CS 413
Information Security

Course Instructor

Naushad Siddiqui

Head of IT Audit, Sui Southern Gas Company

Visiting Faculty, NED University

Director & Faculty Member, ISACA Karachi

CS413 Information Security

Fall 2020

Week 2

Agenda

- Cryptography
- Cryptanalysis
- Dimensions of Cryptographic System
- Steganography
- Symmetric Encipherment
- Asymmetric Encipherment
- Caesar Cipher
- Cryptanalysis of Caesar Cipher
- Monoalphabetic Cipher

Security Mechanism Techniques

Cryptography

Key Terms

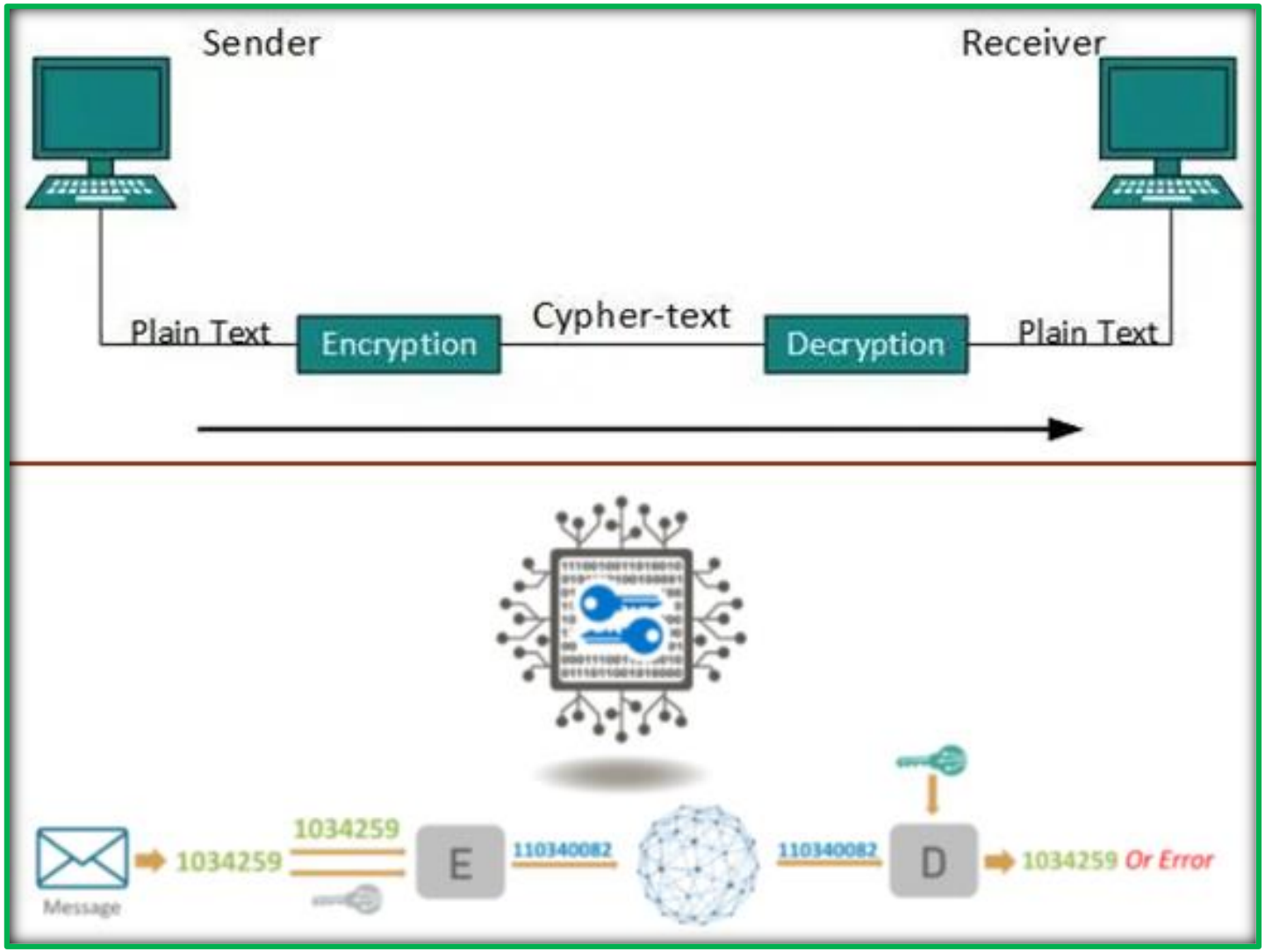
- Plain Text
  - an unencrypted original message
- Cipher Text
  - an encrypted message or coded message
- Cipher
  - algorithm for transforming plaintext to ciphertext
  - used to encrypt & decrypt the message
  - encryption algorithm performs transformations on plain text
  - decryption algorithm is inverse of encryption algorithm
- Key
  - information used in cipher, known only to sender / receiver
  - determines the output of the cipher algorithm and is needed to encrypt and decrypt a message

Cryptography (2)

Key Terms

- Encipher (encrypt)
  - converting plaintext to ciphertext
- Decipher (decrypt)
  - recovering plaintext from ciphertext
- Cryptography
  - study of encryption principles/methods
- Cryptanalysis (codebreaking)
  - study of principles/ methods of deciphering ciphertext without knowing key
- Cryptology
  - field of both cryptography and cryptanalysis

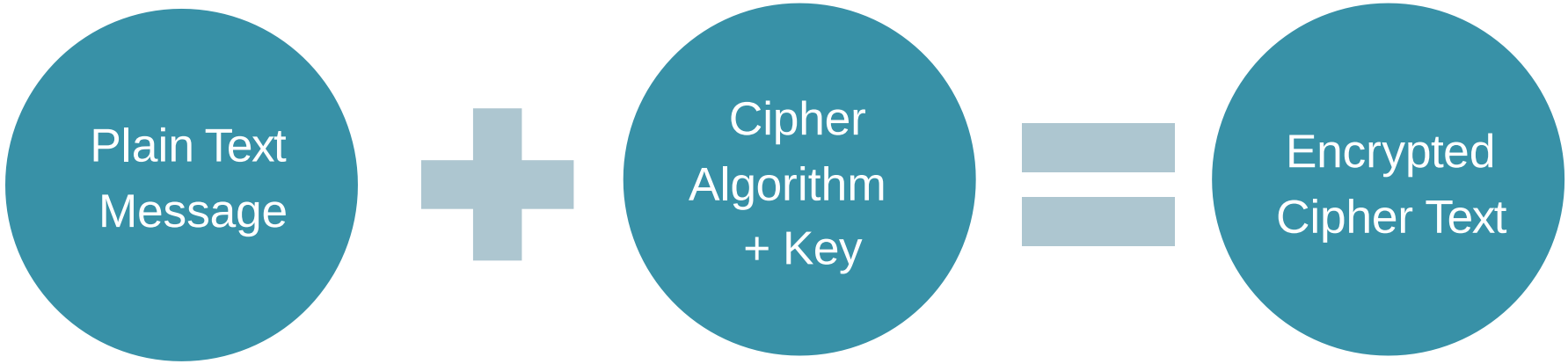
Cryptography (3)



Cryptography (4)

Encryption

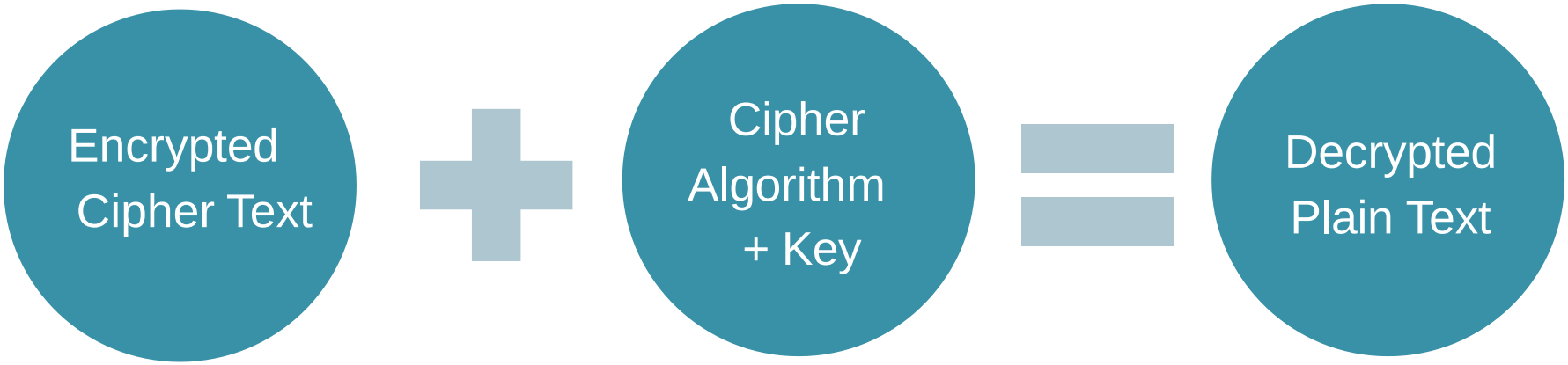
- A process that encodes a message or file so that it can only be read by certain people.
- Encryption uses an algorithm to scramble, or encrypt, data and then uses a key for the receiving party to unscramble, or decrypt, the information.



Cryptography (5)

Decryption

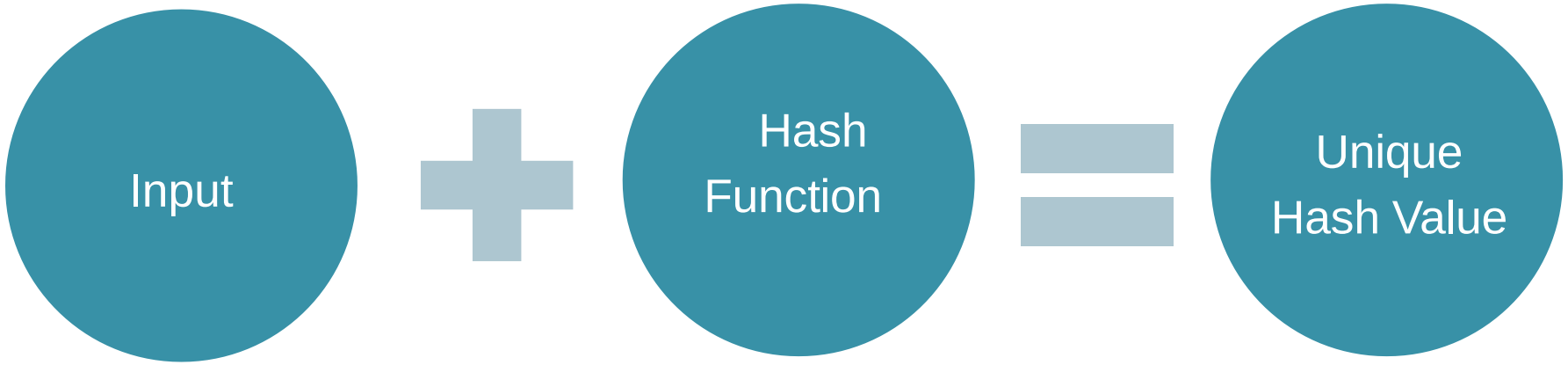
- The conversion of encrypted data into its original form is called Decryption.
- It is generally a reverse process of encryption.
- It decodes the encrypted information so that an authorized user can only decrypt the data because decryption requires a secret key or password.



Cryptography (6)

Hashing

- Hashing is the process of converting an input (data) into a fixed size string of text.
- It's a one-way function, meaning you can't use a hash value to determine its input data.
- Hashing is used to provide data integrity because each unique input will have a unique output.
- We use hashing to verify that something has not been tampered with.



Cryptanalysis

Cryptanalysis Definition

- Cryptanalysis is the study of ciphertext, ciphers and cryptosystems with the aim of understanding how they work and finding and improving techniques for defeating or weakening them.
- Cryptanalysis is the process of studying cryptographic systems to look for weaknesses or leaks of information.
- A systematic approach of finding weaknesses in cryptographic algorithms and using these weaknesses to decipher the ciphertext without knowing the secret key.

Cryptanalysis (2)

Cryptanalyst

- ***Cryptanalyst*** seeks to decrypt ciphertexts without knowledge of the plaintext source, encryption key or the algorithm used to encrypt it.
- Cryptanalysts' research results are used by cryptographers to improve and strengthen or replace flawed algorithms.

Cryptanalysis (3)

Cryptanalysis Attacks

- Typical objective is to recover the key in use rather than simply to recover the plaintext of a single ciphertext.
- Two general approaches:
  1. Cryptanalytic attack
  2. Brute-force attack

- ***Cryptanalytic Attacks*** rely on the nature of the algorithm plus perhaps some knowledge of the general characteristics of the plaintext or even some sample plaintext-ciphertext pairs.
- ***Brute-force Attacks*** try every possible key on a piece of ciphertext until an intelligible translation into plaintext is obtained. On average, half of all possible keys must be tried to achieve success.

Cryptanalysis (4)

Brute Force Attack

- A brute-force attack involves trying every possible key until an intelligible translation of the ciphertext into plaintext is obtained.
- It is the most basic attack, and is proportional to key size
- On average, half of all possible keys must be tried to achieve success.

Key Size (bits)	Number of Alternative Keys	Time required at 1 decryption/μs	Time required at 10^6 decryptions/μs
32	$2^{32} = 4.3 \times 10^9$	2^{31} μs = 35.8 minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	2^{55} μs = 1142 years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	2^{127} μs = 5.4×10^{24} years	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	2^{167} μs = 5.9×10^{36} years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	2×10^{26} μs = 6.4×10^{12} years	6.4×10^6 years

This Table shows how much time is required to conduct a brute-force attack

Cryptanalysis (5)

Types of Attacks on Encrypted Messages

Type of Attack	Known to Cryptanalyst
Ciphertext only	• Encryption algorithm and Ciphertext
Known plaintext	• Encryption algorithm and Ciphertext • One or more plaintext-ciphertext pairs formed with the secret key
Chosen plaintext	• Encryption algorithm and Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key
Chosen ciphertext	• Encryption algorithm and Ciphertext • Purported ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
Chosen text	• Encryption algorithm and Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key • Purported ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key

Dimensions of Cryptographic System

(1) The types of operations used

This dimension covers the type of operations used for transforming plaintext to ciphertext

- All encryption algorithms are based on two general principles
  - ***Substitution:*** in which each element in the plaintext (bit, letter, group of bits or letters) is mapped into another element
  - ***Transposition:*** in which elements in the plaintext are rearranged. The fundamental requirement is that no information be lost (that is, that all operations are reversible)
- Most systems, referred to as product systems, involve multiple stages of substitutions and transpositions

Dimensions of Cryptographic System (2)

(2) The number of keys used

- ***Symmetric:*** If both sender and receiver use the same key, the system is referred to as symmetric, single-key, secret-key, or conventional encryption.
- ***Asymmetric:*** If the sender and receiver use different keys, the system is referred to as asymmetric, two-key, or public-key encryption.

(3) The way in which the plaintext is processed

- A ***block cipher*** processes the input one block of elements at a time, producing an output block for each input block.
- A ***stream cipher*** processes the input elements continuously, producing output one element at a time, as it goes along.

Steganography

- Steganography is the technique of hiding secret data within an ordinary, non-secret, file or message in order to avoid detection; the secret data is then extracted at its destination.
- The use of steganography can be combined with encryption as an extra step for hiding or protecting data
- It is the art and science of hiding information by embedding messages within other, seemingly harmless messages.

Steganography (2)

Various techniques have been used historically; some examples are as follows:

- The sequence of first letters of each word of the overall message spells out the hidden message.
- Character marking: Selected letters of printed or typewritten text are overwritten in pencil. The marks are ordinarily not visible unless the paper is held at an angle to bright light.
- Invisible ink: A number of substances can be used for writing but leave no visible trace until heat or some chemical is applied to the paper.
- Pin punctures: Small pin punctures on selected letters are ordinarily not visible unless the paper is held up in front of a light.
- Typewriter correction ribbon: Used between lines typed with a black ribbon, the results of typing with the correction tape are visible only under a strong light.

Steganography (3)

Advantages:

- It can be employed by parties who have something to lose should the fact of their secret communication (not necessarily the content) be discovered.
- Encryption flags traffic as important or secret or may identify the sender or receiver as someone with something to hide.

Disadvantages:

- It requires a lot of overhead to hide a relatively few bits of information, although using some scheme may make it more effective.
- Once the system is discovered, it becomes virtually worthless.

Steganography (4)

Steganography in Old Days

Steganography in Modern Days

Self Study

Symmetric & Asymmetric Encipherment

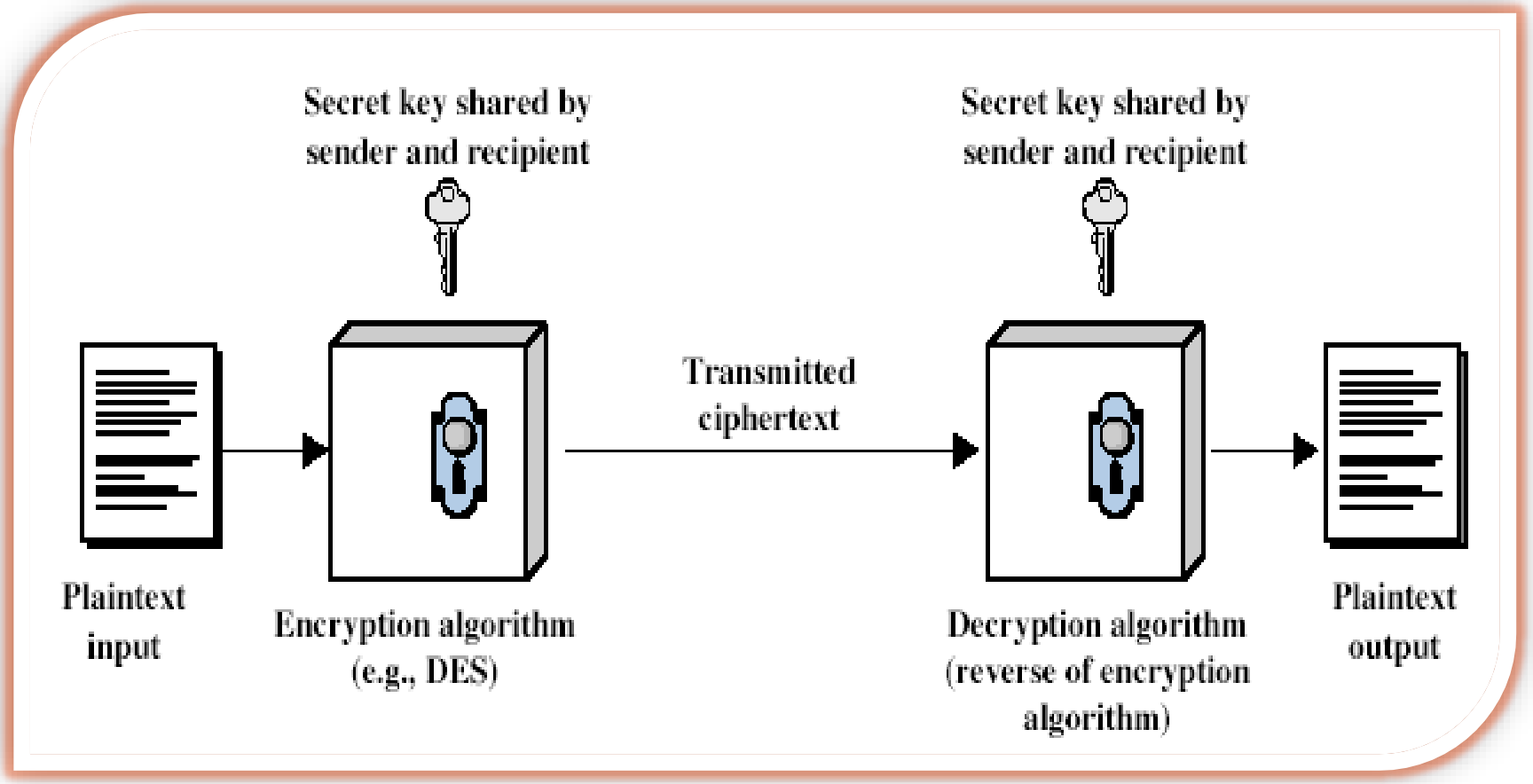
Symmetric Encipherment

- Symmetric encryption, also referred to as conventional encryption or single-key encryption, was the only type of encryption in use prior to the development of public-key encryption in the 1970s.
- It remains by far the most widely used of the two types of encryption.
- All traditional schemes are symmetric / single key / private-key encryption algorithms, with a single key, used for both encryption and decryption.
- It’s also known as secret-key encryption or private-key encryption.

Symmetric Encipherment (2)

- Both the sender and receiver share the same key and use it to encrypt and decrypt all messages.
- Since both sender and receiver are equivalent, either can encrypt or decrypt messages using that common key.
- Symmetric encryption is much more efficient at encrypting large amounts of data than its counterpart, asymmetric encryption.
- The downside of symmetrical encryption is that it makes it hard to initiate communication the first time. How to securely transmit the private key to each user?

Symmetric Encipherment (3)



Symmetric Encipherment (4)

- It is impractical to decrypt a message on the basis of the cipher- text plus knowledge of the encryption/decryption algorithm
- Do not need to keep the algorithm secret, we only need to keep the key secret.
- This feature of symmetric encryption is what makes it feasible for widespread use.
- Mathematically it can be considered a pair of functions

$$Y = E_K(X)$$

$$X = D_K(Y)$$

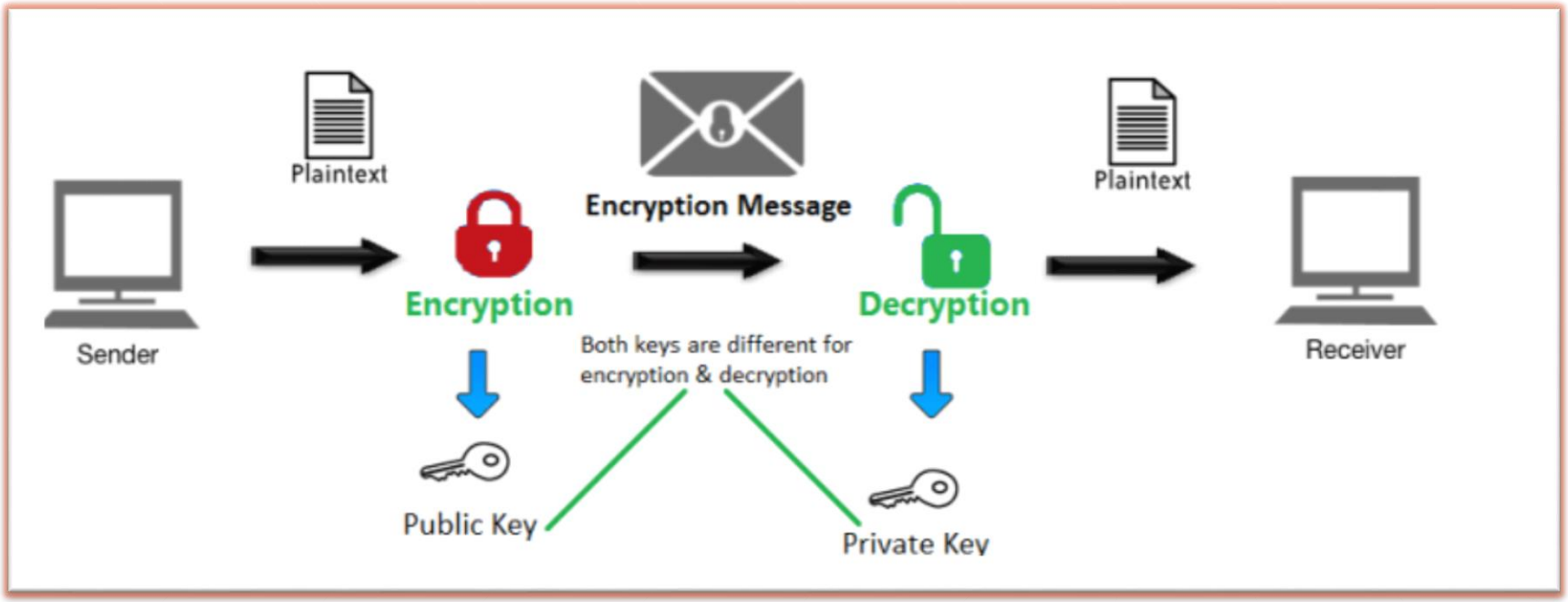
where

- X is plaintext
- Y is ciphertext
- K is key
- E_K is encryption algorithm
- D_K is decryption algorithm

Asymmetric Encipherment

- Asymmetric encryption uses two keys, a public key and a private key created as a matched pair.
  - Private Key: Kept secret and never shared.
  - Public Key: Shared with others.
- Commonly referred to as:
  - Public Key Encryption
  - Public Key Infrastructure (PKI) Encryption
- Anything encrypted with the private key can only be decrypted with the matched public key.
- Anything encrypted with the public key can only be decrypted with the matched private key.

Asymmetric Encipherment (2)



Asymmetric Encipherment (3)

- When using asymmetric encryption, both sender and receiver have to generate a key pair on their computers. This is done using some algorithm, such as RSA, Diffie-Hellman or ElGamal.
- These algorithm will generate a pair of public and private keys that are mathematically linked to each other.
- Asymmetric cryptography is typically used to authenticate data using digital signatures.
- A digital signature is a mathematical technique used to validate the authenticity and integrity of a message, software or digital document.
- It is the digital equivalent of a handwritten signature or stamped seal.

Classical Substitution Ciphers

Classical Substitution Ciphers

- Classical Encryption techniques enables us to illustrate the basic approaches to symmetric encryption used today and the types of cryptanalytic attacks that must be anticipated.
- The two basic building blocks of all encryption technique are ***Substitution*** and ***Transposition***.
- In a **Substitution Cipher**, any character of plain text from the given fixed set of characters is substituted by some other character from the same set depending on a key.
- For example with a shift of 1, A would be replaced by B, B would become C, and so on
- Letters of plaintext are replaced by other letters or by numbers or symbols
- If plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns

Caesar Cipher

- The Caesar cipher is one of the earliest known and simplest ciphers.
- It is a type of substitution cipher in which each letter in the plaintext is 'shifted' a certain number of places down the alphabet.
- The early Greeks described several substitution ciphers, the first attested use in military affairs of one was by Julius Caesar.

- Example:

```
meet me after the toga party
PHHW PH DIWHU WKH WRJD SDUWB
```

Caesar Cipher (2)

Let's 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

Then the algorithm can be expressed as follows. For each plaintext letter *p*, substitute the ciphertext letter *C*:

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

We define a mod n to be the remainder when a is divided by n. For example, 11 mod 7 = 4.

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

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

Caesar Cipher (3)

- This mathematical description uses modulo (clock) arithmetic. Here, when you reach Z you go back to A and start again. Mod 26 implies that when you reach 26, you use 0 instead (i.e. the letter after Z, or 25 + 1 goes to A or 0).
- Example: howdy (7,14,22,3,24) encrypted using key f (i.e. a shift of 5) is MTBID

Cryptanalysis of Caesar Cipher

- With a Caesar cipher, there are only 26 possible keys, of which only 25 are of any use, since mapping A to A etc doesn't really obscure the message! i.e. A maps to A,B,..Z
- Could simply try each in turn
- In a brute force search, for any given ciphertext, just try all shifts of letters
- Need to be able to recognize when have an original message (i.e. is it English or whatever).
- Can try each of the keys (shifts) in turn, until can recognize the original message.
- Example "GCUA VQ DTGCM"

When broken gives "easy to break", with a shift of 2 (key C).

Cryptanalysis of Caesar Cipher (2)

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

Cryptanalysis of Caesar Cipher (3)

Example:

- 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 follows

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

Monoalphabetic Cipher

- With only 25 possible keys, the Caesar cipher is far from secure.
- A dramatic increase in the key space can be achieved by allowing an arbitrary substitution, where the translation alphabet can be any permutation of the 26 alphabetic characters.
- Rather than just shifting the alphabet, could shuffle (jumble) the letters arbitrarily
- A monoalphabetic cipher is any cipher in which the letters of the plain text are mapped to cipher text letters based on a single alphabetic key.
- Examples of monoalphabetic ciphers would include the Caesar-shift cipher, where each letter is shifted based on a numeric key, and the atbash cipher, where each letter is mapped to the letter symmetric to it about the center of the alphabet.

Monoalphabetic Cipher (2)

- Monoalphabetic cipher is one where each symbol in plain text is mapped to a fixed symbol in cipher text.
- The relationship between a character in the plain text and the characters in the cipher text is one-to-one.
- Each alphabetic character of plain text is mapped onto a unique alphabetic character of a cipher text.
- Monoalphabetic Cipher is described as a substitution cipher in which the same fixed mappings from plain text to cipher letters across the entire text are used.

Monoalphabetic Cipher (3)

- Each plaintext letter maps to a different random ciphertext letter, hence key is 26 letters long

```
Plain:  abcdefghijklmnopqrstuvwxyz
Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

Plaintext:  ifwewishtoreplaceletters
Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA
```

Monoalphabetic Cipher (4)

Monoalphabetic Cipher Security

- Note that even given the very large number of keys, being 10 orders of magnitude greater than the key space for DES, the Mono-Alphabetic substitution cipher is not secure, because it does not sufficiently obscure the underlying language characteristics.
- Although having a total of 26! = 4 x 1026 keys
- With so many keys, might think is secure …
- … but would be !!!WRONG!!!
- Problem is language characteristics

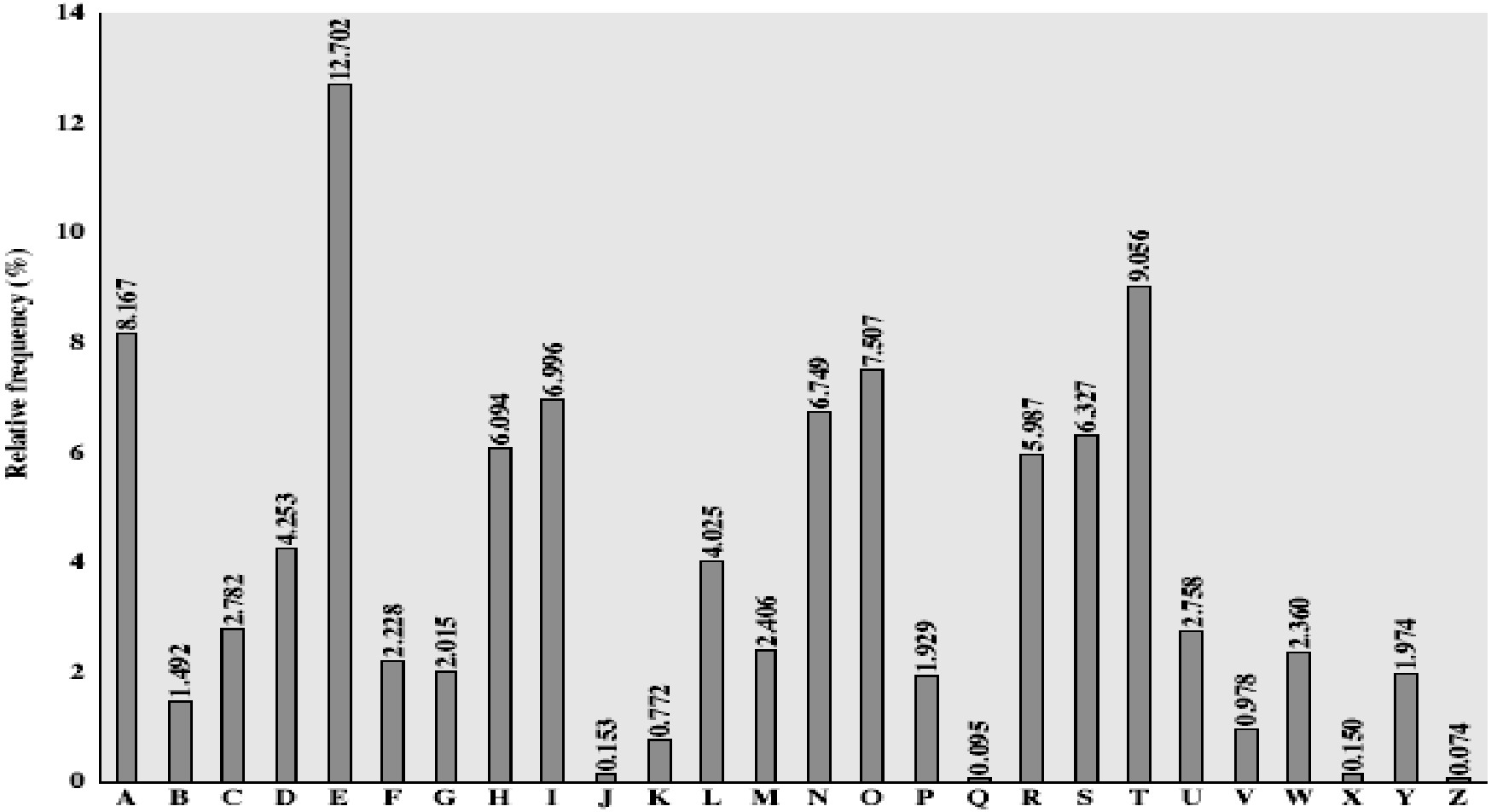
Monoalphabetic Cipher (5)

Language Redundancy and Cryptanalysis

- Human languages are redundant e.g. "th lrd s m shp hrd shll nt wnt"
- Letters are not equally commonly used
- In English, E is by far the most common letter followed by T,R,N,I,O,A,S
- Other letters like Z,J,K,Q,X are fairly rare
- As the example shows, we don't actually need all the letters in order to understand written English text.
- Here vowels were removed, but they're not the only redundancy.
- Basic idea is to count the relative frequencies of letters, and note the resulting pattern.

Monoalphabetic Cipher (6)

English Letter Frequencies



Monoalphabetic Cipher (7)

Use in Cryptanalysis

- Key concept - monoalphabetic substitution ciphers do not change relative letter frequencies
- Discovered by Arabian scientists in 9th century
- Calculates letter frequencies for ciphertext
- Compare counts/plots against known values
- If Caesar cipher look for common peaks / troughs

```
peaks at: A-E-I triple, NO pair, RST triple
troughs at: JK, X-Z
```

- For monoalphabetic must identify each letter

```
tables of common double/triple letters help
```

Monoalphabetic Cipher (8)

Use in Cryptanalysis

- The simplicity and strength of the monoalphabetic substitution cipher meant it dominated cryptographic use for the first millennium AD.
- It was broken by Arabic scientists. The earliest known description is in Abu al-Kindi's "A Manuscript on Deciphering Cryptographic Messages", published in the 9th century but only rediscovered in 1987 in Istanbul
- Monoalphabetic ciphers are easy to break because they reflect the frequency data of the original alphabet.
- The cryptanalyst looks for a mapping between the observed pattern in the ciphertext, and the known source language letter frequencies.

Monoalphabetic Cipher (9)

Example

- Given ciphertext:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

- Count relative letter frequencies

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

- Comparing this breakdown with English Letter Frequency chart, it seems likely that cipher letters P and Z are the equivalents of plain letters e and t, but it is not certain which is which.
- The letters S, U, O, M, and H are all of relatively high frequency and probably correspond to plain letters from the set {a, h, i, n, o, r, s}.

Monoalphabetic Cipher (10)

***Example** (…cont…)*

- The letters with the lowest frequencies (namely, A, B, G, Y, I, J) are likely included in the set {b, j, k, q, v, x, z}
- A powerful tool is to look at the frequency of two-letter combinations, known as ‘digrams’. A table could be drawn up showing the relative frequency of digrams.
- The most common such digram’ is ‘th’. 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
- Guess P & Z are e and t
- Guess ZW is ‘th’ and hence ZWP is ‘the’

Monoalphabetic Cipher (11)

***Example** (…cont…)*

```
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
```

- 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
```

End of Week 02