CRYPTOGRAPHIC ALGORITHMS AND PROTOCOLS

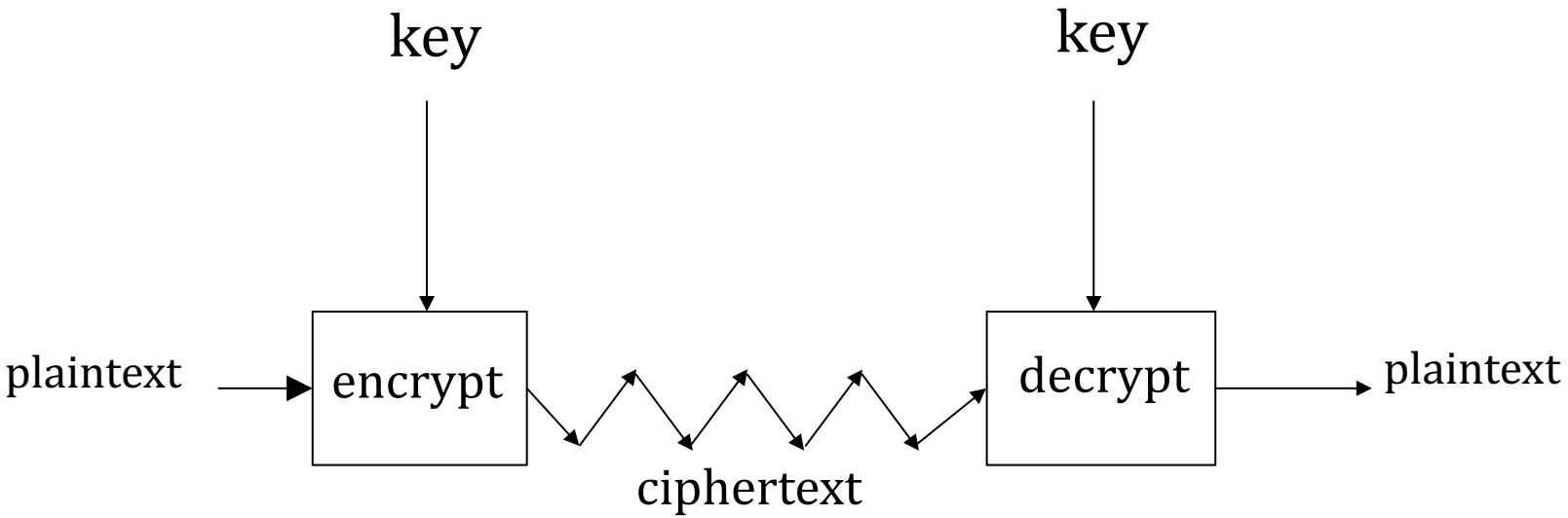
PART I: CRYPTOGRAPHY

2.1. Classical Ciphers

Crypto

- **Basic assumptions**
  - Only the key is secret
  - The system is completely known to the attacker
  - That is, crypto algorithms are not secret
- **This is known as Kerckhoffs' Principle**
- **Why do we make this assumption?**
  - Experience has shown that secret algorithms are weak when exposed
  - Secret algorithms never remain secret
  - Better to find weaknesses beforehand

Crypto as Black Box



A generic view of symmetric key crypto

Simple Substitution

- **Plaintext**: fourscoreandsevenyearsago
- **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
Ciphertext	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

- **Ciphertext**:
  IRXUVFRUHDQGVHYHQBHDUVDJR
- Shift by 3 is “Caesar’s cipher”

Ceasar's Cipher Decryption

- Suppose we know a Ceasar's cipher is being used:

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
Ciphertext	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

- Given ciphertext:
  VSRQJHEREVTXDUHSDQWV
- Plaintext:
  spongebobsquarepants

Not-so-Simple Substitution

- Shift by n for some $n \in \{0,1,2,\ldots,25\}$
- Then key is n
- Example: key n = 7

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
Ciphertext	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

Cryptanalysis I: Try Them All

- A simple substitution (shift by n) is used
  - But the key is unknown
- Given ciphertext: CSYEVIXIVQMREXIH
- How to find the key?
- Only 26 possible keys → try them all!
- **Exhaustive key search, or brute force attack**
- Solution: key is n = 4

Least-Simple Simple Substitution

- In general, simple substitution key can be any **permutation** of letters
  - Not necessarily a shift of the alphabet
- For example

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

- Then $26! > 2^{88}$ possible keys!

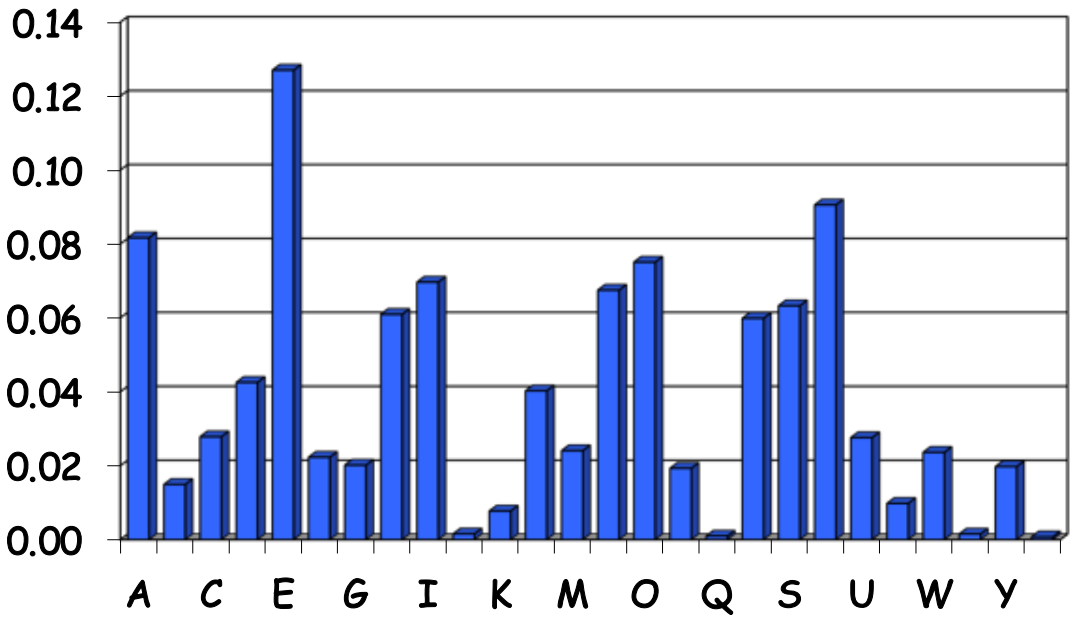
Cryptanalysis II: Be Clever

- We know that a simple substitution is used
- But not necessarily a shift by n
- Find the key given the ciphertext:

PBFPVYFBQXZTYFPBFEQJHDXXQVAPTPQJKTOYQWIPBVWLXT
OXBTFXQWAXBVCXQWAXFQJVWLEQNTOZQGGQLFXQWAKVW
LXQWAEBIPBFXFQVXGTVJVWLBTPQWAEBFPBFHCVLXBQUFE
VWLXGDPEQVPQGVPPBFTIXPFHXZHVFAGFOTHFEFBQUFTDH
ZBQPOTHXTYFTODXQHFTDPTOGHFQPBQWAQJJTODXQHFOQP
WTBDHHIXQVAPBFZQHCFWPFHPBFIPBQWKFABVYYDZBOTH
PBQPQJTQOTOGHFQAPBFEQJHDXXQVAVXEBQPEFZBVFOJIWF
FACFCCFHQWAUVWFLQHGFXVAFXQHFUFHILTTAVWAFFAWTE
VOITDHFHFQAITIXPFHXAFQHEFZQWGFLVWPTOFFA

Cryptanalysis II

- Cannot try all 2^{88} simple substitution keys
- Can we be more clever?
- English letter frequency counts…



Cryptanalysis II

- Ciphertext:

PBFPVYFBQXZTYFPBFEQJHDXXQVAPTPQJKTOYQWIPBVWLXTOXBTFXQW
AXBVCXQWAXFQJVWLEQNTOZQGGQLFXQWAKVWLXQWAEBIPBFXFQVXG
TVJVWLBTPQWAEBFPBFHCVLXBQUFEVWLXGDPEQVPQGVPPBFTIXPFHX
ZHVFAGFOTHFEFBQUFTDHZBQPOTHXTYFTODXQHFTDPTOGHFQPBQWA
QJJTODXQHFOQPWTBDHHIXQVAPBFZQHCFWPFHPBFIPBQWKFABVYYDZ
BOTHPBQPQJTQOTOGHFQAPBFEQJHDXXQVAVXEBQPEFZBVFOJIWFFACFC
CFHQWAUVWFLQHGFXVAFXQHFUFHILTTAVWAFFAWTEVOITDHFHFQAIT
IXPFHXAFQHEFZQWGFLVWPTOFFA

- Analyze this message using statistics below

Ciphertext frequency counts:

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
21	26	6	10	12	51	10	25	10	9	3	10	0	1	15	28	42	0	0	27	4	24	22	28	6	8

Cryptanalysis: Terminology

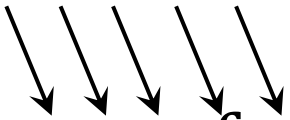
- Cryptosystem is **secure** if best know attack is to try all keys
  - Exhaustive key search, that is

- Cryptosystem is **insecure** if ***any*** shortcut attack is known

Transposition Ciphers

- now consider classical transposition or permutation ciphers
- these hide the message by rearranging the letter order without altering the actual letters used
- can recognise these since have the same frequency distribution as the original text

Rail Fence cipher

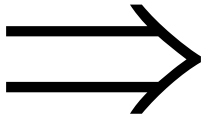
- Write message letters out diagonally over a number of rows
- Then read off cipher row by row
- eg. write message “meet me after the toga party” as:
  - m e m a t r h t g p r y
  - 
  - e t e f e t e o a a t
- Giving ciphertext
  - MEMATRHTGPRYETEFETEOAAT

Double Transposition

- **Plaintext**: *attackxatxdawn*

	col 1	col 2	col 3
row 1	a	t	t
row 2	a	c	k
row 3	x	a	t
row 4	x	d	a
row 5	w	n	x

Permute rows and columns



	col 1	col 3	col 2
row 3	x	t	a
row 5	w	x	n
row 1	a	t	t
row 4	x	a	d
row 2	a	k	c

- **Ciphertext**: *xtawxnattxadakc*
- Key is matrix size and permutations: (3,5,1,4,2) and (1,3,2)

One-Time Pad

- The cipher will be secure if a truly random key as long as the message, called One-Time pad, is used.
- Is unbreakable since ciphertext bears no statistical relationship to the plaintext
- Since for **any plaintext** & **any ciphertext** there exists a key mapping one to other
- Can only use the key **once** though
- Problems in generation & safe distribution of key

One-Time Pad: Encryption

e=000	h=001	i=010	k=011	l=100	r=101	s=110	t=111

Encryption: Plaintext ⊕ Key = Ciphertext

	h	e	i	l	h	i	t	l	e	r
Plaintext	001	000	010	100	001	010	111	100	000	101
Key	111	101	110	101	111	100	000	101	110	000
Ciphertext	110	101	100	001	110	110	111	001	110	101
	s	r	l	h	s	s	t	h	s	r

One-Time Pad: Decryption

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Decryption: Ciphertext ⊕ Key = Plaintext

	s	r	l	h	s	s	t	h	s	r
Ciphertext	110	101	100	001	110	110	111	001	110	101
Key	111	101	110	101	111	100	000	101	110	000
Plaintext	001	000	010	100	001	010	111	100	000	101
	h	e	i	l	h	i	t	l	e	r

One-Time Pad

Double agent claims sender used the following "**key**"

e=000	h=001	i=010	k=011	l=100	r=101	s=110	t=111

	s	r	l	h	s	s	t	h	s	r
Ciphertext	110	101	100	001	110	110	111	001	110	101
Key	101	111	000	101	111	100	000	101	110	000
Plaintext	011	010	100	100	001	010	111	100	000	101
	k	i	l	l	h	i	t	l	e	r

One-Time Pad

Or sender is captured and claims the key is…

e=000	h=001	i=010	k=011	l=100	r=101	s=110	t=111

	s	r	l	h	s	s	t	h	s	r
Ciphertext	110	101	100	001	110	110	111	001	110	101
Key	111	101	000	011	101	110	001	011	101	101
Plaintext	001	000	100	010	011	000	110	010	011	000
	h	e	l	i	K	e	s	i	k	e

One-Time Pad Summary

- **Provably** secure…
  - Ciphertext provides **no** info about plaintext
  - All plaintexts are equally likely
- …but, only when be used correctly
  - Pad must be random, used only once
  - Pad is known only to sender and receiver
- Note: pad (key) is same size as message
- So, why not distribute msg instead of pad?