Diyala University
College of Engineering
Computer & Software
Engineering Department

Fourth Year 2012/2013

ENCRYPTION AND DATA SECURITY

Chapter2: Part1

PRESENTED BY
DR. ALI J. ABBOUD

TEXTBOOKS

- **Applied Cryptography – 2nd edition**
  Bruce Schneier Available online.
- **Security In Computing – 4th edition**
  Pfleeger and Pfleeger.
- http://www.wikipedia.org/
- **Papers assigned for reading**

CLASS STRUCTURE

- **Chapter 1**: **Introduction to encryption and Data security.**
- ***Chapter 2: Encryption systems.***
- **Chapter 3**: **Secure encryption systems.**
- **Chapter 4**: **Security involving programs.**
- **Chapter 5**: **Protection Services for users of operating systems.**
- **Chapter 6**: **Data Base security.**
- **Chapter 7**: **Computer Network Security.**
- **Chapter 8**: **Communication Security.**

Chapter 2: Encryption Systems

- ***2.1 Terminology and Background***
- ***2.2 Substitution Ciphers***
- ***2.3 Transpositions (Permutations)***
- ***2.4 Making "Good" Encryption Algorithms***
- ***2.5 The Data Encryption Standard (DES)***
- ***2.6 The AES Encryption Algorithm***
- ***2.7 Public Key Encryption***
- ***2.8 Uses of Encryption***
- ***2.9 Summary***

In This Chapter

important tool

- rooted in some heavy-duty math
  - number theory
  - group & field theory
  - computational complexity
  - probability

- our goal:
  - be able to intelligently use cryptosystems
  - not design/break cryptosystems

Text's Notation

- S *sender*
- R *recipient*
- T *trans. medium*
- O *outsider* or *intruder*

O might try to:

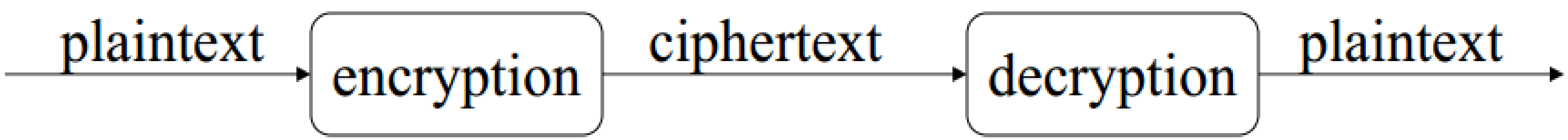
- block
- intercept
- modify
- fabricate

Block it, by preventing its reaching R, thereby affecting the availability of the message.
Intercept it, by reading or listening to the message, thereby affecting the confidentiality of the message.
Modify it, by seizing the message and changing it in some way, affecting the message's integrity.
Fabricate an authentic-looking message, arranging for it to be delivered as if it came from S, thereby also affecting the integrity of the message.

Terminology

- encryption (or *encipher*)
- decryption (or *decipher*)
- note: encode/decode different meaning
- plaintext
- ciphertext

Graphical View



Notation

- denote plaintext $P = <p_1,p_2, \ldots , p_n>$
- denote ciphertext $C = <c_1,c_2, \ldots , c_n>$
- Example:
  - plaintext “I like cheesy poofs”
  - P = <I, ,L,I,K,E, ,C,H,E,E,S,Y, ,P,O,O,F,S>
  - ciphertext “X QXVC JMCCZB ARREZ”
  - C = <X, ,Q,X,V,C, ,J,M,C,C,Z,B, ,A,R,R,E,Z>
- More formally:
  - $C=E(P)$ $\qquad$ $P=D(C)$
  - $P=D(E(P))$

How Codes Are Different

- code uses *linguistic units*
- codebook is the key

word	code
bored	9685
car	2307
poultry	7902
students	1092
.	.
.	.

bored students

↓

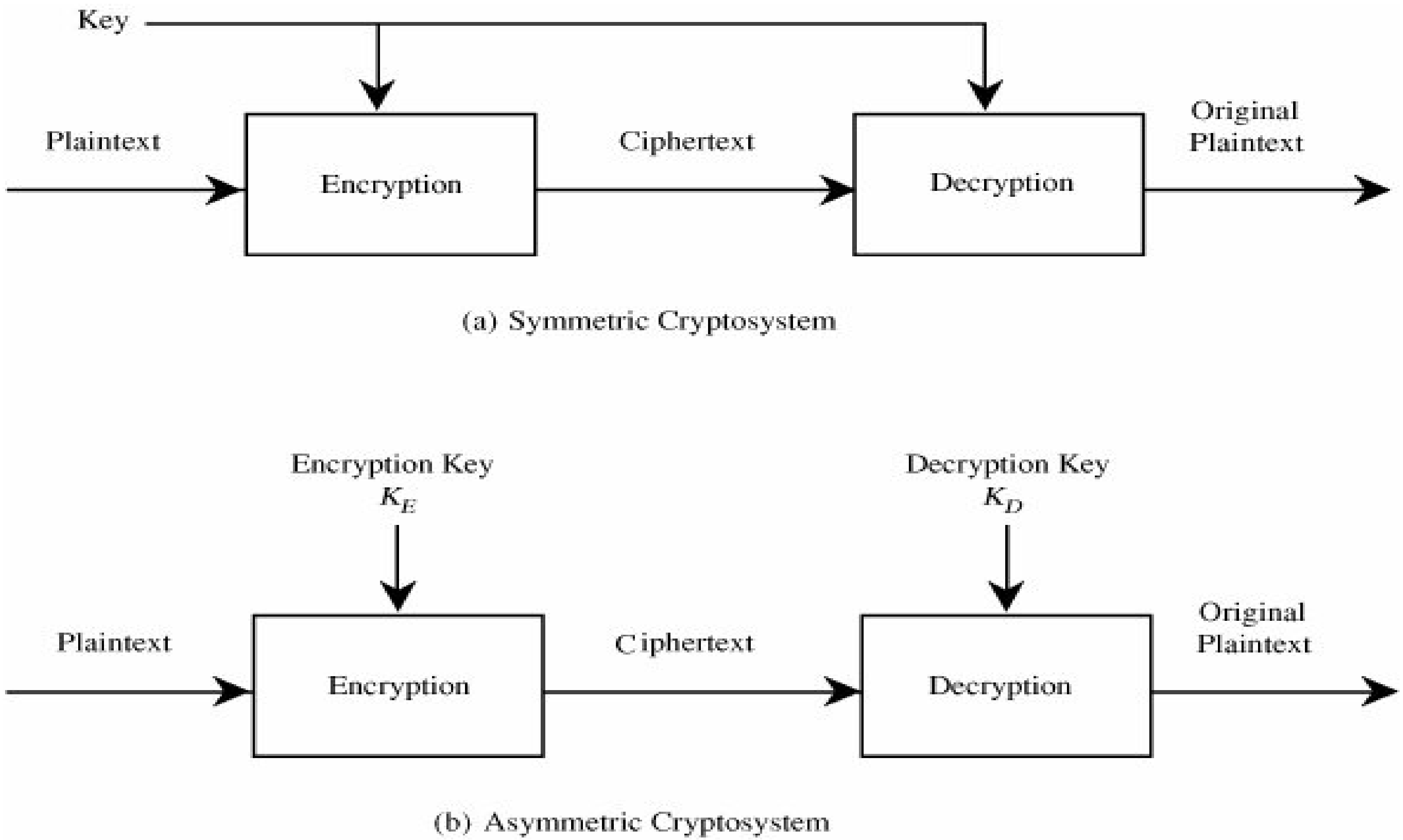
9685 1092

- may use phrases as well
- *e.g.*, "return to base for supplies" enciphered GIDIZZLEDUNK

Cryptographic Keys

- most algorithms use keys
- encryption:
  - $C = E(K, P)$
  - $P = D(K, C)$
  - $P = D(K, E(K,P))$
- Cryptographic algorithm (aka *cipher*)
  - mathematical function used for encrypt
- Cryptosystem consists of:
  - cryptographic algorithm
  - set of all possible plaintexts
  - set of all possible ciphertexts

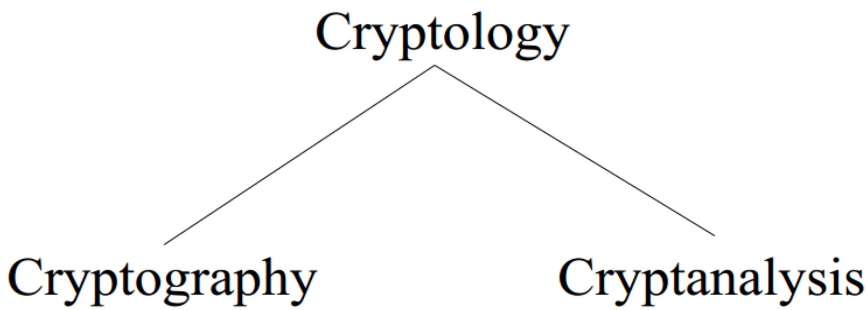
Encryption Algorithms



Asymmetric Algorithm

- encryption, decryption keys different
- encryption key: K_E
- decryption key: K_D
  - $C = E(K_E, P)$
  - $P = D(K_D, C)$
  - $P = D(K_D, E(K_E, P))$

Cryptanalysis



- Cryptanalyst tries to **break** an algorithm
- Categories (*due to Lars Knudsen*)
  - **total break** - find the key K such that D(K,C)=P
  - **global deduction** - find alternative algorithm, A, equivalent to D(K,C) without knowing K
  - **instance (or *local*) deduction** - find the plaintext of an intercepted ciphertext
  - **information deduction** - get some information about the key or plaintext, *e.g.*, first bits of the key, info about the form of the plaintext, …
- Attempt at cryptanalysis called an **attack**

How is Cryptanalysis Done?

- Analyst works with whatever is available:
  - encrypted messages
  - known algorithms
  - intercepted plaintext
  - known or suspected plaintext
  - properties of the likely plaintext
  - properties of computers
  - properties of network protocols

Character Arithmetic

- Usually don't consider case
- Can do arithmetic on letters
- Example: A+2, Y+5, *etc*.

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M
Code	0	1	2	3	4	5	6	7	8	9	10	11	12

Letter	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Code	13	14	15	16	17	18	19	20	21	22	23	24	25

- What if you go past the end, *e.g.* Y+3?

modular arithmetic – quick review

a and b are integers, $b \geq 1$

divide a by b (using regular long division)

result is:

- q (quotient)
- r (remainder or residue)

$a = qb + r$, where $0 \leq r < b$

$r = a \bmod b$

Cryptographic Elements

- Primitive operations:
  - **substitutions** - exchange one letter for another
  - **transpositions –** rearrange the order of the letters

Keyword Mixed Alphabet

- Form ciphertext alphabet by:
  - pick a keyword
  - spell it without duplicates
  - then, fill in the rest of the alphabet in order
- Example, keyword *VACATION*

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

- Encrypt "*I should be sailing*" as:
  - DQBK SGTA IQVD GDJN

Another Substitution

- Shift plaintext chars. three characters

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

- Example:
  - P = "*Old School cracked me up*"
  - C = ROG VFKRRO FUDFNHG PH XS

Another Substitution

- Shift plaintext chars. three characters

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

notice wrap

- Example:
  - P = "*Old School cracked me up*"
  - C = ROG VFKRRO FUDFNHG PH XS

Another Substitution

- Shift plaintext chars. three characters

A:	A	B	C	D	E	F	G	H	I	J	K	L	M
C:	D	E	F	G	H	I	J	K	L	M	N	O	P

A:	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
C:	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

- Algorithm called *Caesar Cipher*

notice wrap

Caesar Example

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
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

- What is: VFUXEV LV D IXQQB VKRZ ?

Caesar Cipher (more formal def)

- encryption:
  - $E_K(m) = m + 3 \bmod 26$
- decryption:
  - $D_K(c) = c - 3 \bmod 26$
- review:
  - if *a* and *m* are positive integers, *a* mod *m* is the remainder when *a* is divided by *m*
- Caesar cipher special case of **shift cipher**

Shift Cipher

- encryption:
  - $E_K(m) = m + K \bmod 26$
- decryption:
  - $D_K(c) = c - K \bmod 26$
- example: *k=5*

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
C: 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

- "summer vacation was too short" encrypts to
  - XZRR JWAF HFYN TSBF XYTT XMTW Y

Breaking Shift Ciphers

- How difficult?
- How many possibilities?
- Example:
  - AKZC JAQA IZMI TTGN CVVG APWE

First 13 Possibilities

0	A	K	Z	C	J	A	Q	A	I	Z	M	I	T	T	G	N	C	V	V	G	A	P	W
1	B	L	A	D	K	B	R	B	J	A	N	J	U	U	H	O	D	W	W	H	B	Q	X
2	C	M	B	E	L	C	S	C	K	B	O	K	V	V	I	P	E	X	X	I	C	R	Y
3	D	N	C	F	M	D	T	D	L	C	P	L	W	W	J	Q	F	Y	Y	J	D	S	Z
4	E	O	D	G	N	E	U	E	M	D	Q	M	X	X	K	R	G	Z	Z	K	E	T	A
5	F	P	E	H	O	F	V	F	N	E	R	N	Y	Y	L	S	H	A	A	L	F	U	B
6	G	Q	F	I	P	G	W	G	O	F	S	O	Z	Z	M	T	I	B	B	M	G	V	C
7	H	R	G	J	Q	H	X	H	P	G	T	P	A	A	N	U	J	C	C	N	H	W	D
8	I	S	H	K	R	I	Y	I	Q	H	U	Q	B	B	O	V	K	D	D	O	I	X	E
9	J	T	I	L	S	J	Z	J	R	I	V	R	C	C	P	W	L	E	E	P	J	Y	F
10	K	U	J	M	T	K	A	K	S	J	W	S	D	D	Q	X	M	F	F	Q	K	Z	G
11	L	V	K	N	U	L	B	L	T	K	X	T	E	E	R	Y	N	G	G	R	L	A	H
12	M	W	L	O	V	M	C	M	U	L	Y	U	F	F	S	Z	O	H	H	S	M	B	I

Last 13 Possibilities

13	N	X	M	P	W	N	D	N	V	M	Z	V	G	G	T	A	P	I	I	T	N	C	J	R
14	O	Y	N	Q	X	O	E	O	W	N	A	W	H	H	U	B	Q	J	J	U	O	D	K	S
15	P	Z	O	R	Y	P	F	P	X	O	B	X	I	I	V	C	R	K	K	V	P	E	L	T
16	Q	A	P	S	Z	Q	G	Q	Y	P	C	Y	J	J	W	D	S	L	L	W	Q	F	M	U
17	R	B	Q	T	A	R	H	R	Z	Q	D	Z	K	K	X	E	T	M	M	X	R	G	N	V
18	S	C	R	U	B	S	I	S	A	R	E	A	L	L	Y	F	U	N	N	Y	S	H	O	W
19	T	D	S	V	C	T	J	T	B	S	F	B	M	M	Z	G	V	O	O	Z	T	I	P	X
20	U	E	T	W	D	U	K	U	C	T	G	C	N	N	A	H	W	P	P	A	U	J	Q	Y
21	V	F	U	X	E	V	L	V	D	U	H	D	O	O	B	I	X	Q	Q	B	V	K	R	Z
22	W	G	V	Y	F	W	M	W	E	V	I	E	P	P	C	J	Y	R	R	C	W	L	S	A
23	X	H	W	Z	G	X	N	X	F	W	J	F	Q	Q	D	K	Z	S	S	D	X	M	T	B
24	Y	I	X	A	H	Y	O	Y	G	X	K	G	R	R	E	L	A	T	T	E	Y	N	U	C
25	Z	J	Y	B	I	Z	P	Z	H	Y	L	H	S	S	F	M	B	U	U	F	Z	O	V	D

Monoalphabetic Ciphers

- simple substitutions, *e.g.*, shift, keyword mixed, newspapaer cryptogram ... are **monoalphabetic ciphers**
- how many possible substitution alphabets?
- can we try all permutations?
- how would you try to break them?

monoalphabetic – brute force

- how many possible substitution alphabets?
  - $26! \approx 4 * 10^{26}$
- can we try all permutations?
  - sure. have some time?
  - at 1 test/μsec, about 12 trillion years.
- how would you try to break them?
  - use what you know to reduce the possibilities

breaking substitutions

- how do you break the newspaper cryptogram?
  - look at common letters (E, T, O, A, N, ...)
  - single-letter words (*I*, and *A*)
  - two-letter words (*of, to, in,* ...)
  - three-letter words (*the, and*, ...)
  - double letters (*ll, ee, oo, tt, ff, rr, nn*, ...)
  - other tricks?

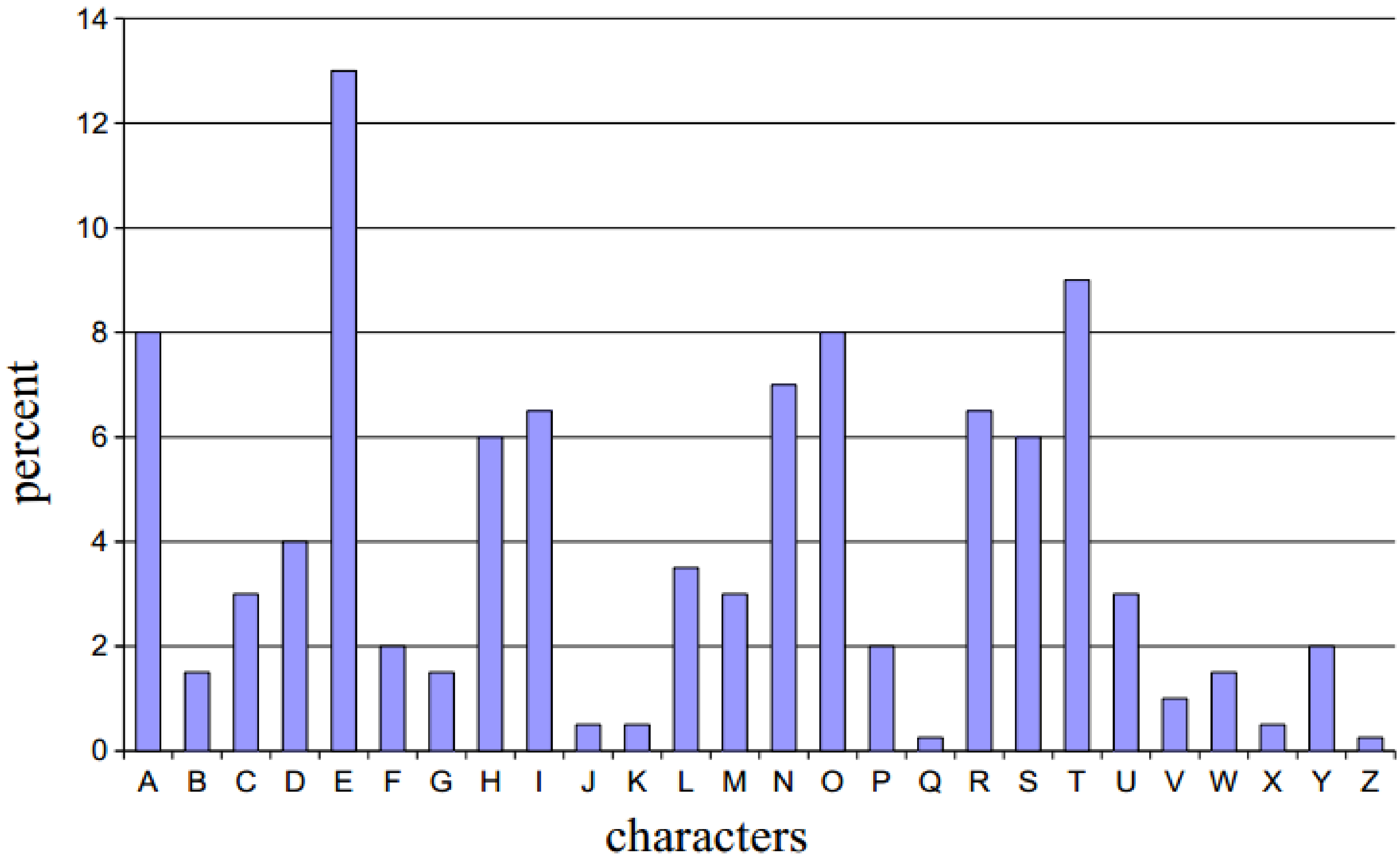
breaking substitutions (cont'd)

- use language statistics of plaintext
  - English, java, TCP packet headers, etc.
- example:
  - frequencies in English

char:	A	B	C	D	E	F	G	H	I	J	K	L	M
pct:	8	1.5	3	4	13	2	1.5	6	6.5	0.5	0.5	3.5	3

char:	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
pct:	7	8	2	0.25	6.5	6	9	3	1	1.5	0.5	2	0.25

Character Frequencies (English)



Common English Digrams and Trigrams

Digrams	Trigrams
EN	ENT
RE	ION
ER	AND
NT	ING
TH	IVE
ON	TIO
IN	FOR
TF	OUR
AN	THI
OR	ONE

End of Chapter2/ Part1