Information Security

CS3002

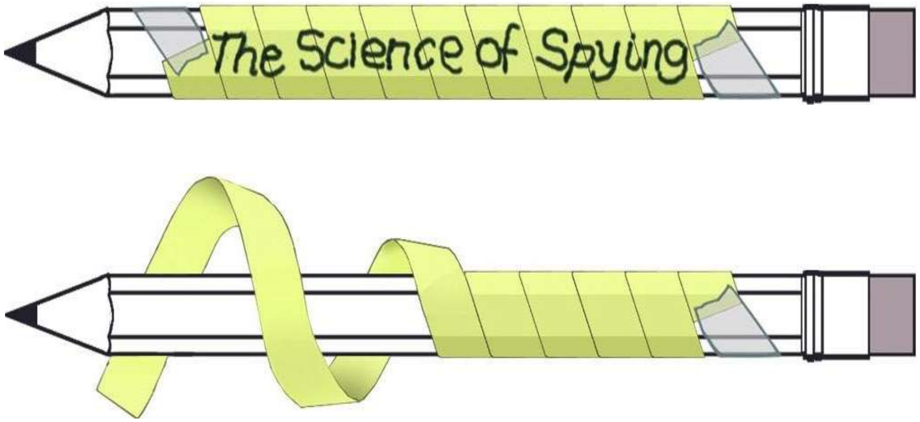
Lecture 2

29th August 2023

Dr. Rana Asif Rehman

Email: r.asif@nu.edu.pk

Old Cryptography Methods



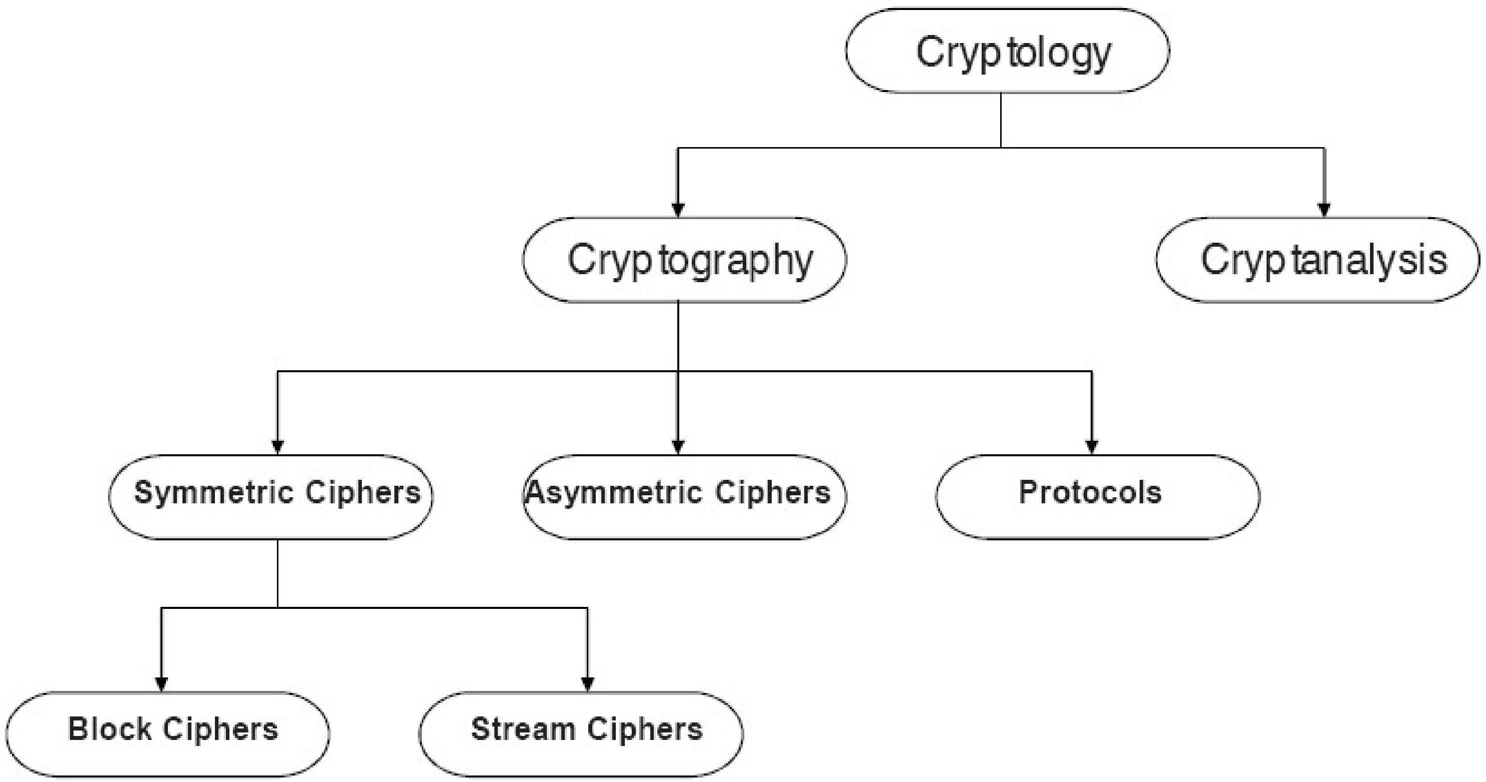
Some Basic Terminology

- **Plaintext** - original message
- **Ciphertext** - coded message
- **Encryption algorithm** - algorithm for transforming plaintext to ciphertext
- **Key** - info used in cipher known only to sender/receiver
- **Encipher (encrypt)** - converting plaintext to ciphertext

Some Basic Terminology (cont.)

- **Decryption algorithm** - algorithm for transforming ciphertext to plaintext
- **DeCipher –** converting ciphertext to plaintext .
- **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

Classification of the Field of Cryptology

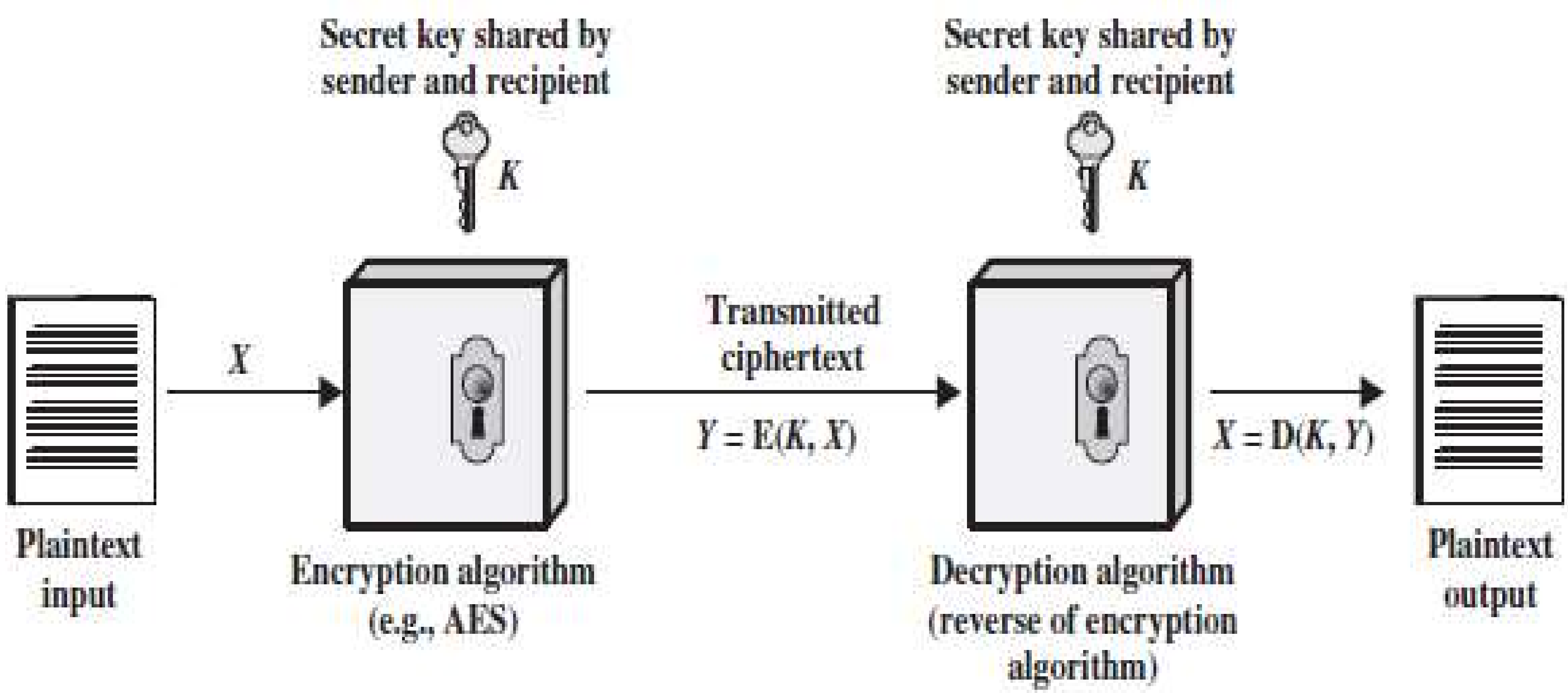


Adopted with thanks from: Chapter 1 of Understanding Cryptography by Christof Paar and Jan Pelzl

Symmetric Encryption

- Or conventional / private-key / single-key
- Sender and recipient share a common key
- All classical encryption algorithms are private-key
- Was only type prior to invention of public-key in 1970's
- And by far most widely used

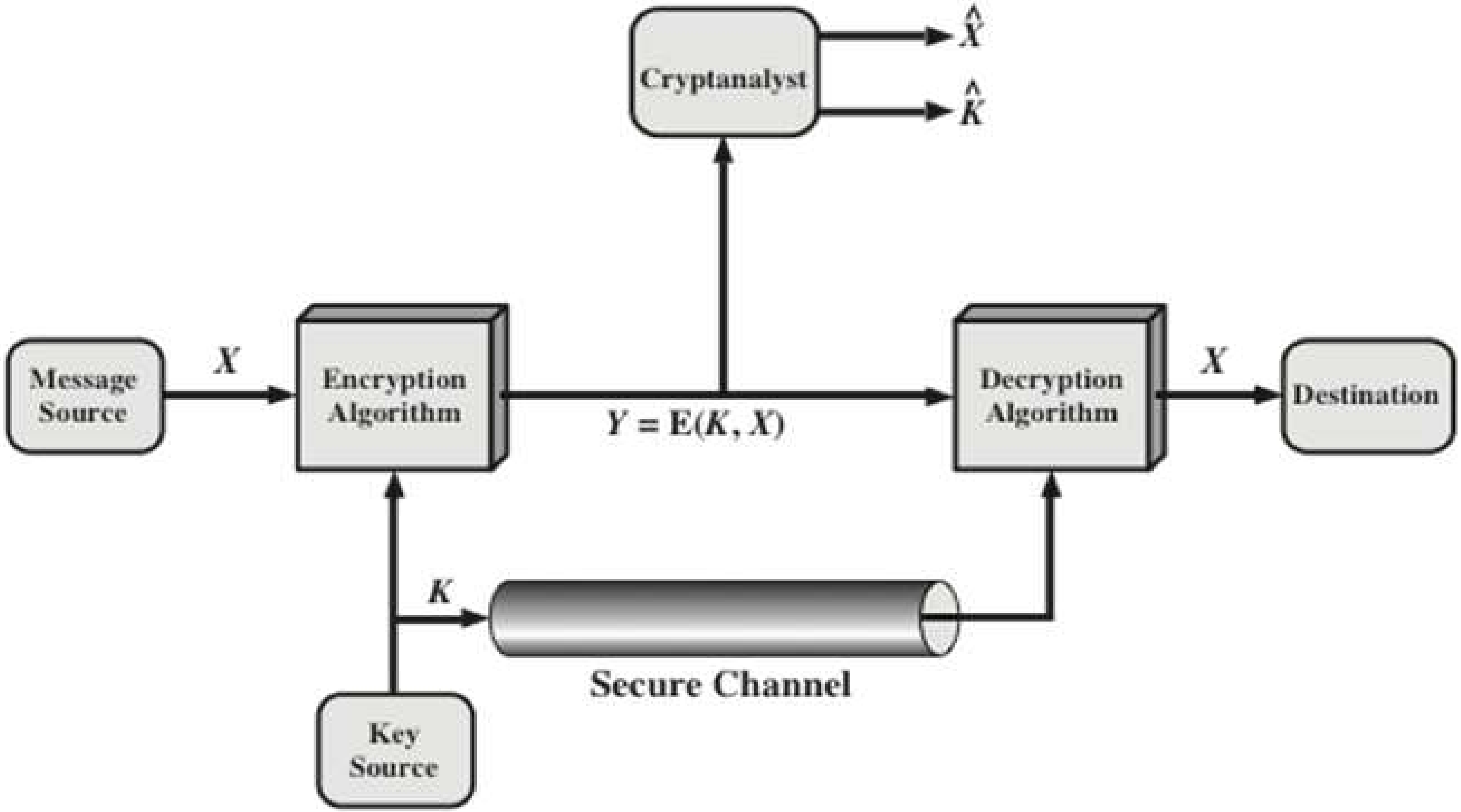
Simplified Model of Symmetric Encryption



Requirements

- Two requirements for secure use of symmetric encryption:
  - A strong encryption algorithm
    - The opponent should be unable to decrypt ciphertext or discover the key even if he or she is in possession of a number of ciphertexts together with the plaintext that produced each ciphertext.
  - A secret key known only to sender / receiver
    - Sender and receiver must have obtained copies of the secret key in a secure fashion and must keep the key secure.

Model of Symmetric Cryptosystem



Model of Symmetric Cryptosystem (cont.)

With the message X and the encryption key K as input, the encryption algorithm forms the ciphertext $Y = [Y_1, Y_2, \ldots, Y_N]$. We can write this as

$$Y = \mathrm{E}(K, X)$$

The intended receiver, in possession of the key, is able to invert the transformation:

$$X = \mathrm{D}(K, Y)$$

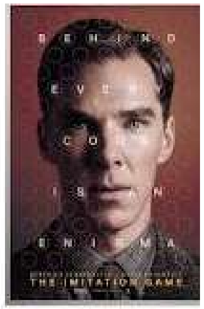
An opponent, observing *Y* but not having access to *K* or *X*, may attempt to recover *X* or *K* or both *X* and *K*.

Cryptography

- Characterize cryptographic system by:
  - Type of encryption operations used
    - Substitution / transposition / both (product systems)
  - Number of keys used
    - Single-key or private / two-key or public
  - Way in which plaintext is processed
    - Block / Stream

Turing's work during the Second World War was so crucial that Winston Churchill acknowledged his role, saying that Turing made the single biggest contribution to Allied victory.

The Imitation Game
2014 · Drama/Thriller · 1h 54m



8/10
IMDb

91%
Rotten Tomatoes

92% liked this film
Google users

In 1939, newly created British intelligence agency MI6 recruits Cambridge mathematics alumnus Alan Turing (Benedict Cumberbatch) to crack Nazi codes, including Enigma -- which cryptanalysts had thought unbreakable. Turing's team, including Joan Clarke (Keira Knightley), analyze Enigma messages while he builds a machine to decipher them. Turing and team finally succeed and become heroes, but in 1952, the quiet genius encounters disgrace when authorities reveal he is gay and send him to prison.

Initial release: 7 January 2015 (Egypt)

Director: Morten Tyldum

Box office: 233.6 million USD

Awards: Academy Award for Best Writing Adapted Screenplay, MORE

Nominations: Academy Award for Best Picture, MORE

http://www.iwm.org.uk/history/how-alan-turing-cracked-the-enigma-code

Classical Substitution Ciphers

- Where letters of plaintext are replaced by other letters or by numbers or symbols.
- Or if plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns.

Classical Symmetric Ciphers

- Classical Substitution Ciphers
  - Caesar Cipher
  - Monoalphabetic Cipher
  - Playfair Cipher
  - Hill Ciphers
  - Polyalphabetic Ciphers
    - Vigenere Cipher
    - Vernam Cipher
  - One-Time Pad

1. Caesar Cipher

- Earliest known substitution cipher.
- First attested use in military affairs.
- Replaces each letter by 3rd letter on
- Example:

```
meet me after the toga party
phhw ph diwhu wkh wrjd sduwb
```

Caesar Cipher (cont.)

- Can define transformation as:

```
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
```

- Then have caesar cipher as:

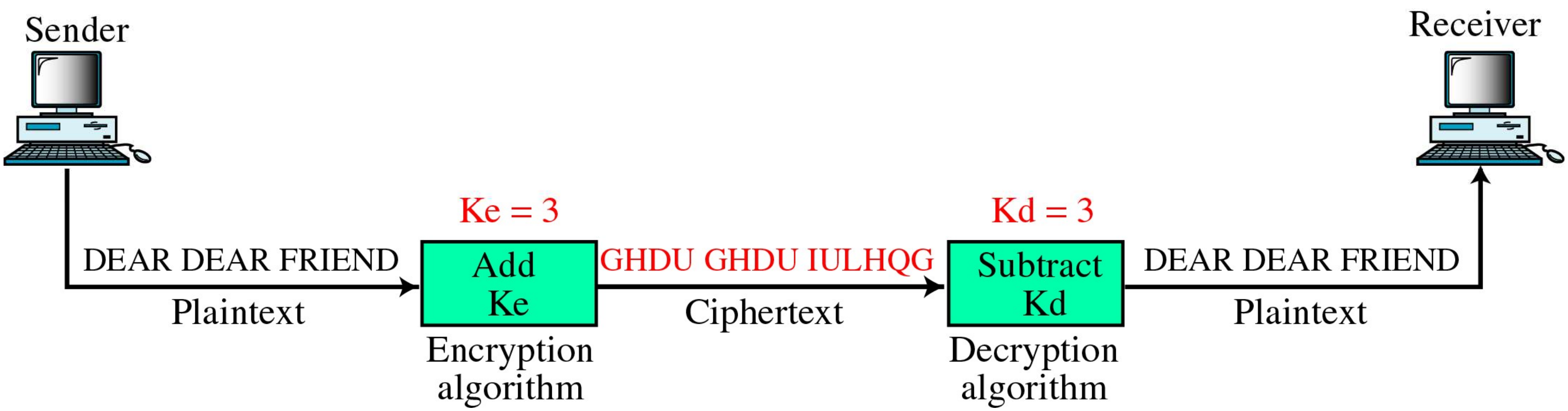
a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12

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

c = E(k, p) = (p + k) mod (26)

p = D(k, c) = (c – k) mod (26)

Caesar Cipher (cont.)



Cryptanalysis of Caesar Cipher

- Only have 25 possible Keys
  - A maps to B,C,..Z
- Could simply try each in turn
- A **brute force search**

  1. The encryption and decryption algorithms are known.

  2. There are only 25 keys to try.

  3. The language of the plaintext is known and easily recognizable.

  - e.g. Break ciphertext "GCUA VQ DTGCM"

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

Brute Force Feasibility

- In most networking situations, we can assume that the algorithms are known.
- What generally makes brute-force cryptanalysis impractical is the use of an algorithm that employs a large number of keys.
  - For example, the triple DES algorithm makes use of a 168-bit key, giving a key space of 2^{168} or greater than $3.7 * 10^{50}$ possible keys.

2. Monoalphabetic Cipher

- With only 25 possible keys, the Caesar cipher is far from secure.
- Rather than just shifting the alphabet could shuffle (jumble) the letters arbitrarily [permutation]
- Each plaintext letter maps to a different random ciphertext letter
- Hence key is 26 letters long

```
Plain:   ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

Plaintext:   IFWEWISHTOREPLACELETTERS
Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA
```

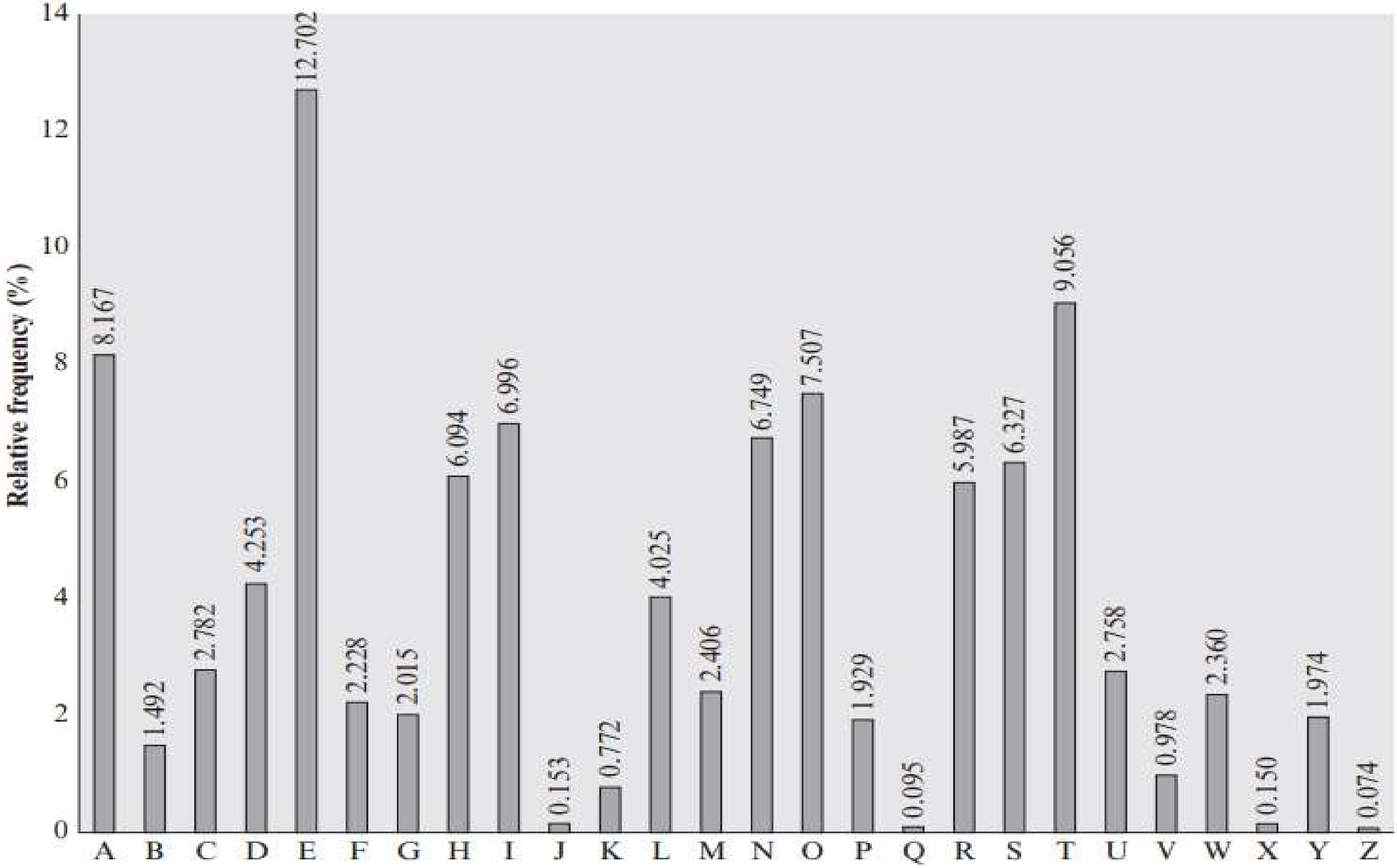
Monoalphabetic Cipher Security

- Now have a total of $26! = 4 \times 10^{26}$ keys
- With so many keys, might think is secure
- But would be **!!!WRONG!!!**
- Problem is language characteristics

Language Redundancy and Cryptanalysis

- Human languages are **redundant** e.g. "Th lrd s m shphrd 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
- Have tables of single, double & triple letter frequencies for various languages

English Letter Frequencies



Use in Cryptanalysis

- Key concept - monoalphabetic substitution ciphers do not change relative letter frequencies.
- Calculate 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

Language Redundancy and Cryptanalysis

- Cipher Text

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

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								

Example Cryptanalysis

- given cipher text:

  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

So far, then, we have

```
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
```

Example Cryptanalysis (cont.)

- 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 Ciphers

- Monoalphabetic ciphers are easy to break because they reflect the frequency data of the original alphabet.
- Smooth out frequency distribution by assigning multiple ciphertext codes (numbers) to each letter in the plaintext alphabet.
- If the number of symbols assigned to each letter is proportional to the relative frequency of that letter, then single-letter frequency information is completely obliterated.
  – Multiple letter patterns can still survive.
- To encrypt a character *m*, pick one of the codes at random

P-40

Homophonic Substitution Ciphers - Example

- Map letters to codes 00, ..., 99

Letter	Homophones
A	17 19 34 41 56 60 67 83
I	08 22 53 65 88 90
L	03 44 76
N	02 09 15 27 32 40 59
O	01 11 23 28 42 54 70 80
P	33 91
T	05 10 20 29 45 58 64 78 99

- encipher P L A I N P I L O T

  91 44 56 65 59 33 08 76 28 78