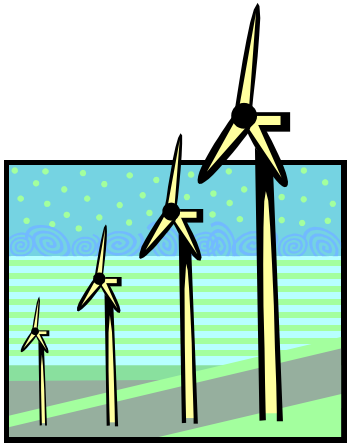
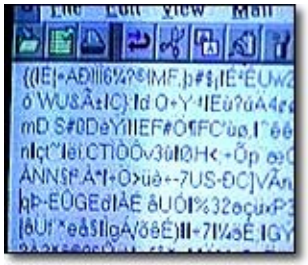
CSCD 303
Fall 2017



Lecture 19
Cryptography - Basics

Cryptography

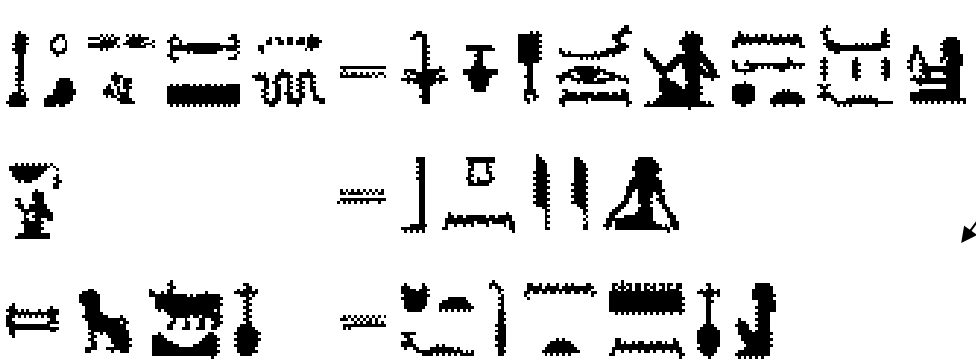


- Introduction
  - **Cryptography** is a science
- For exchanging information between intended recipients
  - Has been used for over 4000 years

  - What is the main use of Crytography today?
  - Cryptography used for
    - Data integrity
    - Insure privacy
    - Identify users

Cryptography History

- Cryptography has long history dates from Egyptians about 4000 years ago
- Ancient Egyptians enciphered some of their hieroglyphic writing on monuments



Egyptian Crypto

Hieroglyphic encipherments of proper names and titles, with cipher hieroglyphs at left, plain equivalents at right

- Ancient Hebrews enciphered certain words in their scriptures
- One of the most famous uses comes from Roman times
- More on this later ....

Cryptography History

Blaise de Vignere

- Continued ...

- 1300's, Geoffrey Chaucer included several ciphers in his works ... father of English Literature, famous poet
- 1460's, Leon Alberti devised a cipher wheel, and described the principles of frequency analysis
- 1585, Blaise de Vigenère published a book on cryptology and described polyalphabetic substitution cipher
  - This cipher is used to this day ...

Cryptography Background

- A complete non-technical account of cryptography from its beginning through early 1960's is

  D. Kahn, The Codebreakers, Macmillan Publishing Company, 1976.

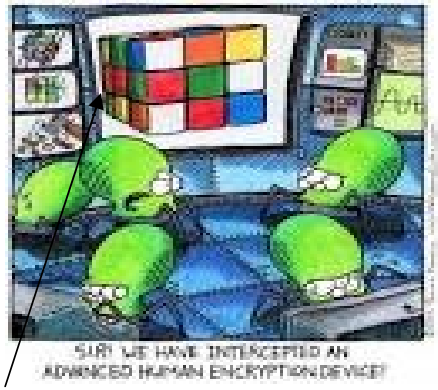
  – Relates historical aspects which were most significant to development of modern cryptography

  Simon Singh, The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography, Anchor, 2011

- Summary of important developments in 1970's and their relation to cryptography today see:

  A. Menezes, P. van Oorschot, and S. Vanstone, Handbook of Applied Cryptography, CRC Press, 1997
  http://www.cacr.math.uwaterloo.ca/hac/

Cryptography Definitions



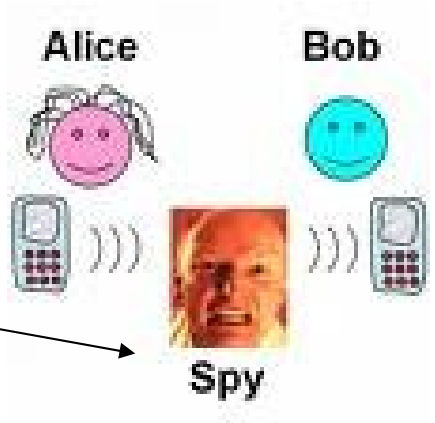
Sir, we have intercepted an advanced human encryption device

- Terms
  - Encryption (E)
    - Process of encoding message so that its meaning is not obvious
  - Decryption (D)
    - Reversal process transform message back to original form
  - Plaintext (P)
    - Original message
  - Ciphertext (C)
    - Encrypted form of original message

Cryptography Definitions

- Terms
  - Cryptanalyst
    - Studies encryption and encrypted messages
    - Works for unauthorized interceptor
  - Cryptographer
    - Works on behalf of a legitimate sender or receiver

Cryptography guards against what security problems?



Cryptography Definitions

- Formal Notation

  $C = E(P)$ and $P = D(C)$

  where C represents ciphertext

  E is encryption rule

  D is decryption rule

  Cryptosystem is where

  $P = D(E(P))$

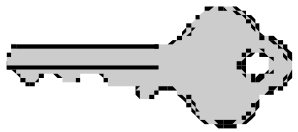
  where P represents Plaintext

  want to convert a message but also want to be able to get it back again

Cryptography Concepts

- In cryptosystems idea of key is extremely important
  - A key is used to both encrypt and decrypt messages
  - May be different keys depending upon the crypto algorithm
- Key length is also important in determining a crypto system's strength

Cryptosystem



- A Cryptosystem is set of rules for how to encrypt plaintext and how to decrypt ciphertext
- Process is similar to using mass produced house locks ... An aside ... Defcon looks at lock insecurity

  http://www.thesidebar.org/insecurity/?p=515

  – Few well-known companies produce standard locks that differ according to physical key
  – You and neighbour have same lock model
  – But your key will only open your lock
- So … have a few well-examined encryption algorithms that everyone uses
  – People using algorithm have different keys

Cryptosystem

- How Secure is it?
  - Unconditionally secure if
    - Can't be broken even with infinite resources
    - Implies cryptanalysis is impossible even if every key were tried
    - Still couldn't determine the correct key

Cryptosystem

- How Secure is it?
  - Computationally secure if
    - Possible to break encryption but practically infeasible to do so
    - Time and resources are > value of encrypted info
    - Measured in 10's or 1000's of years
    - Most modern crypto systems are in that category

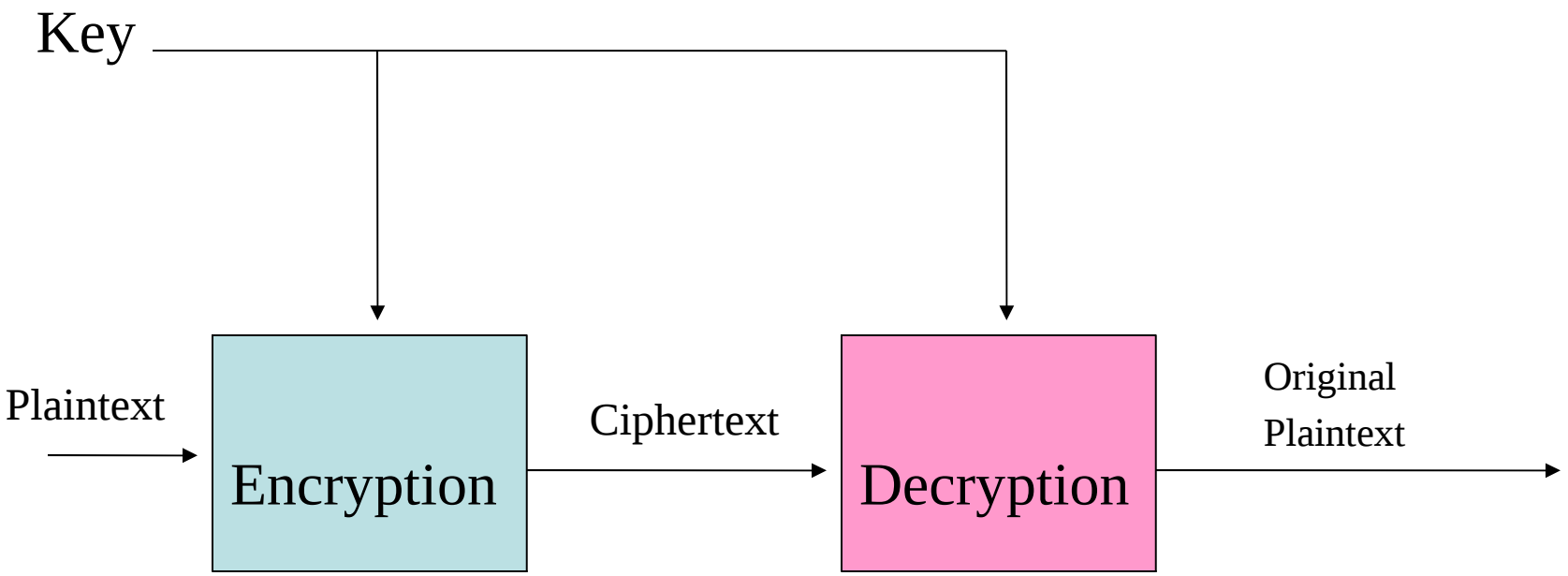
Definitions

- What is Symmetric Encryption and how does it differ from Asymmetric Encryption?

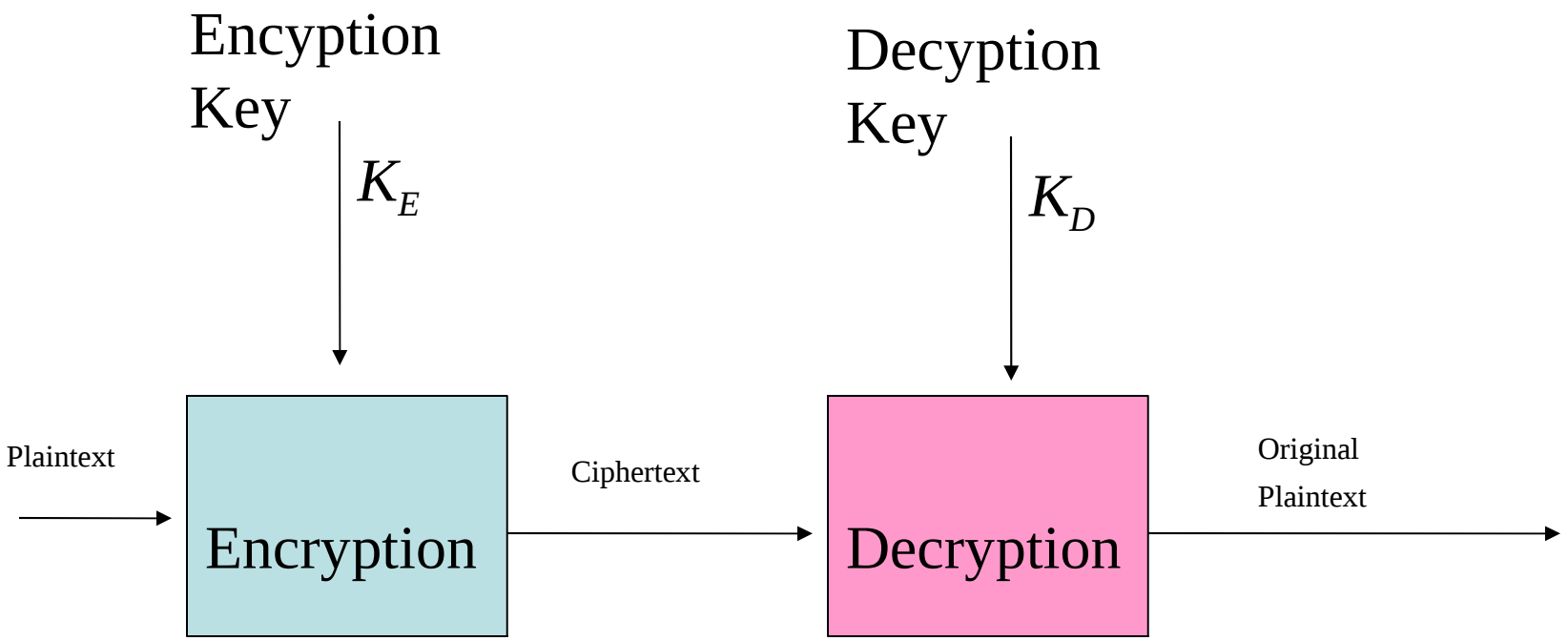
Cryptography – Types

- Symmetric
  - When encryption and decryption keys are the same
    - D and E are mirror images of each other
      - $P = D(K, E(K,P))$
- Asymmetric
  - When the encryption and decryption keys are different
    - $P = D(K_D\ E(K_E, P))$

Cryptography – Types
Symmetric



Cryptography – Types
Asymmetric



Crypto Analysis

- Cryptanalyst's job is to break an encryption
  - Deduce original message from ciphertext
  - If actual decryption algorithm can be deduced, can break encryption of all messages sent by sending party

  - How do you break an algorithm?
    - Use a variety of information
      - Encrypted messages,
      - Known encryption algorithms,
      - Intercepted plaintext, math or statistical tools, ingenuity and luck!

Breakable Encryption

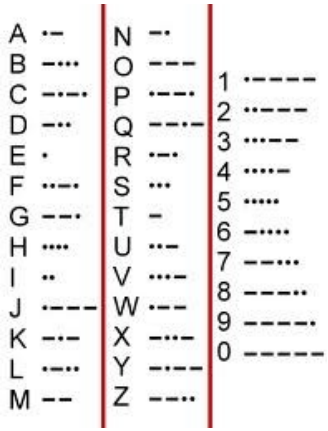
- Breakable Algorithm
  - Given enough time and data, analyst can determine algorithm
    - Yet, may be impractical to try to break
    - Example
      - 25 character message – just uppercase letters
      - So, 26^{25} possibilities
      - If computer can perform 10^{10} operations/sec then finding correct decipherment would take 10^{11} years
      - What to do?
        - Cryptanalyst will reduce search space

Encryption Techniques

- Two types of Encryption Techniques

1. A transposition cipher an encoding process that does not change any letters of original message, but changes **position of letters**
   - One simple transposition cipher reverses order of letters
   - For example, message
     - THE GAME IS AFOOT becomes
       EHT EMAG SI TOOFA
   - Easy to recognize and decode
   - Analogy, transposition ciphers are like jigsaw puzzles
     - All pieces are present, just matter of putting them in correct order

Encryption Techniques

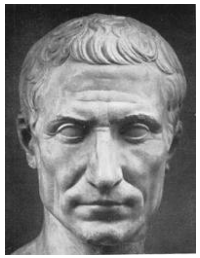


2. A substitution cipher, process that maintains order of letters but changes their identity
   - Each letter is replaced by another letter or symbol
   - Ex., Morse Code is substitution cipher, each letter is replaced by specific set of dots and dashes
   - Many substitution ciphers use only one alphabet, and are called monoalphabetic
   - Substitute one and only one letter for a particular letter in message

Substitution cipher

- For example, every T in message is replaced by the same substitute letter or symbol
- What's the problem with that?
  - Cipher scheme easy to remember, but also vulnerable to "cracking" using frequency analysis - letter counting
  - Encoded message derived using monoalphabetic substitution, can be "cracked"
  - By comparing frequency of letter occurrences in coded message with frequency of letter occurrences in language used for message

Substitution cipher



- What was first recorded use of Substitution Cipher?
  - **Caesar Cipher**
  - Julius Caesar was first to use this crypto scheme
    - Also called a shift cipher
    - A key number, k is agreed upon by sender and receiver
    - Then standard alphabet is shifted k positions so that the kth letter is substituted for letter A, the k+1st for B, etc and the alphabet is wrapped to maintain a one-to-one correspondence

Substitution cipher

- Example Caesar Cipher
  - Caesar used a shift of 3 places, so a plaintext letter, pi was enciphered as a ciphertext letter, ci by the rule:

$$Ci = E(pi) = pi + 3$$

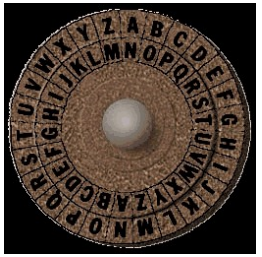
  - Example
    - T R E A T Y I M P O S S I B L E
    - w u h d w b l p s r v v l e o h

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

Early Ciphers

- Needed to be easy
  - Not written down
  - Very easy to break
  - Secure encryption shouldn't allow an interceptor to use small piece of ciphertext to predict entire pattern

Analysis of Caesar Cipher



Many clues from the ciphertext

a) Breaks between words are preserved

b) Double letters are preserved = SS = vv

c) Letters always map to the same substituted letter

T I E -> w l h

Key Substitution Cipher

- Other Substitution Ciphers
  - Permutation is reordering of elements of a sequence
  - One way to scramble letters of an alphabet is to use a

    key, A word that controls permutation
  - If key is an actual word, sender or receiver first writes alphabet and then writes key under it

Key Substitution Cipher

Use: **word** as the key

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

w o r d a b c e f g h u j k l m n p q s t u v x y z

Key is short so most plaintext letters are one or two positions off

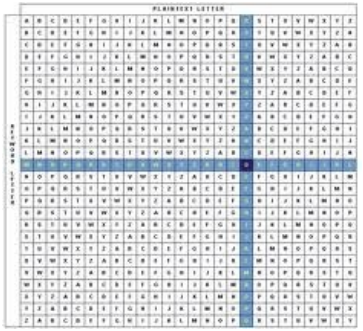
Longer keywords – distance is greater and less predictable

Use: **professional** as key

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

p r o f e s i n a l b c d g h j k m q t u v w x y z

Other Substitution Schemes



- To make substitution ciphers more secure, use more than one alphabet

- Such ciphers are called polyalphabetic,
  - Same letter can be represented by different letters when encoded
  - One-to-many correspondence makes frequency analysis much more difficult in order to crack code

- One such cipher named for Blaise de Vigenere, a 16th century Frenchman
  - The **Vigenere cipher**

Vigenere cipher

- ... is a polyalphabetic cipher based on using successively shifted alphabets
  - A different shifted alphabet for each of 26 English letters
- Based on table shown in next slide plus use of keyword
  - Letters of keyword determine shifted alphabets used in encoding process

Vigenère Tableau

ABCDEFGHIJKLMNOPQRSTUVWXYZ
BCDEFGHIJKLMNOPQRSTUVWXYZA
CDEFGHIJKLMNOPQRSTUVWXYZAB
DEFGHIJKLMNOPQRSTUVWXYZABC
EFGHIJKLMNOPQRSTUVWXYZABCD
FGHIJKLMNOPQRSTUVWXYZABCDE
GHIJKLMNOPQRSTUVWXYZABCDEF
HIJKLMNOPQRSTUVWXYZABCDEFG
IJKLMNOPQRSTUVWXYZABCDEFGH
JKLMNOPQRSTUVWXYZABCDEFGHI
KLMNOPQRSTUVWXYZABCDEFGHIJ
LMNOPQRSTUVWXYZABCDEFGHIJK
MNOPQRSTUVWXYZABCDEFGHIJKL
NOPQRSTUVWXYZABCDEFGHIJKLM
OPQRSTUVWXYZABCDEFGHIJKLMN
PQRSTUVWXYZABCDEFGHIJKLMNO
QRSTUVWXYZABCDEFGHIJKLMNOP
RSTUVWXYZABCDEFGHIJKLMNOPQ
STUVWXYZABCDEFGHIJKLMNOPQR
TUVWXYZABCDEFGHIJKLMNOPQRS
UVWXYZABCDEFGHIJKLMNOPQRST
VWXYZABCDEFGHIJKLMNOPQRSTU
WXYZABCDEFGHIJKLMNOPQRSTUV
XYZABCDEFGHIJKLMNOPQRSTUVW
YZABCDEFGHIJKLMNOPQRSTUVWX
ZABCDEFGHIJKLMNOPQRSTUVWXY

Historical Note

Standard Vigenere was main cryptographic system used by Confederated States during American Civil War, following four key phrases used by Confederates have survived to this day:

- **IN GOD WE TRUST**
- **COMPLETE VICTORY**
- **MANCHESTER BLUFF**
- and, as the war-luck turned:
- **COME RETRIBUTION**

Vigenere Cipher

- For example, suppose we wish to encipher the plaintext message:

  TO BE OR NOT TO BE THAT IS THE QUESTION

  Keyword = RELATIONS

- We begin by writing keyword, repeated as many times as necessary, above plaintext message.
- To derive ciphertext using tableau, for each letter in plaintext, find intersection of row given by corresponding keyword letter and column given by plaintext letter to get ciphertext letter

Vigenere Cipher

Keyword: RELAT IONSR ELATI ONSRE LATIO NSREL

Plaintext: TOBEO RNOTT OBETH ATIST HEQUE STION

Ciphertext: KSMEH ZBBLK SMEMP OGAJX SEJCS FLZSY

- Decipherment of an encrypted message is equally straightforward. One writes the keyword repeatedly above message:

Keyword: RELAT IONSR ELATI ONSRE LATIO NSREL

Ciphertext: KSMEH ZBBLK SMEMP OGAJX SEJCS FLZSY

Plaintext: TOBEO RNOTT OBETH ATIST HEQUE STION

- Use keyword letter to pick a column of table and then trace down column to row containing ciphertext letter. The index of that row is plaintext letter

Vigenere Cipher

- The strength of Vigenere cipher against frequency analysis can be seen in previous example
- Note there are 7 'T's in plaintext message
  - T's encrypted by 'H,' 'L,' 'K,' 'M,' 'G,' 'X,' and 'L'
- This successfully masks frequency characteristics of English 'T'
- Thus, any message encrypted by a Vigenere cipher is a collection of as many simple substitution ciphers as there are letters in the keyword

Cracking the Vigenere Cipher

- For 300 years Vigenere cipher was considered to be practically unbreakable
- Then in 1863 Prussian military officer devised method to determine length of keyword and then divide message into simpler forms to which letter frequency analysis could be applied
- For further information see URLs

Applet that shows how it can be broken
http://math.ucsd.edu/~crypto/java/EARLYCIPHERS/Vigenere.html

Another Analysis
http://www.simonsingh.net/The_Black_Chamber/vigenere_cracking.html

One-time Pad

- Supposed to be in theory perfect cipher
- Name comes from method
  - Large, non-repeating set of keys written to pads of paper by women in DOD!!
- If keys are 20 characters long, one key per page and had to send a message of 300 characters ....
- Then, would use next 15 pages of keys
- Sender would write keys one at a time above plain text and encipher plaintext with Vigenère Tableau chart
- Sender then destroys keys

- Describes how this was used

http://www.thefullwiki.org/Venona_project

One-time Pad

– For encryption to work, receiver needs same pad as sender
  – Then, takes correct number of keys and deciphers message as if it were a plain substitution with a long key

- Problem with this method?
- One-time pad has some problems
  – Need to synchronize between sender and receiver
  – Need for unlimited number of keys
  – Key generation is not hard but
    » Distribution, storing and accounting for keys is hard … ongoing problem

  Example:
  http://www.fourmilab.ch/onetime/otpjs.html

One-time Pad

- Random Numbers
  - Close approximation of a one-time pad is random-number generator
  - Computer random numbers are not absolutely random
  - Really sequence with a long period

    If wanted to use random number generator to send a message,
    - Generate 300 random numbers and scale them to be between 0 and 25
    - Use one number to encipher each character

Book Ciphers

Another way to generate "random numbers" is to use books, music or other objects with structure

- Sender and receiver need access to same object
- Example
  - Sender and receiver agree to use same phone book and start on page 35
  - Use two middle digits of each 7 digit phone number
    - (ddd – DDdd) mod 26 as key letter for substitution cipher
    - Use Vigenère Tableau chart

Book Ciphers

- Passage from Descarte's Meditation:
  - I am, I exist, that is certain.
- Example message: Machines cannot think

Plaintext: MAC H I NESCA NNOTT HIN K

- Key: i a m i e x i s t t h a t i s c e r t
- Then use a table, like Vignere tableau

Cipher: u a o p m k m k v t u n h b l j m e d

Book Ciphers

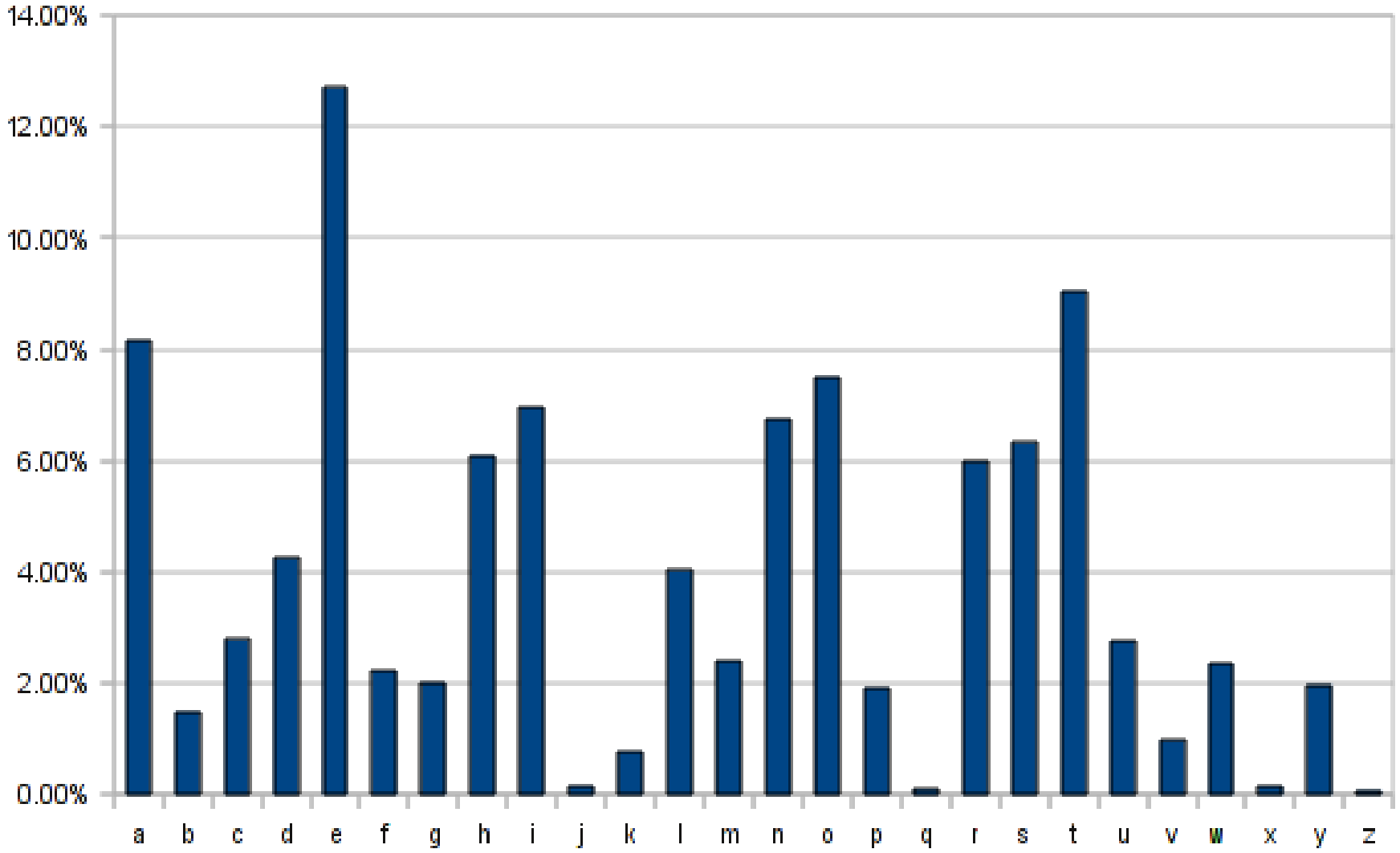
- How to Break it?
  - Neither original message or key text is evenly distributed
  - Cluster around high frequency letters
    - 50% of all letters, A E O T N I
    - Compute probability of both being one of 6 is

      .5 X .5 = .25 or 1 in 4 chance that both letters are in message and key
    - Otherwise need to consider 26^{19} possible encodings

Character Frequencies

- In most languages letters are not equally common
- In English **e** is by far most common letter
  - Have tables of single double & triple letter frequencies
  - These are different for different languages

Frequency of Letters in English



Encryption Techniques

- Transposition
  - Rearranging letters of message
  - Want **diffusion** – wide spreading of information across ciphertext
  - Try to break established pattern
  - Column transpositions

    c1 c2 c3 c4 c5

    c6 c7 c8 c9 c10

    c11 c12 etc.

Encryption Techniques

- Transposition
  - Form ciphertext by reading from columns

This is a message to show how a columnar transposition works

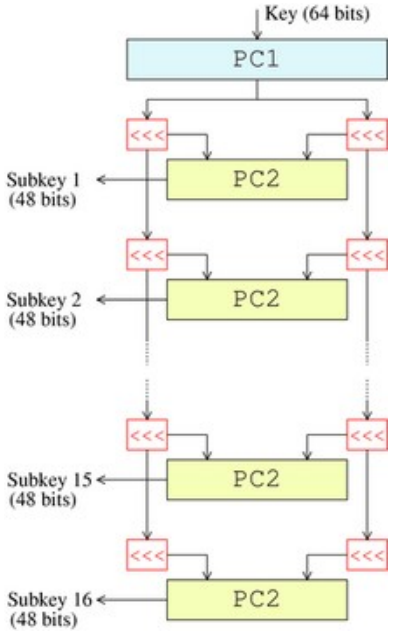
Thisi
sames
saget
oshow
howac
olumn
artra
nspos
ition
works

tssoh oaniw hasso lrsto imghw
utpir seeoa mrook lstwc nasns

Length of message just happens to be a multiple of 5

If message length is not equal length of a row use some infrequent letters to fill in gaps

Encryption Techniques

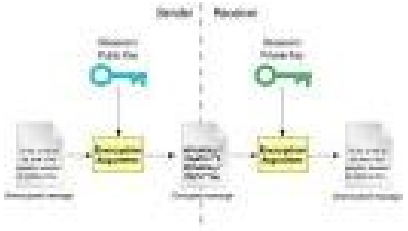


- Combination Approach
  - Substitution and Transposition
  - Cipher building blocks
  - Combination of two ciphers
    - Product Cipher – ciphers are performed one right after another $E_2(E_1(P, K_1)\ K_2)$
    - Just because you apply two ciphers doesn't mean result is stronger than each individual cipher

Encryption Techniques

- Properties of Trustworthy Encryption Systems
  - Commercial users have requirements must be satisfied when using encryption
    - Encryption is **commercial grade** if it meets these requirements:
      - Based on sound mathematics – derived from solid principles
      - Analysed by experts, found to be sound review by critical outside experts is essential
      - Stood Test of Time – new algorithm gains popularity, people continue to review it both for math foundations and way it builds upon those foundations
        - Flaws of algorithms are discovered soon after their release

Encryption Techniques



- Three Popular Algorithms
  - DES – Data Encryption Standard
  - RSA – Rivest – Shamir – Adelman
  - AES – Advanced Encryption Standard

    DES and RSA – meet above criteria

    AES – new – meets first two and is starting to achieve widespread adoption

The End

Chapter 10 in our Book.

- Reading: Some reading here, public key for now
  http://en.wikipedia.org/wiki/Cryptography