1} = 1$
  - $10 * 10^{-1} = 1$

Modular Multiplicative Inverse

- Multiplicative Inverse under mode m

$$\boldsymbol{A * A^{-1} \equiv 1\ (mod\ m)}$$

- Examples where $m = 5$:
  - $7 * ? \equiv 1(mod\ 5)$
  - $2 * ? \equiv 1(mod\ 5)$
  - $12 * ? \equiv 1(mod\ 5)$
  - $10 * ? \equiv 1(mod\ 5)$
  - $0 * ? \equiv 1(mod\ 5)$

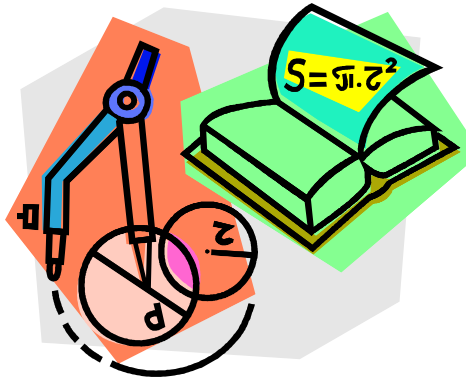


Modular Multiplicative Inverse

- Multiplicative Inverse under mode m

$$\boldsymbol{A * A^{-1} \equiv 1\ (mod\ m)}$$

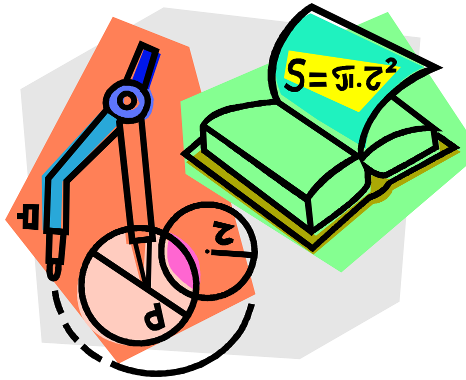
- Examples where m = 5:
  - $7 * 3 \equiv 1\ (mod\ 5)$
  - $2 * 3 \equiv 1 (mod\ 5)$
  - $12 * 3 \equiv 1 (mod\ 5)$
  - $10 * ? \equiv 1 (mod\ 5)$ (there is no modular multiplicative inverse for this integer, why?)
  - $0 * ? \equiv 1 (mod\ 5)$ (Zero has no modular multiplicative inverse)



Modular Multiplicative Inverse

- Is the Multiplicative Inverse of $\mathbf{2\ (mod\ 5)}$ the same as the Multiplicative Inverse of $\mathbf{2\ (mod\ 7)}$?

- $2 * 3 \equiv 1\ (mod\ 5)$
- $2 * 4 \equiv 1\ (mod\ 7)$

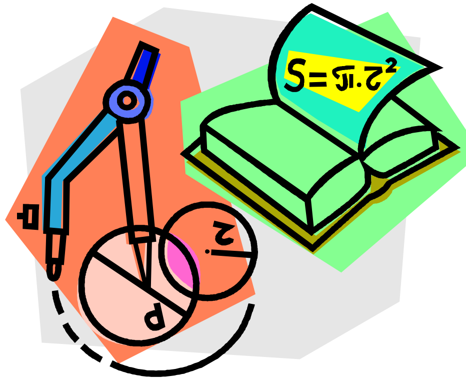


Modular Multiplicative Inverse

- Can you manually calculate the Multiplicative Inverse for $\mathbf{4563210789\ (mod\ 7)}$ **?**

- How about the Multiplicative Inverse of

$$\begin{pmatrix} \mathbf{17} & \mathbf{17} & \mathbf{5} \\ \mathbf{21} & \mathbf{18} & \mathbf{21} \\ \mathbf{2} & \mathbf{2} & \mathbf{19} \end{pmatrix} \mathbf{(mod\ 26)}$$



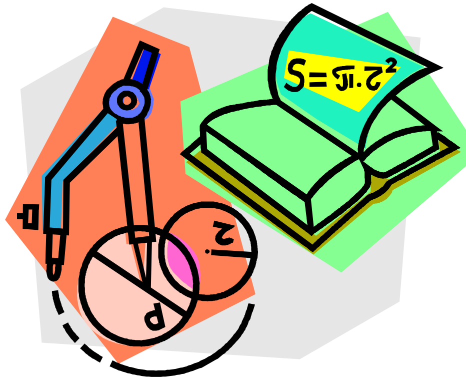
Modular Multiplicative Inverse

- The modular multiplicative inverse of an integer $\boldsymbol{a}$ modulo $\boldsymbol{m}$ is an integer $\boldsymbol{b}$ such that

$$\boldsymbol{ab \equiv 1\ (mod\ m)}$$

- It may be denoted as $\boldsymbol{a^{-1}}$, where the fact that the inversion is $\boldsymbol{m}$-modular is implicit.

- The multiplicative inverse of a modulo $\boldsymbol{m}$ exists if and only if $\boldsymbol{a}$ and $\boldsymbol{m}$ are **coprime** (i.e., if $\mathbf{gcd}(\boldsymbol{a}, \boldsymbol{m}) = 1$)).



Modular Multiplicative Inverse

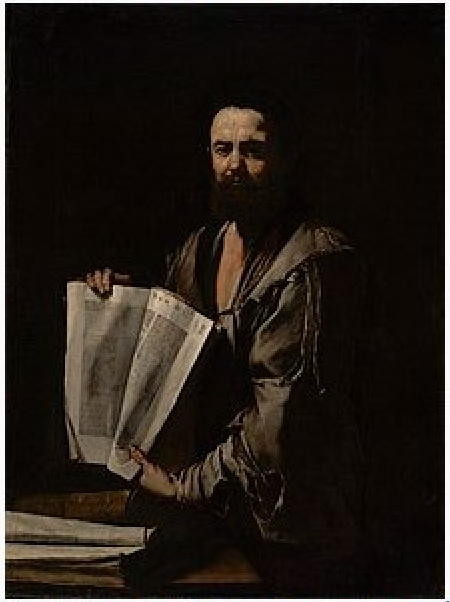
- The modular multiplicative inverse of a modulo m can be found with the **Extended Euclidean algorithm**.

- Euclid [300 BC]

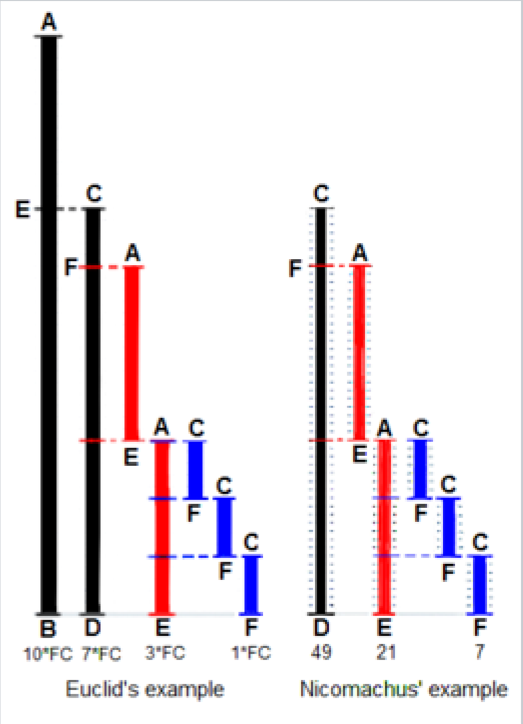
Euclid's algorithm, is an efficient method for computing the greatest common divisor (GCD) of two integers (numbers).

Euclid

Εὐκλείδης



Euclid by Jusepe de Ribera, c. 1630–1635[2]



Euclid's method for finding the greatest common divisor (GCD) of two starting lengths BA and DC, both defined to be multiples of a common "unit" length. The length DC being shorter, it is used to "measure" BA, but only once because the remainder EA is less than DC. EA now measures (twice) the shorter length DC, with remainder FC shorter than EA. Then FC measures (three times) length EA. Because there is no remainder, the process ends with FC being the GCD. On the right Nicomachus's example with numbers 49 and 21 resulting in their GCD of 7 (derived from Heath 1908:300).

Source: https://en.wikipedia.org/wiki/Euclid

The Extended Euclidean Algorithm

- Bézout [1730 AD]
  If a, b are co-prime integers, there exists integers x, y such that: $ax + by = 1$.

- Euclid [300 BC]
  His extended algorithm allows us to find x and y

Working with modulus a, $y = 1/b$, Similarly, if we choose b as modulus then $x = 1/a$

$$by \equiv 1 \ (mod\ a)$$

$$ax \equiv 1 \ (mod\ b)$$

Étienne Bézout

Born	31 March 1730 Nemours, Seine-et-Marne
Died	27 September 1783 (aged 53) Avon, Île-de-France
Nationality	French
Known for	Bézout's theorem Bézout's identity Bézout matrix Bézout domain
Parents	Pierre Bézout (father) Jeanne-Hélène Filz (mother)
	Scientific career
Fields	Mathematics
Institutions	French Academy of Sciences

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

Matrix Modular Inverse Calculator

- https://www.dcode.fr/matrix-inverse

Hill Cipher Algorithm

- Encryption Algorithm

  $\boldsymbol{C = PK \; mod \; 26}$ where $\boldsymbol{K}$ is an $\boldsymbol{nxn}$ matrix

$$(c_1 \; c_2 \; c_3) = (p_1 \; p_2 \; p_3) \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \bmod 26$$

$$\mathbf{C} = \mathbf{PK} \bmod 26$$

- Must be able to invert the key matrix → $GCD(\det([K]), 26) = 1.$

- Key Characteristics:
  - No more P-C positional correspondence
  - The K-C relationship is complex

Hill Cipher Encryption Example

- Use the encryption key below to encrypt a plaintext that is "paymoremoney"

$$K = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

Hill Cipher Encryption Example

- The first three letters of the "paymoremoney" are represented by the vector (15 0 24).
- $\boldsymbol{C = PK \bmod 26} = (15\ 0\ 24)\ \boldsymbol{K} = (\boxed{303}\ \boxed{303}\ \boxed{531})\ \boldsymbol{mod\ 26} = (17\ 17\ 11) = RRL$
  - $\boxed{((15*17)+(0*21)+(24*2))}, \boxed{((15*17)+(0*18)+(24*2))}, \boxed{((15*5)+(0*21)+(24*19))}$

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12

n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

$$K = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

paymoremoney = RRLMWBKASPDH

Hill Cipher Decryption - **Exercise**

- Decrypt cipher text = "RRLMWBKASPDH" using the key provided earlier.

- Decryption requires using the inverse of the matrix K.

$$\boldsymbol{C = E(K,P) = PK\ mod\ 26}$$

$$\boldsymbol{P = D(K,C) = CK^{-1}\ mod\ 26 = PKK^{-1} = P}$$

$$K = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix} \qquad K^{-1} = \begin{pmatrix} 4 & 9 & 15 \\ 15 & 17 & 6 \\ 24 & 0 & 17 \end{pmatrix}$$

Hill Cipher KPA Attack

- Plaintext "**hill**cipher" was encrypted using a 2x2 Hill cipher to produce the ciphertext **HCRZ**SSXNSP. Find the key.

$$\boldsymbol{C = PK\ mod\ 26}$$

$$\boldsymbol{P = CK^{-1}\ mod\ 26}$$

$$\boldsymbol{K = P^{-1}C mod\ 26}$$

$$\begin{pmatrix} \mathbf{7} & \mathbf{2} \\ \mathbf{17} & \mathbf{25} \end{pmatrix} = \begin{pmatrix} \mathbf{7} & \mathbf{8} \\ \mathbf{11} & \mathbf{11} \end{pmatrix} \boldsymbol{K}\ mod\ 26$$

$$\begin{pmatrix} \mathbf{7} & \mathbf{8} \\ \mathbf{11} & \mathbf{11} \end{pmatrix}^{-1} mod\ 26 = \begin{pmatrix} 25 & 22 \\ 1 & 23 \end{pmatrix}$$

$$\boldsymbol{K} = \begin{pmatrix} 25 & 22 \\ 1 & 23 \end{pmatrix} \begin{pmatrix} \mathbf{7} & \mathbf{2} \\ \mathbf{17} & \mathbf{25} \end{pmatrix} = \begin{pmatrix} 549 & 600 \\ 398 & 577 \end{pmatrix} mod\ 26 = \begin{pmatrix} \mathbf{3} & \mathbf{2} \\ \mathbf{8} & \mathbf{5} \end{pmatrix}$$

Hill Cipher - **Exercise**

- Eve mounts a CPA (Chosen-Plaintext Attack) with P="DONT", intercepts C="ELNI". Find the 2x2 Hill's key
  - To verify your answer: $\mathrm{k} = \begin{pmatrix} \mathbf{10} & \mathbf{9} \\ \mathbf{13} & \mathbf{23} \end{pmatrix}$

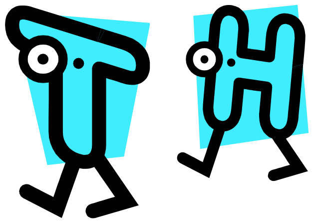
- Repeat with P="DONT", C="ELNK".
  - To verify your answer: $\mathrm{k} = \begin{pmatrix} \mathbf{10} & \mathbf{19} \\ \mathbf{13} & \mathbf{19} \end{pmatrix}$

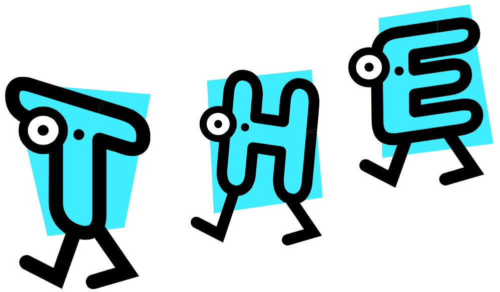
- *One letter change in C changed a **column** in K.*

Polyalphabetic Ciphers

Recall - Obliterating Patterns

- To defeat the cryptanalyst, we must prevent PT’s patterns from appearing in CT; i.e. make CT as random as possible—maximize its entropy.
- How about these attempts:
  - Compose two ciphers (Affine)
  - Encrypt multiple letters of plaintext
    - Encrypt in blocks (Playfair, Hill)
  - Use multiple cipher alphabets
    - Different mappings for same PT letter (Vigenère, Vernam)





Polyalphabetic Ciphers

- Another way to improve on the simple monoalphabetic technique is to use different monoalphabetic substitutions as one proceeds through the plaintext message.

- Polyalphabetic Ciphers:
  - Vigenère Cipher.
  - Vernam Cipher.

Vigenère Cipher

Vigenère Cipher

- This cipher was invented in 1586 by Blaise de Vigenère.
- The best known, and one of the simplest, polyalphabetic ciphers.
- To encrypt a message, a key is needed that is as long as the message.
  - Usually, the key is a repeating keyword.

Blaise de Vigenere

Born	April 5, 1523 Saint-Pourcain-sur-Sioule
Died	February 19, 1596 (aged 72) Paris
Nationality	French
Occupation(s)	diplomat, cryptographer, alchemist

Vigenère Cipher

- Assume a sequence of plaintext letter $P = P_0, P_1, P_2, \ldots, P_{n-1}$ and a key consisting of the sequence of letters $k = k_0, k_1, k_2, \ldots, k_{m-1}$ where typically $\boldsymbol{m} < \boldsymbol{n}$. The ciphertext $C = C_0, C_1, C_2, \ldots, C_{n-1}$ as:

$$C = (P_0 + k_0)\ mod\ 26, (P_1 + k_1)\ mod\ 26, \ldots, (P_{m-1} + k_{m-1})\ mod\ 26, \ldots,$$
$$(P_m + k_0)\ mod\ 26, (P_{m+1} + k_1)\ mod\ 26, \ldots, (P_{2m-1} + k_{m-1})\ mod\ 26, \ldots.$$

- A Polyalphabetic substitution cipher

$$C_i = E(K, P) = (p_i + k_{i\ mod\ m})\ mod\ 26$$
$$P_i = D(K, C) = (C_i - k_{i\ mod\ m})\ mod\ 26$$

Vigenère Cipher Example

Plaintext	w	e	a	r	e	d	i	s	c	o	v	e	r	e	d	s	a	v	e	y	o	u	r	s	e	l	f
Key	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e
Ciphertext	Z	I	C	V	T	W	Q	N	G	R	Z	G	V	T	W	A	V	Z	H	C	Q	Y	G	L	M	G	J

key	3	4	2	4	15	19	8	21	4	3	4	2	4	15
plaintext	22	4	0	17	4	3	8	18	2	14	21	4	17	4
ciphertext	25	8	2	21	19	22	16	13	6	17	25	6	21	19

key	19	8	21	4	3	4	2	4	15	19	8	21	4
plaintext	3	18	0	21	4	24	14	20	17	18	4	11	5
ciphertext	22	0	21	25	7	2	16	24	6	11	12	6	9

Vigenère Cipher

- Strength: There are multiple ciphertext letters for each plaintext letter, one for each unique letter of the keyword.
  - Letter frequency information is obscured.

Plaintext	w	e	a	r	e	d	i	s	c	o	v	e	r	e	d	s	a	v	e	y	o	u	r	s	e	l	f
Key	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e
Ciphertext	Z	I	C	V	T	W	Q	N	G	R	Z	G	V	T	W	A	V	Z	H	C	Q	Y	G	L	M	G	J

Vigenère Cipher Attack

Plaintext	w	e	a	r	e	d	i	s	c	o	v	e	r	e	d	s	a	v	e	y	o	u	r	s	e	l	f
Key	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e	d	e	c	e	p	t	i	v	e
Ciphertext	Z	I	C	V	T	W	Q	N	G	R	Z	G	V	T	W	A	V	Z	H	C	Q	Y	G	L	M	G	J

- Exhaustive Attack: = $26^{|k|} \approx 2^{5|k|}$ ➔ *hopeless!*
- Cryptanalysis
  - Characters that are |K| apart are shifted equally!
  - ➔ Can answer: is the key of a given length?

Vigenère Autokey System

- A keyword is concatenated with the plaintext itself to provide a running key.
- Example:

key: deceptivewearediscoveredsav

plaintext: wearediscoveredsaveyourself

Ciphertext ZICVTWQNGKZEIIGASXSTSLVVWLA

Plaintext

Key

	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 Square can be used for encryption and decryption

Source: https://www.javatpoint.com/vigenere-cipher

Vigenère Autokey System Cryptanalytics

- Vulnerable to cryptanalysis:
  - The key and the plaintext share the same frequency distribution of letters.
  - A statistical technique can be applied.

- Defense: Choose a keyword that is as long as the plaintext and has no statistical relationship to it.

Vernam Cipher

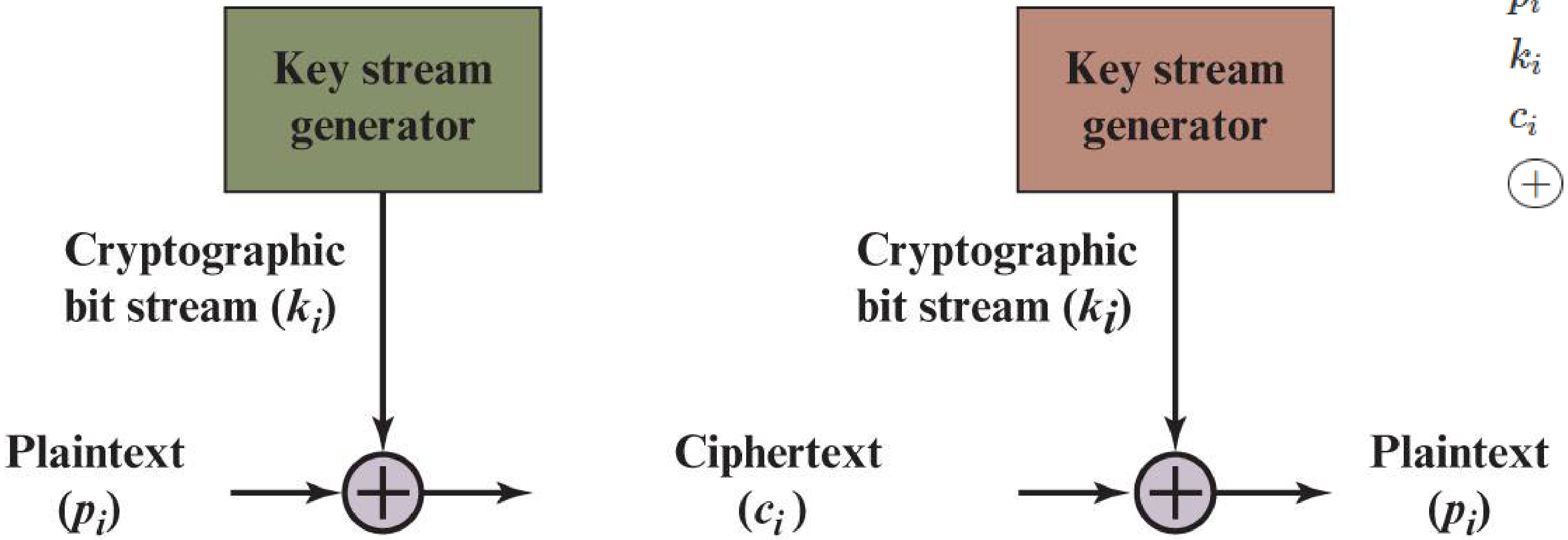
Vernam Cipher

- In Vernam cipher, we choose a keyword that is as long as the plaintext and has no statistical relationship to it.
- The system was introduced by an AT&T engineer named Gilbert Vernam in 1918.
- His system works on binary data (bits) rather than letters.

Gilbert Vernam	
Born	April 3, 1890
Died	February 7, 1960 (aged 69)
Nationality	American
Alma mater	Worcester Polytechnic Institute
Occupation	Cryptographer

Vernam Cipher

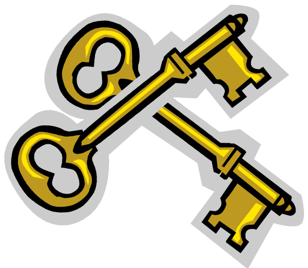
$$c_i = p_i \oplus k_i$$
$$p_i = c_i \oplus k_i$$



p_i = ith binary digit of plaintext
k_i = ith binary digit of key
c_i = ith binary digit of ciphertext
$\oplus$ = exclusive-or (XOR) operation

Vernam Cipher Attack

- Although such a scheme, with a long key, presents formidable cryptanalytic difficulties, it can be broken with:
  - Sufficient ciphertext.
  - The use of known or probable plaintext sequences, or both.



One-Time Pad

One-Time Pad (OTP)

- 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.

Source: https://en.wikipedia.org/wiki/Joseph_Mauborgne

OTP Definition

- Definition
  1. $|K| = |P|$
  2. K is random
  3. $c = E(k,p) = k \oplus p$ *(bitwise ^)*
  4. K never re-used (hence the O in OTP)

OTP Examples

ciphertext:
ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS

key: pxlmvmsydofuyrvzwc tnlebnecvgdupahfzzlmnyih

plaintext: mr mustard with the candlestick in the hall

ciphertext:
ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS

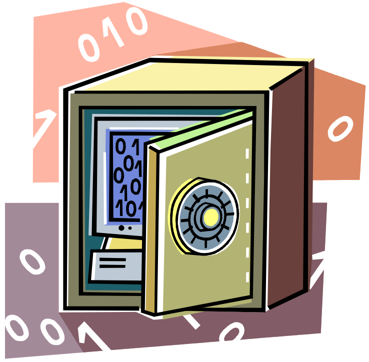
key: pftgpmiydgaxgoufhklllmhsqdqogtewbqfgyovuhwt

plaintext: miss scarlet with the knife in the library

OTP Perfect Secrecy

- 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.

- It boasts perfect secrecy
  - Thwarts *exhaustive* attacks even if Eve had *infinite* classical or quantum computing power!



Generated using: https://worditout.com/word-cloud/create

OTP 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.

Generated using: https://worditout.com/word-cloud/create

OTP Difficulties

- Mammoth key distribution problem:
  - For every message to be sent, a key of equal length is needed by both sender and receiver.

Topics

- Symmetric Cipher Model
- Substitution Techniques
- Transposition Techniques

Transposition Techniques

- **Transposition cipher** performs some sort of permutation on the plaintext letters
  - Rail Fence Cipher
  - Row Transposition Cipher

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

Rail Fence 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
```

Encrypted →

```
MEMATRHTGPRYETEFETEOAAT
```

Rail Fence Cipher Attack

- A pure transposition cipher is trivial to cryptanalyze, because it has the same letter frequencies as the original plaintext.
- Exhaustive: Try all possible numbers of key length on the known ciphertext (start with 2 and increment).
- Known / Chosen Plaintext: Trivial to find the key.

Row Transposition Cipher

- 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:	TTNAAPTMTSUOAODWCOIXKNLYPETZ

Row Transposition Cipher – Double Encrypt

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: TTNAAPTMTSUOAODWCOIXKNLYPETZ

Key: 4 3 1 2 5 6 7
Input: t t n a a p t
m t s u o a o
d w c o i x k
n l y p e t z
Output: NSCYAUOPTTWLTMDNAOIEPAXTTOKZ

Today’s Topics

- Symmetric Cipher Model
- Substitution Techniques
- Transposition Techniques

Today’s lecture slides were prepared from “Cryptography and Network Security”, 8/e, by William Stallings, Chapter 3 – “Classical Encryption Techniques”.