Cryptography

The Making and Breaking of Secret Codes.

Need for Cryptography

- Many areas of human endeavor require secret communication.
- Modern methods of communication more open and subject to interception.
  - Telegraph, radio, internet.
- The use is rapidly increasing.
- Electronic commerce requires it.

Codes & Ciphers

- Convenience codes.
  - Publicly known - no secrecy involved.
  - Morse code - telegram & radio.
  - ASCII code - computer.
  - Zip, area, …
- Secret codes or ciphers.
  - Today's topic.

Summary

- Four codes used over time -- and how to break them.
  - Substitution ciphers.
    - Caesar’s cipher.
    - Monoalphabetic ciphers.
    - Polyalphabetic ciphers.
  - Computer enabled ciphers.
    - Public key ciphers.

Caesar’s Cipher

- Used in the gallic wars
    - Documented by Suetonius in *Lives of the twelve Caesar's*
  - **ABCDEFGHIJKLMNOPQRSTUVWXYZ**
  - **DEFGHIJKLMNOPQRSTUVWXYZABC**
  - Help me → KHOS PH
- Algorithm -- simple shift
- Key -- number, the amount of shift

Breaking the Cipher

- Find the key -- there are 26 possibilities. We can check them one by one until one makes sense.
  - If we know a simple shift code is being used.

Monoalphabetic Ciphers

- **Example:**
  - **ABCDEFGHIJKLMNOPQRSTUVWXYZ**
  - **QAZWSXEDCRFVTGBYHNUJMIKOLP**
  - Help me → DSVY TS
- Algorithm -- permutation of the alphabet
  - There are 26! -- 4 X 10^{26} possibilities

Key

- Must be enough information to easily construct the permutation
- Key word -- Rice University
  - **ABCDEFGHIJKLMNOPQRSTUVWXYZ**
  - **RICEUNVSTYZABDFGHJKLMOPQWX**
  - Help me → SUAG BU

Breaking the Cipher

- Frequency analysis
  - Mathematics
- Word and language patterns
  - Linguistics
  - Puzzlers
- Persistence
  - *The Gold Bug* -- Edgar Allan Poe

Alphabet Frequency (%)

A	8.2	J	0.2	S	6.3
B	1.5	K	0.8	T	9.1
C	2.8	L	4.0	U	2.8
D	4.3	M	2.4	V	1.0
E	12.7	N	6.7	W	2.4
F	2.2	O	7.5	X	0.2
G	2.0	P	1.9	Y	2.0
H	6.1	Q	0.1	Z	0.1
I	7.0	R	6.0		

Breaking the Cipher (Cont.)

- Frequency analysis invented by middle eastern, Arabian mathematicians in 9th century.
- By the end of the 14th century "anyone" could easily break monoalphabetic ciphers.

Polyalphabetic Ciphers

- Leon Battista Alberti - 1460
  - Use two or more cipher alphabets & alternate them
    - **ABCDEFGHIJKLMNOPQRSTUVWXYZ**
    - **QAZWSXEDCRFVTGBYHNUJMIKOLP**
    - **POLIKUJMNHYTGBVFREDCXSWZAQ**
    - Help me → DKVF TK
  - 1.6 x 10^{53} combinations

Blaise de Vigenere - 1560

- Introduced a convenient keyword
  - Made the use of the algorithm easier
- Use 26 cipher alphabets
  - **`ABCDEFGHIJKLMNOPQRSTUVWXYZ`**
  - **`BCDEFGHIJKLMNOPQRSTUVWXYZA`**
  - **`CDEFGHIJKLMNOPQRSTUVWXYZAB`**
  - **`DEFGHIJKLMNOPQRSTUVWXYZABC`**
  - **`EFGHIJKLMNOPQRSTUVWXYZABCD`**
  - **`etc`**

Keyword BOZ

- ABCDEFGHIJKLMNOPQRSTUVWXYZ
- BCDEFGHIJKLMNOPQRSTUVWXYZA
- OPQRSTUVWXYZABCDEFGHIJKLMN
- ZABCDEFGHIJKLMNOPQRSTUVWXY
- Help me → ISKQ AD
- THE → UVD, HGF, or SIS

Use of the Vigenere Cipher

- Ignored for about 200 years
- Invention of telegraph made codes more important
  - Messages easy to intercept
  - Greatly increased volume of traffic
- Became known as *le chiffre indechiffrable*

Breaking the Vigenere Cipher

- Charles Babbage
  - Invented an early mechanical computer
  - C. 1854 broke the Vigenere code
  - Did not publish the result
- Frederich Wilhem Kasiski (Prussian)
  - 1863 published the way to break the code

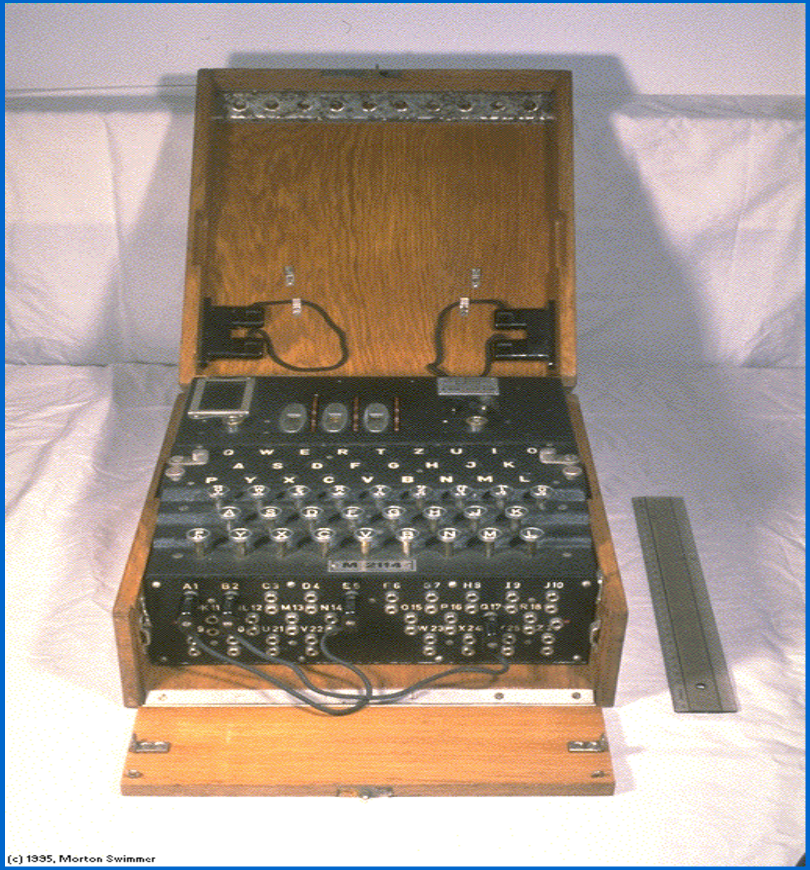
Breaking the Cipher (cont.)

- Weak point is the keyword
  - Look for repeating patterns in the cipher
    - Using BOZ, THE → UVD, HGF, or SIS
      - How far apart are appearances of same pattern?
    - Allows determination of the length of the keyword
      - Determines the number of cipher alphabets used
- Frequency analysis on each cipher alphabet
- Requires a lot of traffic

20^{th} Century Ciphers

- Radio (Marconi ~ 1900)
  - Greater ease of communication
  - Greater ease of interception
- Electro-mechanical devices
  - Encryption and decryption can be semi-automated
  - Polyalphabetic ciphers with more alphabets

The Enigma Machine



- Invented in 1918 by Arthur Scherbius and Richard Ritter
- Keyboard
- 3 rotors or scramblers
- Reflector
- Output lights
- Plug wiring

The Enigma Machine (Cont.)

- The use of the rotors and reflector causes it to rotate through a cycle of about 17,000 cipher alphabets.
- Plug wiring changes the cycle.
- Starting position determines which cycle and where in the cycle the message starts.
- Over 10^{16} different starting positions.

Key

- Determines the starting position
- Two keys used
  - Daily key used only to encrypt a message key
  - Message key unique to each message

Importance in World War 2

- All countries had similar machines
  - Many were confident it was unbreakable
- Poland & England broke enigma
- USA broke Japanese codes
- One of the keys to Allied victory in WW2
- Battle of the Atlantic
- Battle of Midway

Cracking Enigma (Poland)

- Polish mathematicians in 1930's
  - Espionage by the French
  - Marian Rejewski
  - Broke Enigma by 1934
    - Noticed patterns in the day key
  - Germans improved the Enigma
  - Gave everything to the Allies 2 weeks before the invasion of Poland

Methods

- Look for mathematical patterns
- Exploit the known structure of the machine
- Exploit defective practices
  - Use of daily key to encrypt repeated message key

Cracking Enigma (England)

- Bletchley Park, Alan Turing & ULTRA
  - Continued with the Polish plan
  - Cillies --- obvious message keys
  - Cribs --- routine messages
  - Bombes --- sets of enigma machines
  - Espionage --- find the code books

Advances in Enigma

- Number of rotors increased to 5 or 6
  - Greatly increased the length of the cycle
- Complexity of the plug wiring increased
  - Increased the number of available cycles
- Elimination of cillies --- use of randomly generated message keys

Computers and Ciphers

- The ASCII code turns messages into numbers:

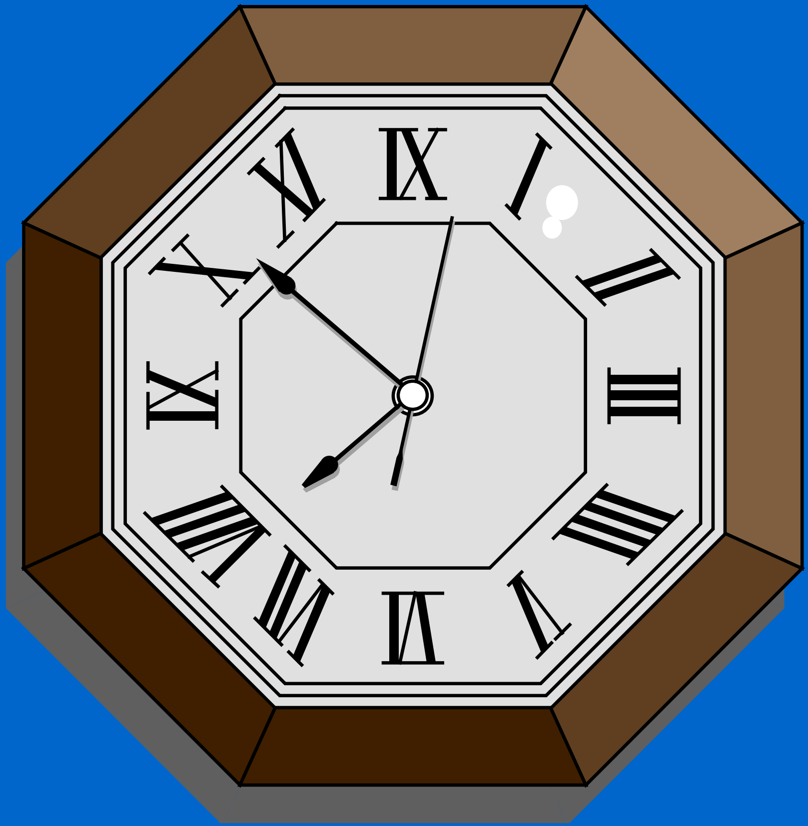
H	e	l	p	!
1001000	1100101	1101100	1110000	0100001

- Help! -->10010001100101110110011100000100001
- = 19,540,949,025
- ASCII code is the computer’s alphabet
- A cipher can be any function that is 1-1

Modular Arithmetic

- A mod(N) = remainder when A is divided by N
- Example:
- `1 mod(3) = 1, 5 mod(3) = 2,  9 mod(3) = 0`
- `2 mod(3) = 2, 6 mod(3) = 0, 10 mod(3) = 1`
- `3 mod(3) = 0, 7 mod(3) = 1, 11 mod(3) = 2`
- `4 mod(3) = 1, 8 mod(3) = 2, 12 mod(3) = 0`

Clock Arithmetic



- The clock uses arithmetic mod(12) to measure hours
- It uses arithmetic mod(60) to measure minutes and seconds

Cipher With Modular Arithmetic

Encryption Algorithm		
M	M^3	M^3 mod(11)
1	1	1
2	8	8
3	27	5
4	64	9
5	125	4
6	216	7
7	343	2
8	512	6
9	729	3
10	1000	10

Decryption Algorithm		
C	C^7	C^7 mod(11)
1	1	1
2	128	7
3	2187	9
4	16384	5
5	78125	3
6	279936	8
7	823543	6
8	2097152	2
9	4782969	4
10	10000000	10

Data Encryption Standard (DES)

- Originally called Lucifer
  - Invented at IBM by Horst Feistal
  - Adopted by US government in 1975
- There are 2^{56} ($\sim 10^{17}$)possible secret keys
  - Called a 56 bit cipher
- Now out of date
  - Advanced Encryption Standard adopted in 2001

Public Key Codes

- Encryption algorithm and key are public information
  - Anyone can use it to communicate with the holder of the decryption algorithm
  - This eliminates the need to secretly convey the key
- Decryption key is not public, and is very hard to discover

The RSA Code

– Ronald Rivest, Adi Shamir & Leonard Adelman -- 1977

- 2 very large primes P & Q (private)
- N = P x Q & number E (public)
- Message M (a number)
- Encrypt the message with the formula
- $C = M^E \bmod(N)$

Decryption in RSA

- Decrypter knows a secret number D with
- $E \times D \mod((P-1)\times(Q-1)) = 1$
- $C^D \mod (N) = (M^E)^D \mod(N)$
- $= M^{ED} \mod(N)$
- $= M$ (Theorem of Euler)

Example

- Take P = 89,833 and Q = 945,677 (private)
- N = PxQ = 84,953,001,941 (public)
- E = 1,080,461 (public)
- Help! → 19,540,949,025 = M
- C = $19{,}540{,}949{,}025^{1{,}080{,}461}$ mod(N)
- = 6,499,326,013

Example (Cont.)

- To decode use D = 235,877 (private)
- C = 6,499,326,013
- C^D mod(N) = 19,540,949,025
- = M
- → Help!

Breaking RSA (Brute Force)

- Need to find the integer D
- Try all possibilities one by one
- If P & Q are large, there are simply too many choices for D. Say about 10^{200}

Breaking RSA (Factoring)

- The best way is to factor N (= P x Q)
  - In practice both P & Q have 100 to 200 digits
    - The code can be made more secure by choosing larger primes
  - N has as many as 400 digits
  - Knowing P, Q & E, it is easy to find D
- Factoring large numbers is an extremely difficult problem

Example

- 1977 Martin Gardner posed a challenge
  - Factor a number with 129 digits, and use it to decode a message
  - Many participants
- Done in 1994 by a team of 600 volunteers
- Modern RSA uses Ns with over 200 digits

Pretty Good Privacy (PGP)

- Phil Zimmermann --- 1991
  - Encryption for the masses
  - Uses a standard encryption method (like DES) for the message
  - Uses RSA only to encrypt the key
- Conflict with US government
  - Eventually everything was settled in favor of personal privacy

Advanced Encryption Algorithm

- By mid 1990s DES was clearly out of date
- 1997 NIST announced an open competition
  - Many competitors from around the world
  - 15 semi-finalists --- NIST asked for comments
  - 1999 5 finalists
  - Oct. 2000 Rijndael declared the best
  - Nov. 2001 Rijndael adopted as the AES

Rijndael

- Invented by Joan Daemen and Vincent Rijmen.
- Operates on 128 bit blocks
- Uses modular arithmetic and several polynomial mappings
- Has a 128 bit key
  - Or 192 or 256
- Won on the basis of security, performance, efficiency, implementability, and flexibility

The future

- Quantum computing
  - New algorithms for factoring numbers very quickly
  - There are at present no quantum computers
  - Area of intense research
- The invention of new algorithms for solving equations is always possible

National Security Agency (NSA)

- America's Black Chamber
- Largest employer of mathematicians in the world
- Once ultra secret, it is becoming more and more open
- They even run a museum

Bibliography

- *The Code Book*, by Simon Singh, New York: Doubleday, 1999
- *The Codebreakers*, by David Kahn, New York: Scribners, 1996 & 1999
- *Cryptography*, by Lawrence Dwight Smith, New York: Dover
- *Alan Turing: The Enigma*, by David Hodges, London: Vintage, 1992

Web Sites

- The Enigma Machine
  - http://www.math.arizona.edu/~dsl/enigma.htm
- Bletchley Park
  - http://www.cranfield.ac.uk/ccc/bpark/
- RSA Security's Frequently Asked Questions
  - http://www.rsasecurity.com/rsalabs/
- The National Security Agency
  - http://www.nsa.gov/