اونيۏرسيتي مليسيا ڤهڠ السلطان عبدالله
UNIVERSITI MALAYSIA PAHANG
AL-SULTAN ABDULLAH

Data & Network Security

Chapter 3 – Cryptography
Part 1

Outline

3.1 Introduction

3.2 Terminology

- 3.2.1 Plain text, cipher text
- 3.2.2 Encryption, decryption
- 3.2.3 Key
- 3.2.3 Symmetric and Asymmetric Cryptosystem
- 3.2.4 Cryptanalysis & Brute Force Attack

3.4 Cipher Types

3.5 Classical Cryptography

3.5 Symmetric Encryption

3.6 Asymmetric Encryption

3.7 Public Key Cryptosystem (RSA, Key Exchange, Diffie-Hellman)

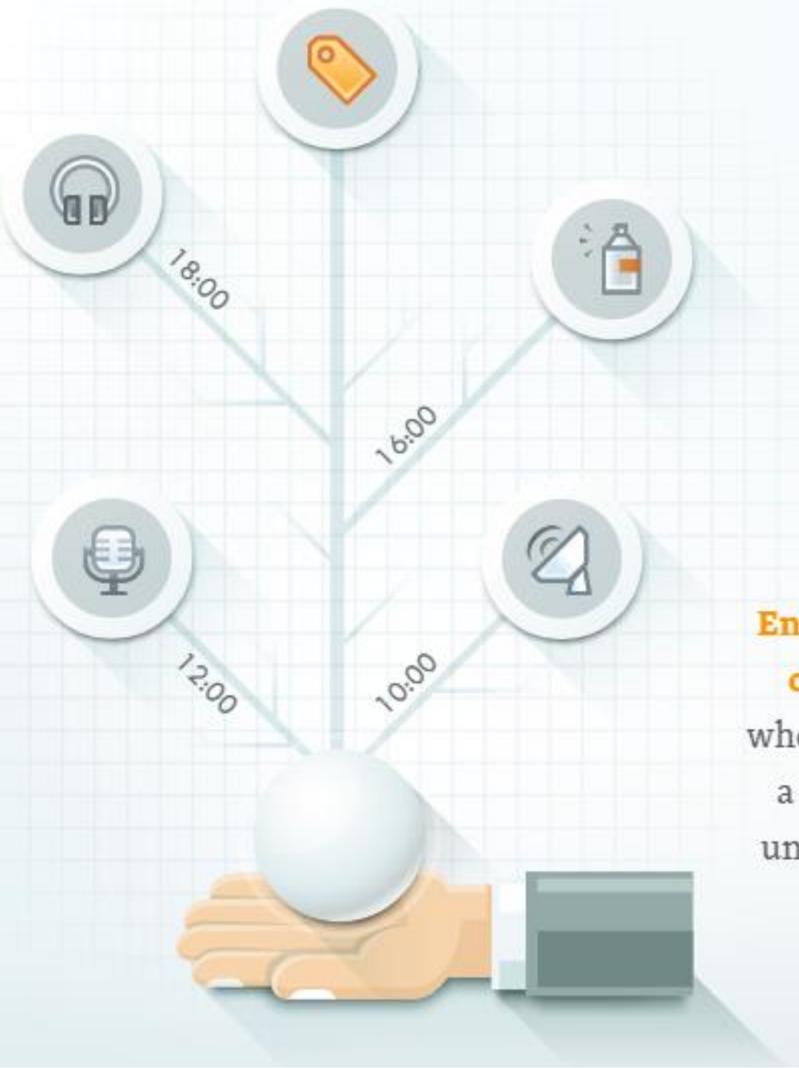
3.8 Others - Message Authentication, Hash Function, Digital Signature

Learning Outcome

At the end of this chapter, the students will be able to

- Define the cryptography definition
- Explain the terminologies used in cryptography.
- Apply some cryptography techniques/ approaches.
- Understand the concept and challenge of Public Key Cryptosystem.
- Understand the concepts of message authentication, hash function, digital signature.

Cryptography is a method of using advanced mathematical principles in storing and transmitting data in a particular form so that only those whom it is intended can read and process it.



Encryption is a key concept in cryptography – It is a process whereby a message is encoded in a format that cannot be read or understood by an eavesdropper.

© Rocheston 2023

Introduction – Cryptography

- Originate from Greek – “kryptos” (secret) and “graphia” (writing).
- Cryptography is defined as secret writing.
- In technical, cryptography is a mapping of readable text to a format that cannot be read (unreadable).
- e.g. ‘ME’ to ‘NB’
- In the early days, cryptography used to be performed by using manual techniques.

Introduction – Cryptography

- 5 Century SM, Sparta people used a method to encrypt messages using a paper made from ‘daun lontar’ attached to a wood.
- Then, information to be hidden will be written on the ‘daun lontar’.
- When the paper is opened from the wood, the words written will be scattered and hard to understand.
- To get back the original message, the paper must be attached back to the same wood.
- In this case – the paper and wood used are key to this system. This encrypted method is called Scytale.

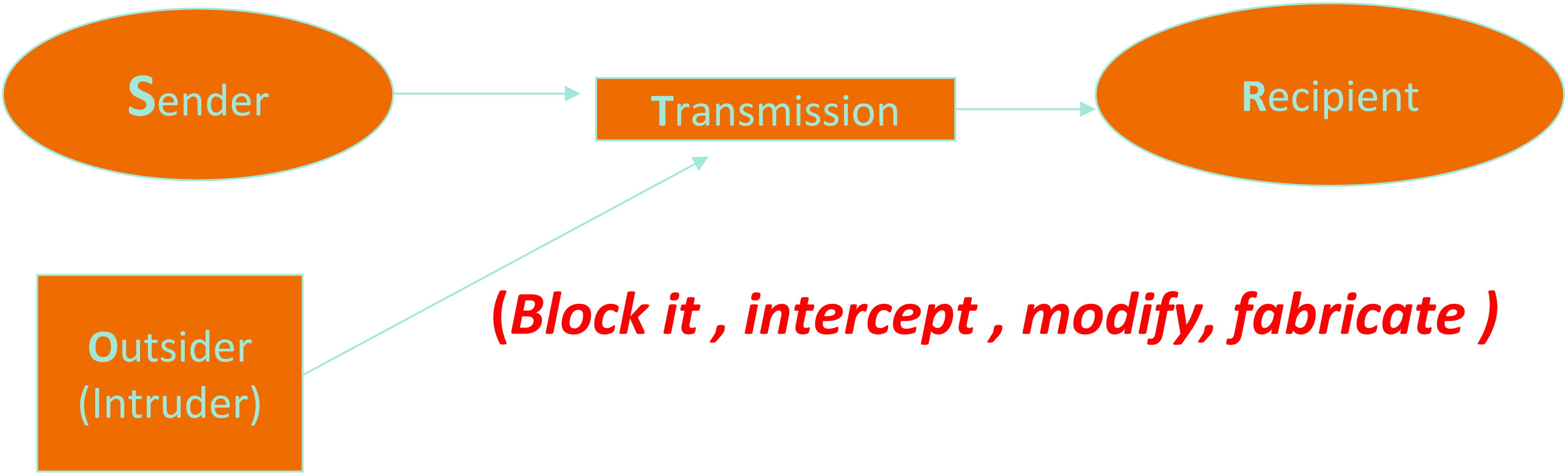
Introduction – Cryptography

- In Julius Caesar (around 2000 years ago), he used a substitution cryptography system created by himself.
- In this method – each word in the text is moved 2 places afterwards in the ABC character table.
- e.g. word A substituted with C, B with D and so on.
- This method is called Caesar cipher.
- However, this method had been broken through analysis towards cipher text. Arabian is the first race that analyzed substitution cipher code.
- Qalqashandi created a technique to solve the code by collecting all the cipher characters and counting the frequency usage of each character.
- Based on this table of frequency, cipher text could be decrypted to get back the original text.

Why Use Cryptography?

- Confidentiality – prevent from message being disclosed to unauthorized users or parties. The message is disclosed to authorized and to the intended parties who have rights to that message only.
- Integrity – ascertain that no modification to the message being received. This is to ensure that the message didn’t modify when sending from sender to receiver.
- Authentication – permit message receiver to verify the original message being sent. This is to make sure that the message can be verified with confidence and prevent it from being disguised.
- Non-repudiation – the sender cannot deny later that he/she has sent the message.

Simple Message Transmission



(Block it , intercept , modify, fabricate)

- Consider the steps involved in sending a message from a sender to a recipient. If the sender entrusts the message to T (transmission), who then delivers it to the recipient, T becomes the transmission medium. If an outsider wants to access the message (to read, change or even destroy it), we call an outsider the intruder.

Terminology

- Human languages take the form of plain text or clear text.
- Message in plain text can be understood by anybody knowing the language.
- Notably, we use plain text during electronic conversations.
- e.g. send an email to someone.
- Clear text or plain text signifies a message that can be understood by the sender, the recipient and by anyone else who gets access to that message.

Cont…

- In normal life, we do not bother about the fact that someone could be overhearing us.
- However, there are situations where we are concerned about the secrecy of our conversations.
- e.g. knowing bank account balance, secret messages from military officers, secret emails, children or primary school students hide their conversation through code language.

Cont…

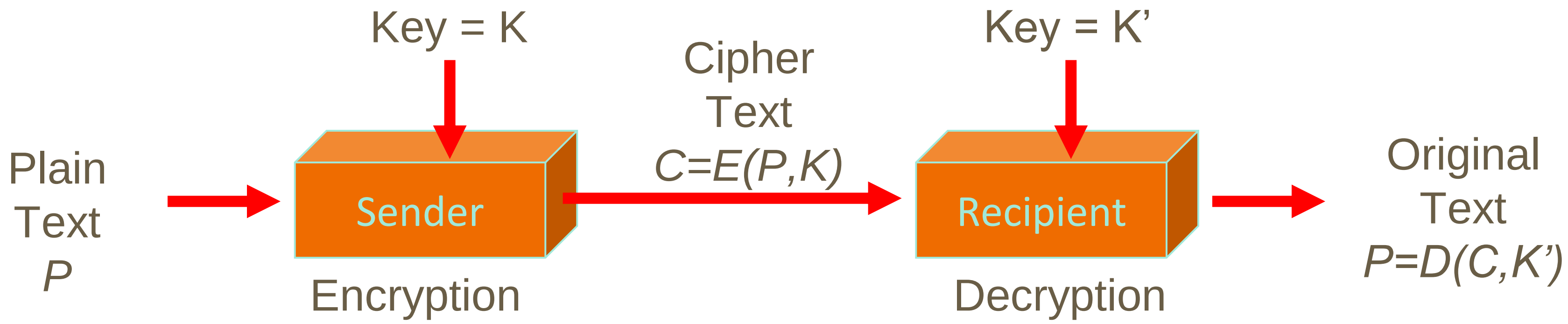
- Given P (plain text) wants to be transferred through a communication channel as a secret message.
- First, the P must be converted to another form. The conversion process is called encryption.
- When this plain text message is codified using the suitable scheme, the resulting message is called cipher text (given C).
- Cipher means a code or a secret message.

Cont…

- Cryptography algorithm is a technique or rule to encrypt that determines how easy or complex the encryption process is.
- Format transformation of the original text, P to a cipher text format, C dependent on an additional parameter, K called key.
- Cipher text, C must undergo an inverse process to get back the plain text, P. This process could be done using a second key, K’.
- This inverse process is called decryption.

Cont…

- The learning about encryption and decryption is called cryptography.
- The process to get the original text from cipher text without key is called cryptanalysis.
- The discipline that combine the 2 divisions (cryptography and cryptanalysis) is called cryptology.



Cont…

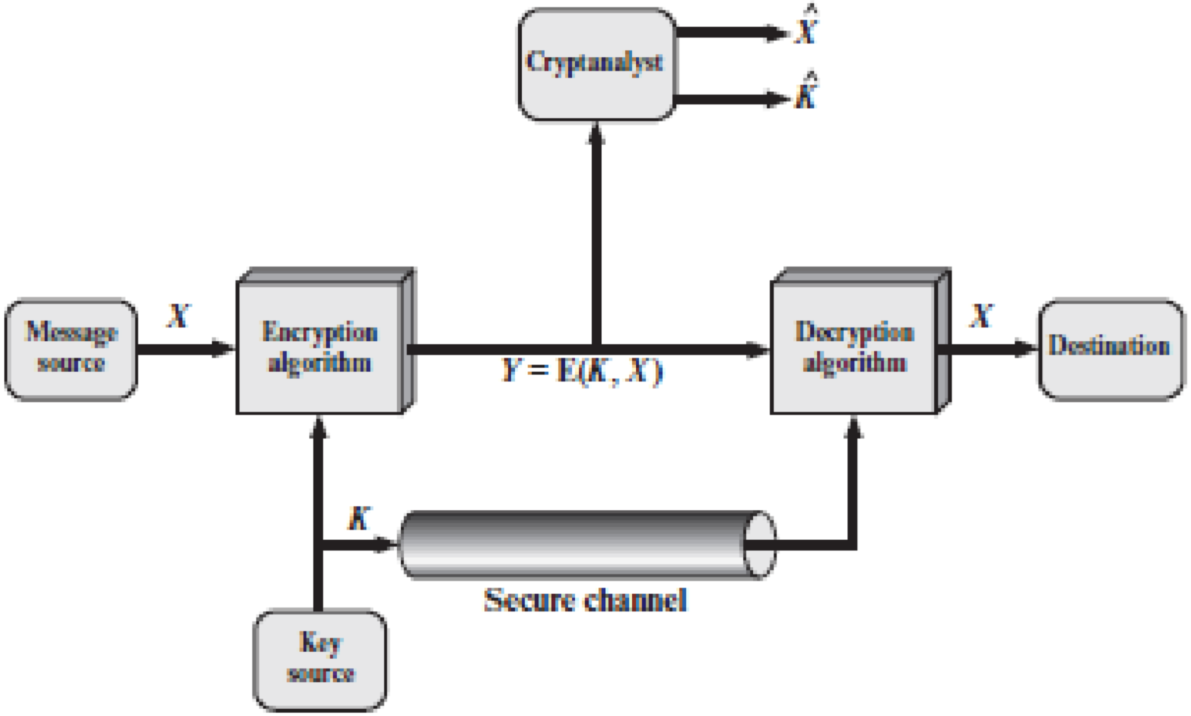
- The study of many schemes used for encryption constitutes the area of study known as **cryptography**. Such a scheme is known as a **cryptographic system**.
- Techniques used for deciphering a message without any knowledge of the enciphering details fall into the area of **cryptanalysis**.
- Cryptanalysis is what the layperson calls "breaking the code."
- The areas of cryptography and cryptanalysis together are called **cryptology**.

Using Key in Cryptography

Encryption Algorithm

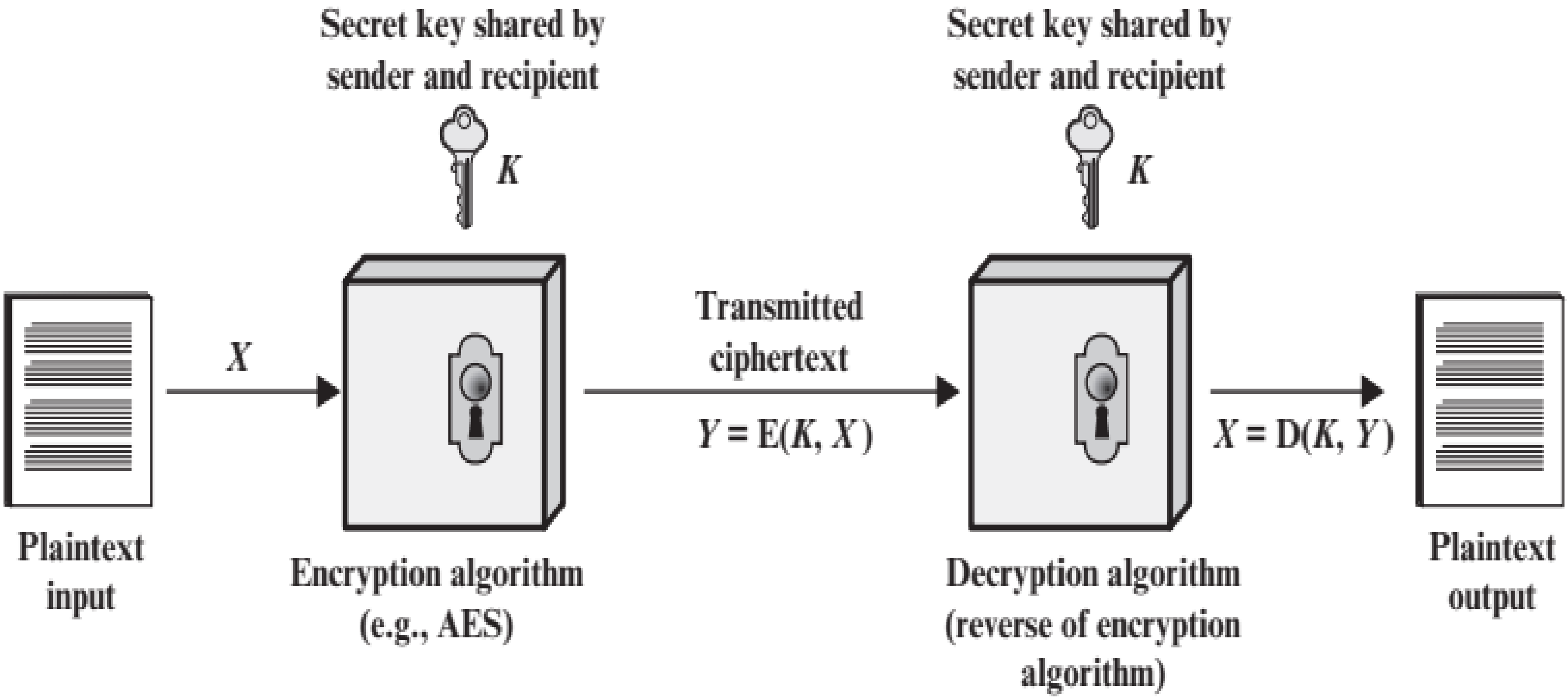
- The cryptographic system involves a set of rules for how to encrypt the plaintext and how to decrypt the cipher text. The encryption and decryption rules, called algorithms, often use a thing called a key, denoted by K.

Symmetric Cryptosystem



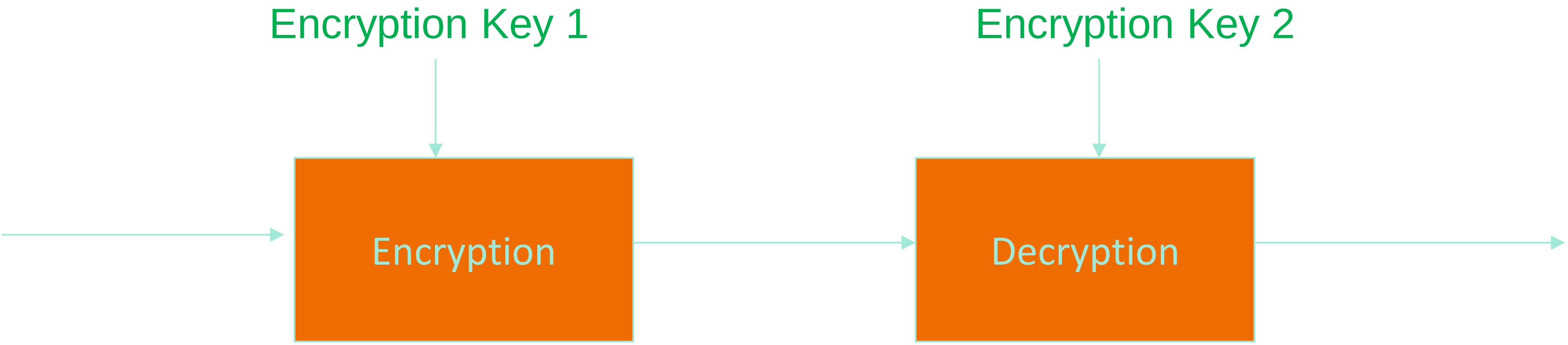
- The keys that were used to encrypt and decrypt are the same and mirror-image process.

Simplified Model of Symmetric Encryption


Secret key shared by sender and recipient
K
Secret key shared by sender and recipient
K
X
Transmitted ciphertext
$Y = E(K, X)$
$X = D(K, Y)$
Plaintext input
Encryption algorithm (e.g., AES)
Decryption algorithm (reverse of encryption algorithm)
Plaintext output

Asymmetric Cryptosystem

$$P = D\,(K_D,\, E(K_E,P))$$



- The process of converting the decrypted message to the original text involves a series of steps and a key that are different from the encrypting process.

Cryptanalysis & Bruteforce Attack

Cryptanalytic Attacks

- rely on:
  - nature of the algorithm
  - some knowledge of the general characteristics of the plaintext
  - some sample plaintext-ciphertext pairs
- exploits the characteristics of the algorithm to attempt to deduce a specific plaintext or the key being used
  - if successful all future and past messages encrypted with that key are compromised

Brute-Force Attack

- try all possible keys on some ciphertext until an intelligible translation into plaintext is obtained
  - on average half of all possible keys must be tried to achieve success

Cipher Types

- The cipher method can be divided into two types:
  - Bit stream cipher
    - Each bit in the plaintext is transformed into a cipher bit one bit at a time.
  - Block cipher
    - The message is divided into blocks and each block of plaintext bits is transformed into encrypted block cipher bits using an algorithm and a key.
    - Example: 8, 16, 32, 64-bit blocks.

Classical Cryptography

- In classic cryptography technique, there are 2 basic components; substitution and transposition.
- Substitution cipher substitutes bit, character or one block of character (e.g. one-character substitutes for another character: C substitutes with F).
- Transposition cipher (or called permutation cipher) arranges back or transposes bits or characters of the original text.

Classical Cryptography

- Example of substitution cipher are Easy substitution, Homophonic substitution , Polyalphabetic substitution and Polygram substitution .
- Example of transposition cipher are Columnar transposition, Rail fence and Vernam cipher.
- Elements of substitution and transposition are also used in modern cryptography algorithms.

Substitution Cipher

- Mono-alphabetic – replace plain text with cipher text using a single alphabet key (eg. Caesar)
- Homophonic – like mono-alphabetic but the substitution is not fixed (eg. A is always replaced with D with key 3 in mono-alphabetic but not in homophonic) (eg. Beale cipher)
- Poly-alphabetic – replace plain text with cipher text using single multiple substitution letters (eg. Vigenere)
- Polygram - a block of alphabets is replaced with another block (eg. Playfair)

Caesar Cipher

- Caesar Cipher – proposed by Julius Caesar.
- Each alphabet in a message is replaced by an alphabet 3 places down the line.
- Very weak scheme of hiding plain text messages – to break it, reverse Caesar Cipher process with the alphabet that is 3 places up the line.
- e.g. A with X, B with Y, C with Z, D with A and so on.
- Good in theory but not so good in practice.
- How to make the cipher more difficult?
- Cipher text alphabets corresponding to the original plain text alphabets may not necessarily be 3 places down the order, instead, can be any number.

Caesar Cipher

- Then have Caesar cipher as: C=Cipher, P=Plain text, K=Key, E=Encyrption, D=Decryption
  - C = E(K, P) = (P + K) mod (26); E = (P+K) mod 26
  - P = D(K, C) = (C – K) mod (26); D = (C-K) mod 26
- Only have 26 possible ciphers
    - A maps to A,B,..Z
- A brute force search - given ciphertext, just try all shifts of letters.
- Break ciphertext “VHFXULWB“.

Caesar Cipher

- The major weakness of Caesar Cipher is its predictability.
- Rather than using a uniform scheme, use random substitution. This means that in each plain text message, each A can be replaced by any other alphabet (B through Z), each B can also be replaced by any other random alphabet (A or C through Z) and so on.
- The crucial difference, there is no relation between the replacement of B and replacement of A. That is, if decides to replace A with D, not necessarily replace each B with E – can replace B with another character.

Caesar Cipher

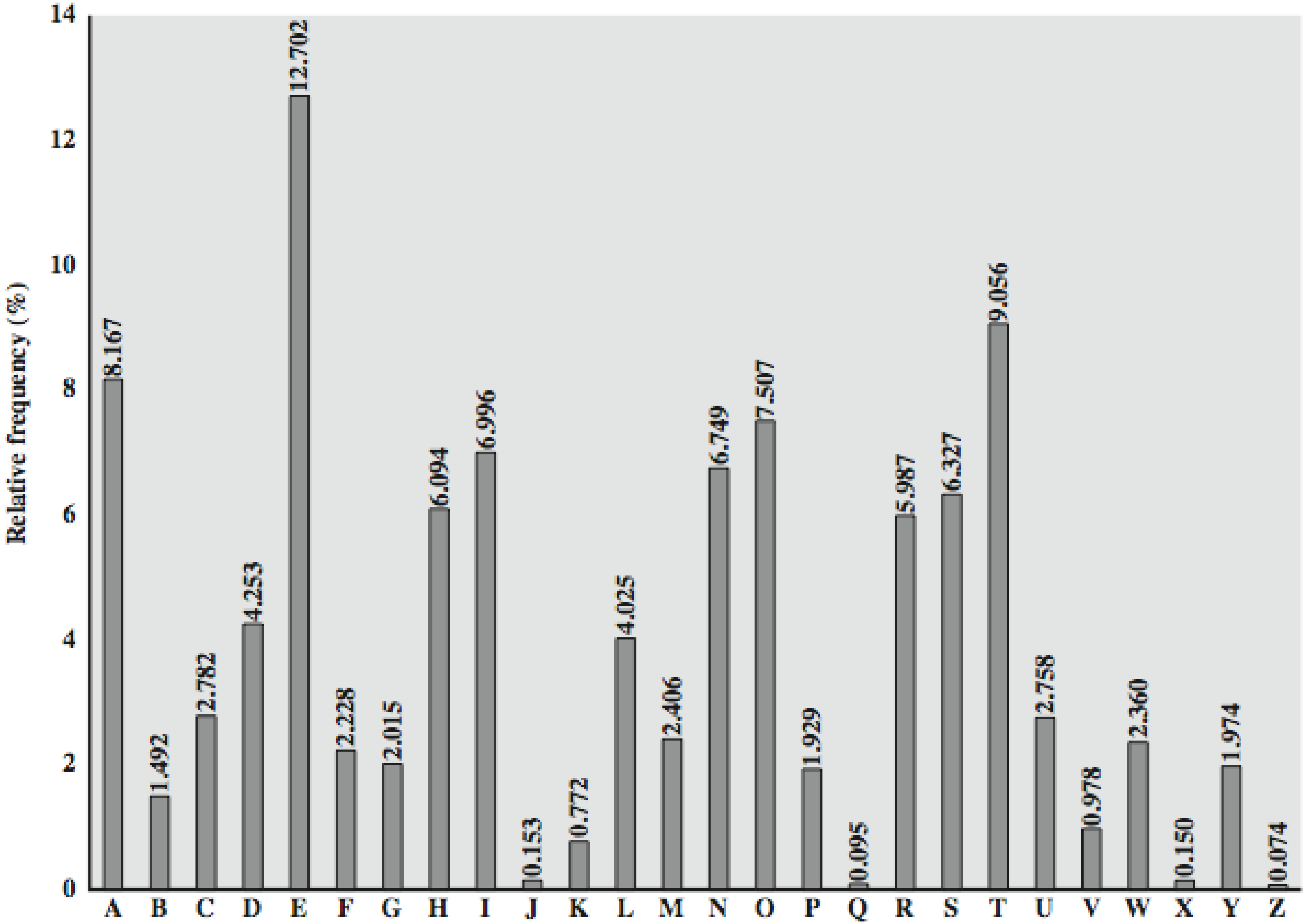
- Now, have a total of 26! = 4 x 1026 keys.
- This is extremely hard to crack. It might take years to try out these many combinations even with the most modern computers.
- There is only one hitch. The cryptanalyst can try different attacks based on her knowledge of the English language.

Caesar Cipher - Language Redundancy and Cryptanalysis

- Human languages are redundant

  eg "th lrd s m shphrd shll nt wnt"
- Letters are not equally commonly used.
- Other letters like Z, J, K, Q, and X are rare.
- Have tables of single, double and triple letter frequencies for various languages.
- In English E is by far the most common letter followed by T, R, N, I, O, A, S.

English Letter Frequencies



Use in Cryptanalysis

- Key concept – mono-alphabetic substitution ciphers do not change relative letter frequencies.
- Discovered by Arabian scientists in the 9th century.
- Calculate letter frequencies for ciphertext.
- Compare counts/plots against known values.
- For mono-alphabetic must identify each letter – tables of common double/triple letters help.

Example Cryptanalysis

- Given ciphertext:

  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
  VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
  EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- Count relative letter frequencies (see text).
- Guess ‘P & Z’ are ‘e’ and ‘t’.
- Guess ‘ZW’ is ‘th’ and hence ZWP is ‘the’.
- Proceeding with trial and error finally get:

  it was disclosed yesterday that several informal but
  direct contacts have been made with political
  representatives of the viet cong in moscow

Homophonic Substitution Cipher

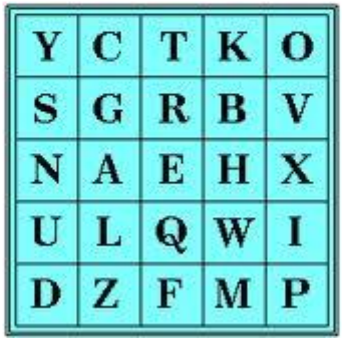
- Very similar to Mono-alphabetic Cipher.
- The difference between the 2 techniques is that the replacement alphabet set in simple substitution technique is fixed (A with D..) whereas, in the case of Homophonic, one plain text alphabet can map to more than one cipher text alphabet.
- e.g. A can be replaced by D, H, P, R; B can be replaced by E, I, Q, S….
- Difficult to analyze compared with mono-alphabetic because the frequency didn’t show the real usage of each alphabet.

Polygram Substitution Cipher

- Rather than replacing one plain text alphabet with one cipher text alphabet at a time, a block of alphabets is replaced with another block.
- It is done by dividing plain text into a group of alphabet. This group can be 2 alphabets or more than that.
- Playfair Cipher and Hill Cipher are examples of cipher that use Polygram Substitution Cipher.

Playfair Cipher

- Not even the large number of keys in a monoalphabetic cipher provides security.
- One approach to improving security was to encrypt multiple letters.
- The Playfair Cipher is an example.
- Invented by Charles Wheatstone in 1854 but named after his friend Baron Playfair.



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

Playfair Cipher

- Playfair cipher algorithm based on a 5 x 5 matrix and one key. This matrix was created using the key. There are 5 rules to obey.
- e.g. given a key = LEDANG and plain text = DATANETWORKSECURITY, what is the cipher text?

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

Playfair Cipher

- Plaintext is encrypted two letters at a time
- Below is a set of rules for Playfair Cipher:
  - divide plain text into a group of 2 alphabets each. DA, TA, NE, TW, OR, KS, EC, UR, IT, Y. If a group lacks one alphabet, fill it with X. Y becomes YX.
  - if a pair is a repeated letter, insert filler like 'X'.
  - if both letters fall in the same row, replace each with a letter to the right (wrapping back to start from the end).
  - if both letters fall in the same column, replace each with the letter below it (wrapping to top from bottom).
  - otherwise, each letter is replaced by the letter in the same row and the column of the other letter of the pair.

Playfair Cipher

- DA, TA, NE, TW, OR, KS, EC, UR, IT, YX.

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

- DA – AN, TA – YF, NE – LD, TW – RY, OR – KT, KS – MR, EC – DB, UR – QS, IT – OQ, YX – ZY
- Cipher Text = ANYFLDRYKTMRDBQSOQZY

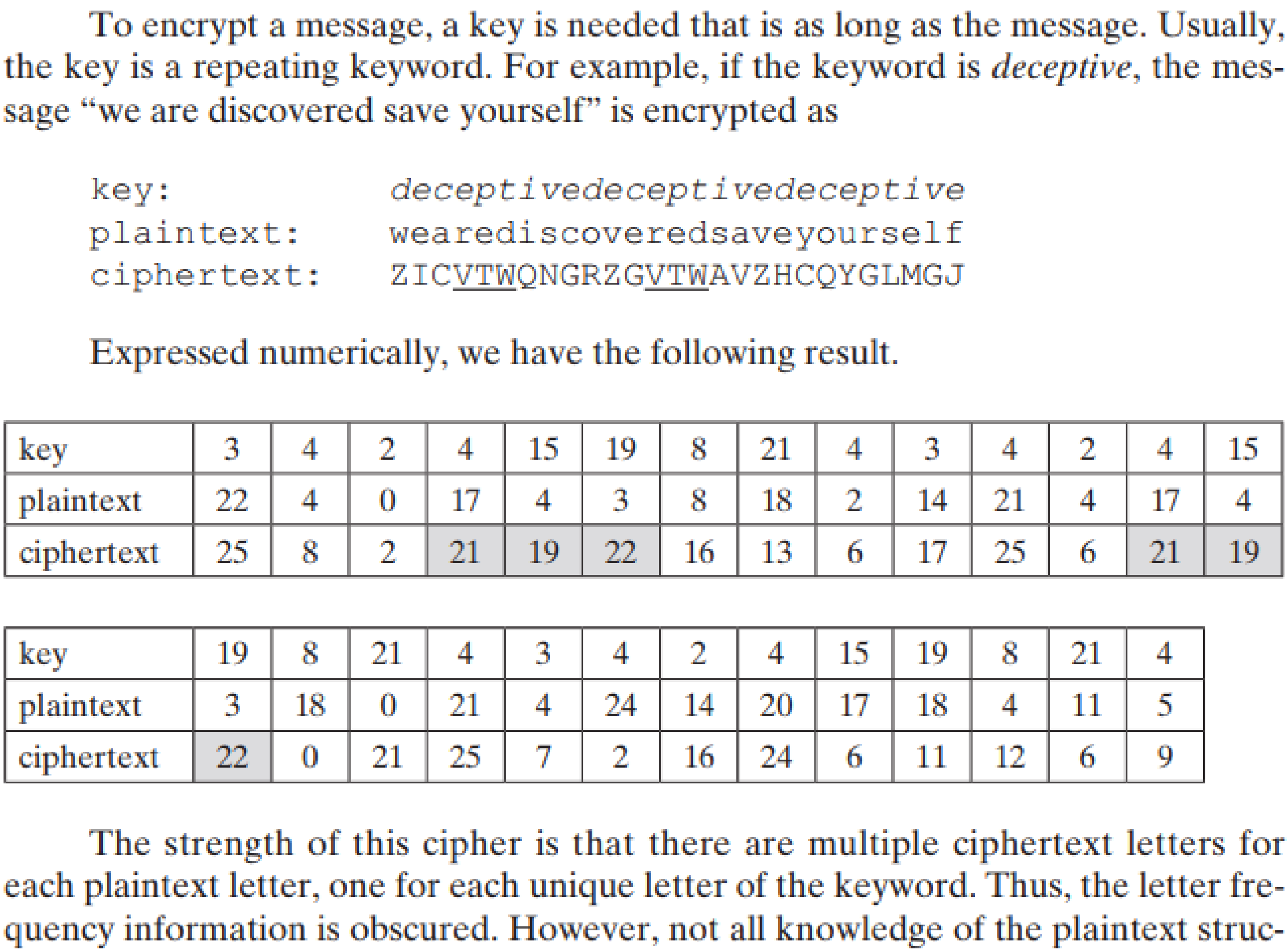
Poly-Alphabetic Substitution Cipher

- Leon Battista invented the Polyalphabetic Cipher in 1568. This cipher has been broken many times, and yet it has been used extensively. The Vigenere Cipher and Beaufort Cipher are examples of it.
- The cipher uses multiple one-character keys. Each of the keys encrypts one plain text character.
- The first key encrypts the first plain text character, the second key encrypts the second plain text character and so on.
- After all the keys are used, they are recycled. Thus, if we have 30 one-letter keys, every 30th character in the plain text would be replaced with the same key.

Vigenere Cipher

- Created by Blaise de Vigenere in the 16th century. In this cipher scheme, one rule set of Mono-alphabetic substitution that is built from 26 Caesar Cipher with a value starting from 0 to 25 used with one value of a key.
- Based on this key, the value for each cipher character is determined.
- e.g. DATANETWORKSECURITY with LEDANG as the key value.

Vigenere 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

```
key:          deceptivedeceptivedeceptive
plaintext:    wearediscoveredsaveyourself
ciphertext:   ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

Expressed numerically, we have the following result.

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

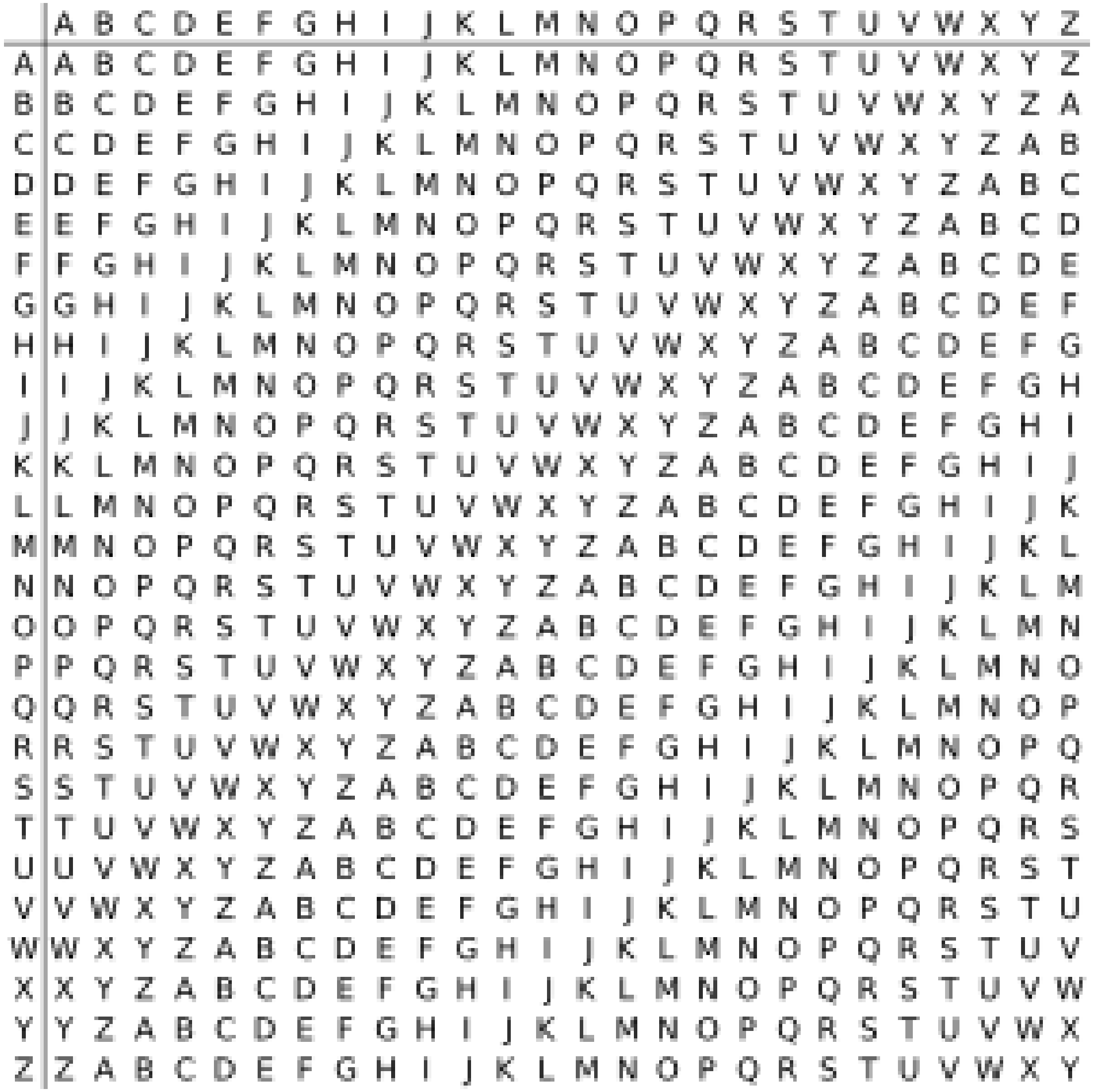
key	19	8	21	4	3	4	2	4	15	19	8	21	4
plaintext	3	18	0	21	4	24	14	20	17	18	4	11	5
ciphertext	22	0	21	25	7	2	16	24	6	11	12	6	9

The strength of this cipher is that there are multiple ciphertext letters for each plaintext letter, one for each unique letter of the keyword. Thus, the letter frequency information is obscured. However, not all knowledge of the plaintext struc-

Vigenere Cipher

Key

Plaintext



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

Vigenere Cipher

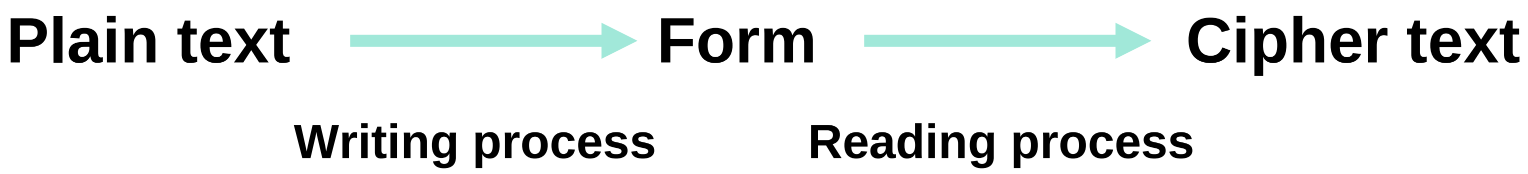
- Eg. 1
- P: DATANETWORKSECURITY
- K: LEDANGLEDANGLEDANGL
- C: OEWAAKEARRXYPGXRVZJ
- Eg.2
- P: HAPPYBIRTHDAY
- K: NETWORKNETWOR
- C: UE
- The first character in plain text, D is moved 11 steps (L key) and so on.
- From this encryption scheme, it is found that the alphabet ‘T’ is encrypted to several alphabets such as ‘W’, ‘E’ and ‘Z’. So, the peak in the frequency alphabet table could be reduced.

Transposition Cipher

- Transposition techniques differ from substitution techniques in the way that they do not simply replace one alphabet with another.
- They also perform some permutations over the plain text alphabets.
- 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

- Usually, the mapping is done with a geometric diagram or matrix.
- The transposition encryption is done in 2 steps:
  - Plain text is arranged in the desired form. This process is referred to as the writing process.
  - Reading process. Is a method to transform plain text that has gone through the writing process to produce cipher text.



Plain text → Form → Cipher text

Writing process

Reading process

Rail Fence Technique

- The Rail Fence is an example of transposition. It uses a simple geometric form as below:
- Encryption: write message letters out diagonally over several rows.

  then read off cipher row by row
- eg. Encrypt message “defend the east wall” with key is 2

d		f		n		t		e		a		t		a		l
	e		e		d		h		e		s		w		l	

Cipher text: dfnteataleedheswl

Rail Fence Technique

- The decryption process for the Rail Fence Cipher involves reconstructing the diagonal grid used to encrypt the message.
- Write the message but leave a dash in place of the spaces yet to be occupied.
- Then, replace all the dashes with the corresponding letters, and read off the plaintext from the table.
- eg. Decrypt message “dfnteataleedheswl” with key 2.

d		f		n		t		e		a		t		a		l
	-		-		-		-		-		-		-		-	

d		f		n		t		e		a		t		a		l
	e		e		d		h		e		s		w		l	

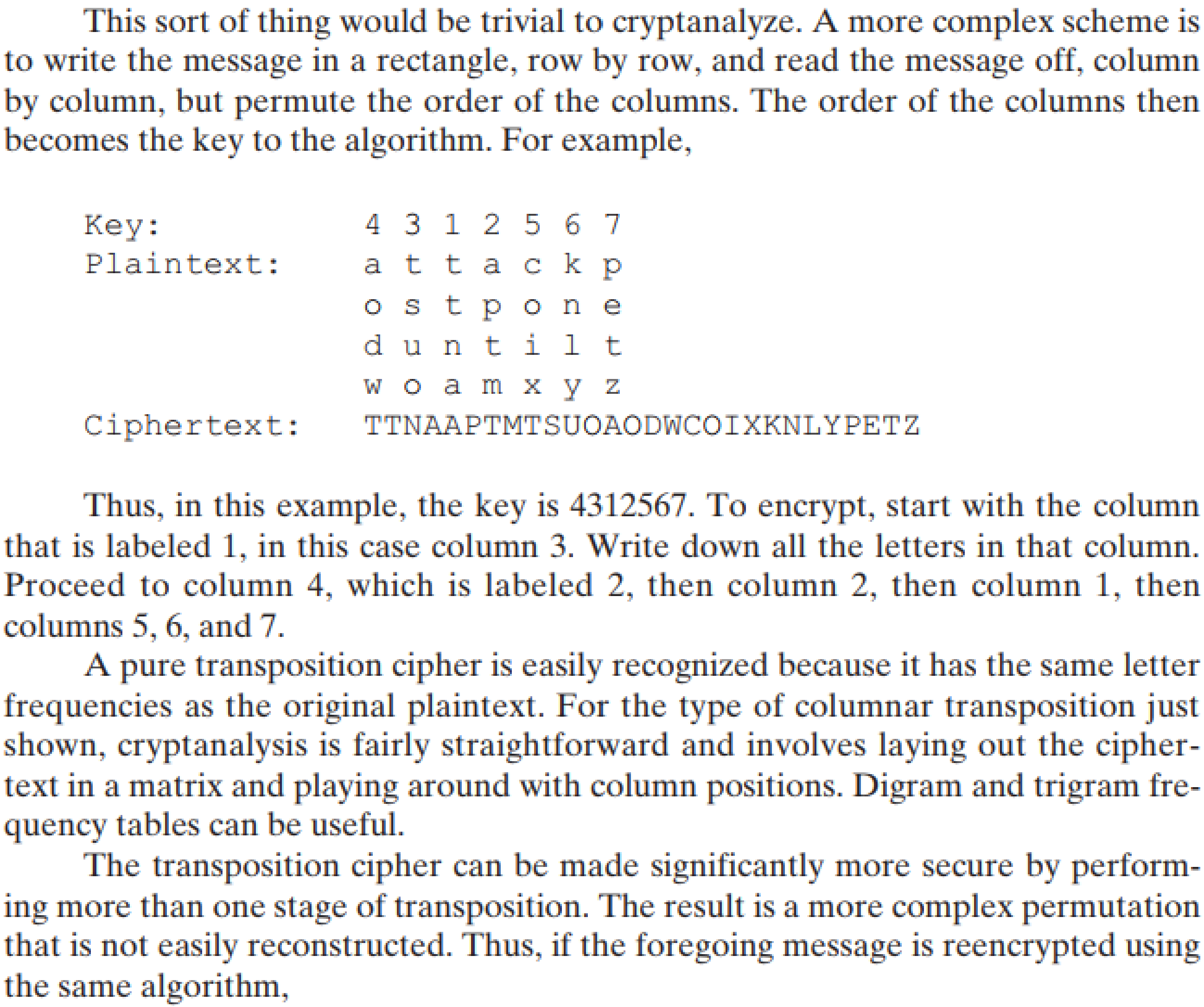
Clear text: defend the east wall

Rail Fence Technique

- Quiz: write the following sentence using Rail Fence technique (the key here is 3)

“defend the east wall of the castle”

Transposition Cipher



This sort of thing would be trivial to cryptanalyze. A more complex scheme is to 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. For example,

```
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
```

Thus, in this example, the key is 4312567. To encrypt, start with the column that is labeled 1, in this case column 3. Write down all the letters in that column. Proceed to column 4, which is labeled 2, then column 2, then column 1, then columns 5, 6, and 7.

A pure transposition cipher is easily recognized because it has the same letter frequencies as the original plaintext. For the type of columnar transposition just shown, cryptanalysis is fairly straightforward and involves laying out the ciphertext in a matrix and playing around with column positions. Digram and trigram frequency tables can be useful.

The transposition cipher can be made significantly more secure by performing more than one stage of transposition. The result is a more complex permutation that is not easily reconstructed. Thus, if the foregoing message is reencrypted using the same algorithm,

Transposition Cipher

- Quiz: encrypt and decrypt the following sentence using a Transposition Cipher with key = 2413.

“dns is great”

<u>Step:</u>

1. Encryption process
2. Cipher text = ?
3. Decryption process
4. Plain text = ?