Symmetric Encryption

- 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
- most widely used

Some Basic Terminology

- **plaintext** - original message
- **ciphertext** - coded message
- **cipher** - algorithm for transforming plaintext to ciphertext
- **key** - info used in cipher known only to sender/receiver
- **encipher (encrypt)** - converting plaintext to ciphertext
- **decipher (decrypt)** - recovering ciphertext from 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

Cryptography

- characterize cryptographic system by:
  - type of encryption operations used
    - substitution / transposition / product
  - number of keys used
    - single-key or private / two-key or public
  - way in which plaintext is processed
    - block / stream

Cryptanalysis

- objective to recover key not just message
- general approaches:
  - cryptanalytic attack
  - brute-force attack

Cryptanalytic Attacks

- **Cipher text only**
  - only know algorithm & cipher text, is statistical, know or can identify plaintext
- **known plaintext**
  - know/suspect plaintext & cipher text
- **chosen plaintext**
  - select plaintext and obtain cipher text
- **chosen cipher text**
  - select cipher text and obtain plaintext
- **chosen text**
  - select plaintext or cipher text to en/decrypt

Brute Force Search

- always possible to simply try every key
- most basic attack, proportional to key size
- assume either know / recognise plaintext

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 $= 5.4 \times 10^{24}$ years	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	2^{167} μs $= 5.9 \times 10^{36}$ years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	2×10^{26} μs $= 6.4 \times 10^{12}$ years	6.4×10^6 years

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 cipher text bit patterns

Caesar Cipher

- earliest known substitution cipher
- by Julius Caesar
- first attested use in military affairs
- replaces each letter by 3rd letter on
- example:

  meet me after the class

  PHHW PH DIWHU WKH FODVV

Caesar Cipher

- can 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 have Caesar cipher as:

  $c = \mathrm{E}(p) = (p + k) \bmod (26)$

  $p = \mathrm{D}(\mathrm{c}) = (\mathrm{c} - k) \bmod (26)$

Cryptanalysis of Caesar Cipher

- only have 26 possible ciphers
  - A maps to A,B,..Z
- could simply try each in turn
- a **brute force search**
- given cipher text, just try all shifts of letters
- do need to recognize when have plaintext
- eg. break cipher text "GCUA VQ DTGCM"

Monoalphabetic Cipher

- rather than just shifting the alphabet
- could shuffle (jumble) the letters arbitrarily
- each plaintext letter maps to a different random cipher text letter
- hence key is 26 letters long

Plain: abcdefghijklmnopqrstuvwxyz

Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

Plaintext: ifwewishtoreplaceletters

Cipher text: WIRFRWAJUHYFTSDVFSFUUFYA

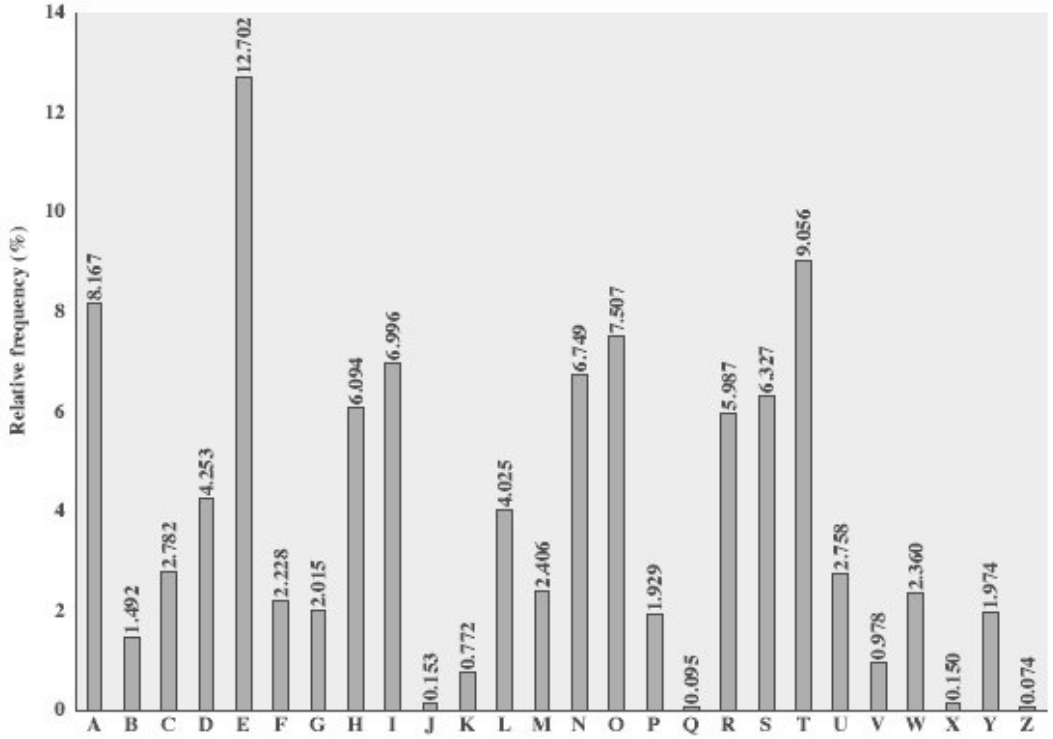
Monoalphabetic Cipher Security

- now have a total of 26! = 4 x 1026 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**
- eg "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
- discovered by Arabian scientists in 9th century
- 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

Example Cryptanalysis

- given ciphertext:

  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
- 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

Playfair Cipher

- not even the large number of keys in a monoalphabetic cipher provides security
- one approach to improving security was to encrypt multiple letters
- the **Playfair Cipher** is an example
- invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair

Play fair Key Matrix

- a 5X5 matrix of letters based on a keyword
- fill in letters of keyword (sans duplicates)
- fill rest of matrix with other letters
- eg. using the keyword MONARCHY

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

Encrypting and Decrypting

- plaintext is encrypted two letters at a time

1. if a pair is a repeated letter, insert filler like 'X'
2. if both letters fall in the same row, replace each with letter to right (wrapping back to start from end)
3. if both letters fall in the same column, replace each with the letter below it (again wrapping to top from bottom)
4. otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair

Security of Play fair Cipher

- security much improved over mono alphabetic
- since have 26 x 26 = 676 diagrams
- would need a 676 entry frequency table to analyse (verses 26 for a mono alphabetic)
- and correspondingly more cipher text
- was widely used for many years
  - eg. by US & British military in WW1
- it **can** be broken, given a few hundred letters
- since still has much of plaintext structure

Hill Cipher:

- Hill cipher was developed by the mathematician Lester Hill in 1929.
- The encryption algorithm takes m successive plaintext letters and substitutes for them m cipher text letters.
- The substitution is determined by m linear equations in which each character is assigned a numerical value (a = 0, b = 1 ... z = 25).

Hill Cipher:

- For m = 3, the system can be described as follows
- $c1 = (k_{11P1} + k_{12P2} + k_{13P3}) \bmod 26$ $c2 = (k_{21P1} + k_{22P2} + k_{23P3}) \bmod 26$ $c3 = (k_{31P1} + k_{32P2} + k_{33P3}) \bmod 26$

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \bmod 26$$

Or

$$C = KP \bmod 26$$

Polyalphabetic Ciphers

- **polyalphabetic substitution ciphers**
- improve security using multiple cipher alphabets
- make cryptanalysis harder with more alphabets to guess and flatter frequency distribution
- use a key to select which alphabet is used for each letter of the message
- use each alphabet in turn
- repeat from start after end of key is reached

Vigenère Cipher

- simplest polyalphabetic substitution cipher
- effectively multiple caesar ciphers
- key is multiple letters long $K = k_1\ k_2\ \ldots\ k_d$
- i^{th} letter specifies i^{th} alphabet to use
- use each alphabet in turn
- repeat from start after d letters in message
- decryption simply works in reverse

Example of Vigenère Cipher

- write the plaintext out
- write the keyword repeated above it
- use each key letter as a caesar cipher key
- encrypt the corresponding plaintext letter
- eg using keyword *deceptive*

key: deceptive deceptive deceptive

plaintext: we are discovered save yourself

ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ

Security of Vigenère Ciphers

- have multiple cipher text letters for each plaintext letter
- hence letter frequencies are obscured
- but not totally lost
- start with letter frequencies
  - see if look mono alphabetic or not
- if not, then need to determine number of alphabets, since then can attach each

One-Time Pad

- if a truly random key as long as the message is used, the cipher will be secure
- called a One-Time pad
- 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

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 out as: meet me after the toga party

  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

Columnar Transposition Ciphers

- a more complex transposition
- write letters of message out in rows over a specified number of columns
- then reorder the columns according to some key before reading off the rows

Key: 3 4 2 1 5 6 7

Plaintext: 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

Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ

Summary

- classical cipher techniques and terminology
- Mono alphabetic substitution ciphers
- cryptanalysis using letter frequencies
- Play fair cipher
- Hill cipher
- polyalphabetic ciphers
- transposition ciphers