CSC3631 Cryptography

Classical Cipher

Changyu Dong

Newcastle University

Classical Ciphers

- Ciphers being created and used in the old days
- Often operate on an alphabet of letters (such as "A-Z"), and are implemented by hand or with simple mechanical devices.
- No longer considered to be secure.
- But the basic ideas: substitution and transposition are still used in modern cryptography.

Shift Cipher

- Shift the letters *k* positions right in an alphabet
- Each letter in the original alphabet is mapped to a different letter in the shifted alphabet.
- *k* is the key
- Caesar's cipher: invented 2000 years ago, shift the alphabet by three positions.
- Standard Alphabet:
ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cryptographic Alphabet:
DEFGHIJKLMNOPQRSTUVWXYZABC
k=3
- "SECURITY" is encrypted as "VHFXULWB"

ROT13

- Commonly found on UNIX systems
- Every letter is rotated by 13 positions

Shift Cipher (mathematically)

- Let's map the English letters to integers, e.g. $A \rightarrow 0, B \rightarrow 1, ..., Z \rightarrow 25$
- The set $\{0, 1, 2, \ldots, 25\}$ can represent the plaintext space and ciphertext space.
- The key is a integer $0 \leq k \leq 25$ (how many positions being shifted)
  - $k = 3 \rightarrow$ Caesar cipher
  - $k = 13 \rightarrow$ ROT13
- The encryption function is addition modulo 26
  - $E_k(x) = x + k \bmod 26$
- The decryption function is subtraction modulo 26
  - $D_k(y) = y - k \bmod 26$

How Secure is a Shift Cipher

- Can an attacker find out the key?
  - Yes, at most 26 tries
  - Key space is too small
  - Exhaustive key search is way too easy.
- A secure cipher must have a key space that is not vulnerable to exhaustive search.

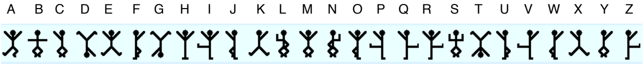
Monoalphabetic Substitution Cipher

- Encrypt messages using a permutation of the original alphabet or a different alphabet.
- Shift cipher is a special case of Monoalphabetic substitution cipher
- Example: permutation

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

  SECURITY → LTEXKOZN

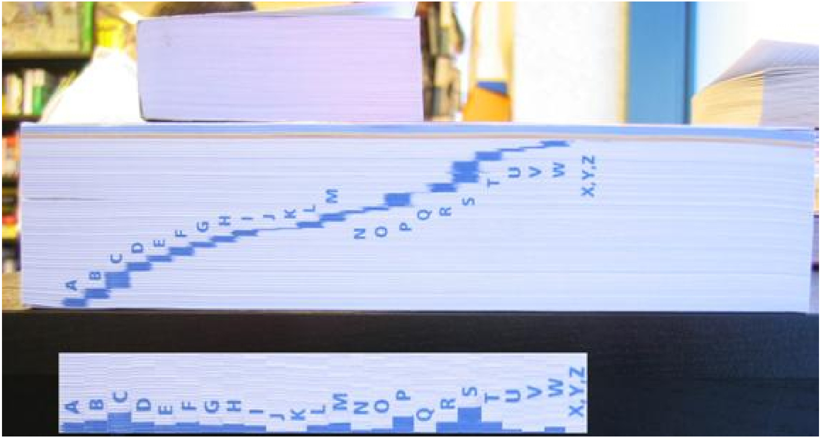
- Example: dancing men (Sherlock Holmes)



How Secure is a Monoalphabetic Substitution Cipher

- More difficult to bruteforce the key (permutation)
  - Key space size is $26! \approx 2^{88}$ (while DES only has a key space of size 2^{56})
  - Exhaustive search won't be successful
- However
  - Each language has its pattern, e.g. frequency of letters and combinations of letters (2-grams, 3-grams)
  - Substitution ciphers preserve this pattern
  - Therefore make it vulnerable to frequency analysis
  - Given enough ciphertext, the cipher can be cracked easily.

Frequency analysis

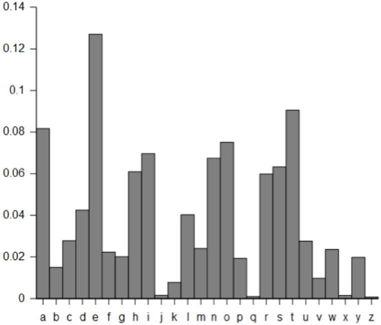


A photographic side view of the Oxford Advanced Learners' Dictionary

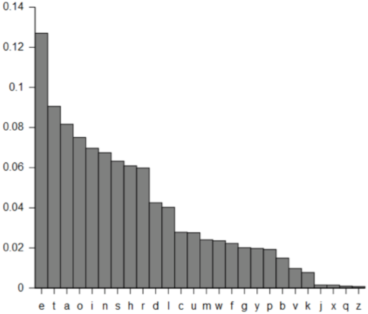
Frequency analysis

Letter	Probability	Letter	Probability
A	.082	N	.067
B	.015	O	.075
C	.028	P	.019
D	.043	Q	.001
E	.127	R	.060
F	.022	S	.063
G	.020	T	.091
H	.061	U	.028
I	.070	V	.010
J	.002	W	.023
K	.008	X	.001
L	.040	Y	.020
M	.024	Z	.001

Frequency analysis

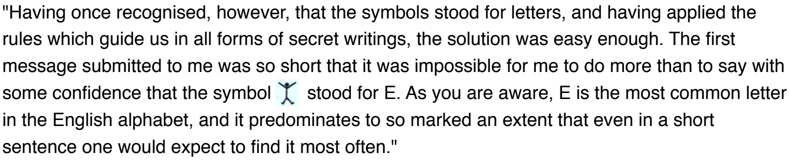


(a) Relative frequencies of English letters



(b) Relative frequencies sorted by frequency

Cryptanalysis of substitution cipher



> "Having once recognised, however, that the symbols stood for letters, and having applied the rules which guide us in all forms of secret writings, the solution was easy enough. The first message submitted to me was so short that it was impossible for me to do more than to say with some confidence that the symbol stood for E. As you are aware, E is the most common letter in the English alphabet, and it predominates to so marked an extent that even in a short sentence one would expect to find it most often."

- Large key space is not sufficient for a secure cipher
- Substitution only provides confusion – making the relationship between the ciphertext and the plaintext complex,
- but not diffusion – dissipating the statistical structure of plaintext over the bulk of ciphertext
  - Each plaintext letter is mapped to one ciphertext letter, and the mapping is deterministic (fixed).
- A secure cipher would need both

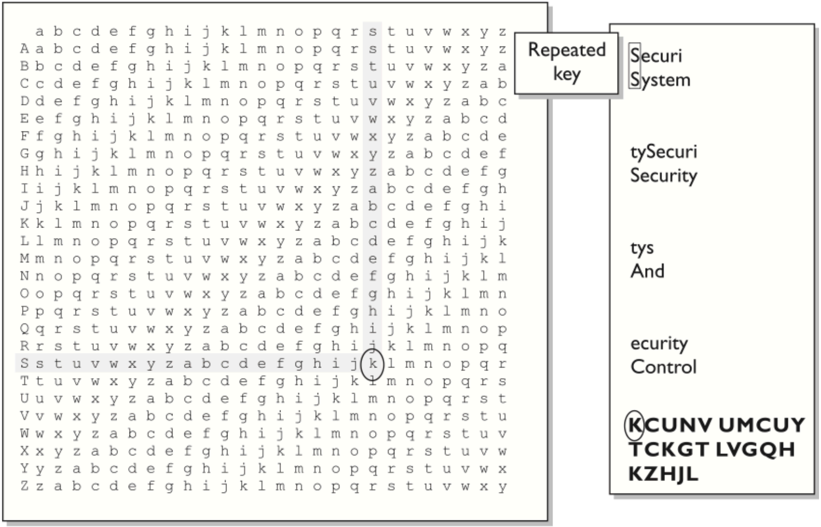
Polyalphabetic Substitution Ciphers

- Use more than one cipher alphabet, and switch between them when encrypting different letters
- Makes frequencies of letters more uniform (diffusion)

The Vigenère Cipher

- Uses multiple shift ciphers
- Key is a short sequence of letters
- Encryption is done by align the plaintext and the key (repeat the key if it is shorter than the plaintext), $c_i = p_i + k_i \bmod 26$
- Decryption is done by align the ciphertext and the key (repeat the key if needed), then $p_i = c_i - k_i \bmod 26$.
- Here p_i, k_i, c_i are the ith letter in the plaintext, key stream, ciphertext respectively.

The Vigenère Cipher



	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

Repeated key

Securi
System

tySecuri
Security

tys
And

ecurity
Control

KCUNV UMCUY
TCKGT LVGQH
KZHJL

key: security, plaintext: systemsecurityandcontrol

How Secure is the Vigenère Cipher

- One letter in ciphertext corresponds to multiple letters in plaintext
  - In the example, “C” in the ciphertext corresponds to “Y”, “E” and “I”
  - Frequency analysis is more difficult.
- The number of the shift ciphers being used depends on the key

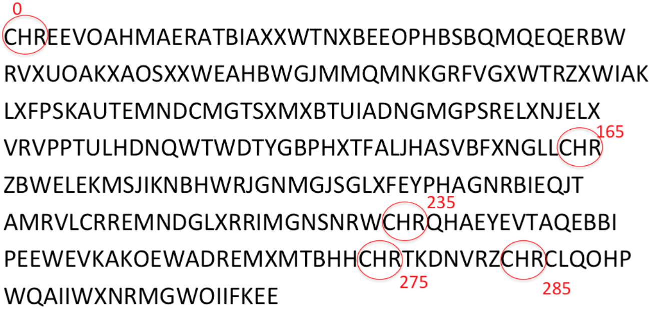
Break The Vigenère Cipher

- Find the length of the key: two identical segments of plaintext will be encrypted to the same ciphertext, if the the distance between them is a multiple of the key length.

Key	K	I	N	G	K	I	N	G	K	I	N	G	K	I	N	G	K	I	N	G	K	I	N	G
PT	t	h	e	s	u	n	a	n	d	t	h	e	m	a	n	i	n	t	h	e	m	o	o	n
CT	D	P	R	Y	E	V	N	T	N	B	U	K	W	I	A	O	X	B	U	K	W	W	B	T

- Then divide the message into many shift cipher encryptions.
- Use frequency analysis to solve the resulting shift ciphers.
- The segment needs to be long enough 3 or more characters.

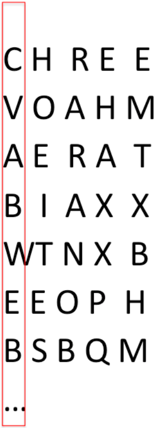
Break The Vigenère Cipher


0
165
235
275
285

CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBW
RVXUOAKXAOSXXWEAHBWGJMMQMNKGRFVGXWTRZXWIAK
LXFPSKAUTEMNDCMGTSXMXBTUIADNGMGPSRELXNJELX
VRVPPTULHDNQWTWDTYGBPHXTFALJHASVBFXNGLLCHR
ZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJT
AMRVLCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBI
PEEWEVKAKOEWADREMXMTBHHCHRTKDNVRZCHRCLQOHP
WQAIIWXNRMGWOIIFKEE

- The shortest distance between CHRs is 285 -275 = 10. Is 10 the key length?

Break The Vigenère Cipher



C	H	R	E	E
V	O	A	H	M
A	E	R	A	T
B	I	A	X	X
W	T	N	X	B
E	E	O	P	H
B	S	B	Q	M
…				

Transposition Ciphers

- Change the position of the letters in the plaintext, without changing the actual letters
- The ciphertext is a permutation of the plaintext.
- So they are also called permutation ciphers

Columnar Transposition

- Plaintext is written row by row in a matrix
- Encryption: write out the columns in an order specified by a key.

Plaintext: attack postponed until two am

Key: 3421567

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

Break Columnar

- Guess the key length: should divide the length of the ciphertext
- Write message out in columns
- Reordering the columns to see whether gives common bigrams (th, he, ...) or N-grams

Kerckhoffs' Principle

- The security of a cryptographic system must depends only only the secrecy of the key, and not on the secrecy of the algorithm.
    - Algorithm is hard to change
    - It is far more difficult to keep the algorithm secret than keep a key secret
    - Public algorithms can be analysed thoroughly to find bugs and flaws, while secret algorithms cannot.
    - Keep the algorithm secret won't add another layer of security, it actually weaken the security.

Reading

- For this lecture:
  - Cryptography made simple § 7 (relevant sections)
  - Cryptography theory and practice: § 1 (relevant sections)
  - Applied cryptography § 1.3