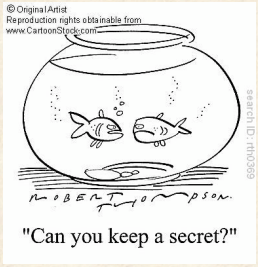
UNIVERSITY of WISCONSIN-MADISON
Computer Sciences Department

CS 202: Introduction to Computation — Professor Andrea Arpaci-Dusseau

How can a computer… send a secret?



Why do we care about security?

Internet security
- Need a secure way to buy things online
- Need to manage bank account securely

Communicating over an unprotected channel

Messages received by many intermediaries, any could be eavesdroppers
- Home router => Roommate
- Building switch => Building manager
- Local network => ISP employee
- Backbone => Government

Secure Telephone Game

Scenario:
- Sender tries to communicate value of a random number (0-100) to a remote receiver
- Message must be "transmitted" from sender to receiver by evil eavesdroppers
- Can the receiver figure out the number and not the eavesdroppers?

How are you going to do this?
- Sender and receiver can communicate out-of-band ahead of time

Results?

Possible outcomes after 1 attempt
- Did receiver figure out number?
- Did eavesdroppers figure out number?

Repeat experiment
- Try to send another number from sender to receiver through eavesdroppers
- Easier for eavesdroppers to figure out secret the more examples they see!

What approaches did everyone use?

Information Security: 4 Goals

Confidentiality
- Eavesdroppers cannot understand messages

Integrity
- Eavesdroppers cannot modify message undetectably

Availability
- Messages reach their destination

Authenticity
- Sender and receiver are who they say they are

What is Cryptography?

Definitions
1. secret writing
2. the enciphering and deciphering of messages in secret code or cipher

Used throughout History
- Ancient ideas (pre-1976)
- Complexity-based cryptography (post-1976)

Essential component of digital world

Themes of Today's Lecture

1. Seeing info does not mean understanding it
2. Creating problems often easier than solving
3. Complete strangers can exchange and agree on secret information

Basic Approach of Cryptography

Theme 1: Seeing info does not mean understanding it

Idea: Convert data to form that doesn't make sense to others
- Clear text (plain text): Initial readable text
- Cipher text: Encrypted version of clear text

Steps
1. Sender: Encrypt clear text to cipher text
2. Cipher text can be stored in readable file or transmitted over unprotected channels
3. Receiver: Decrypt cipher text to clear text

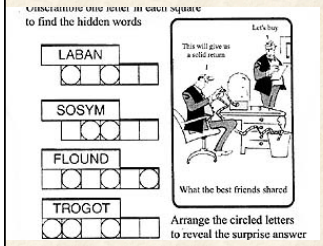
Theme 2: Creating problems often easier than solving

Word Scramble of "scratch"

Algorithm to create?

- tarcchs
- Complexity?

O(N) steps



Algorithm to solve?

Try every permutation

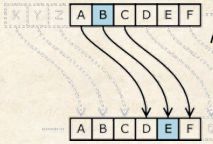
atrcchs

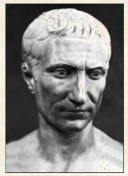
atrccsh…

N * N-1 * N-2 … * 1 = O(N!)

Letter Scrambling: Ancient Cryptographic Idea

Example 1: Caesar or shift cipher (c. 100 BC)





Ex: Clear text -- Attack at noon

Cipher text -- Dwwdfn dw qrrq

How would eavesdropper crack the code?

Enumerate all 26 possibilities until see reasonable

Shift of 1? Cvvcej cv pqqp

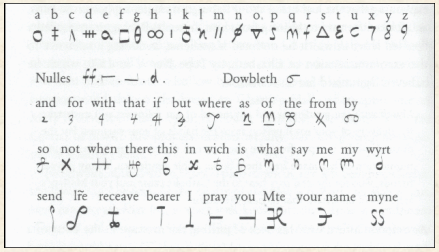
2? Buubdl bu oppo

3? Attack at noon – Got it!

Substitution Cipher

Example 2: Cipher of Queen Mary of Scots, 1587

Wanted to assassinate Queen Elizabeth I



Possibilities?

26 * 25 * 24 * … * 1

Huge number of possibilities if do stupid brute force attack

Still cracked and Queen Mary was executed!

Mafia Boss

Example 3: "Boss of bosses" Bernardo Provenzano, captured after 40 years

Sent "pizzini" (little messages on scraps of paper) using weak cipher

- "…I met 512151522 191212154 and we agreed that we will see each other after the holidays…,"
- 5 = B, 12 = I, 15 = N, etc

"It will keep your kid sister out, but it won't keep the police out." - Bruce Schneier (Cryptographer)

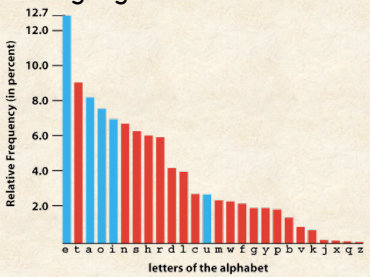
From Discovery Channel, Apr 17 2006

How to Crack the Code?

Use characteristics of language

1. Frequency Analysis

E most common, then T



2. Letter combinations or sequences
   - Bigrams: Pairs of letters
     - ST, NG, TH, and QU are common, NZ and QJ are rare
   - Trigrams: Three letters
     - THE most common

One-time Pad

Approach
- Each letter of message shifted by different amount
- Amount to shift is stored on a pad that can only be used once

Advantage
- Impervious to frequency analysis

Disadvantage
- How do you get one-time pad to sender and receiver???

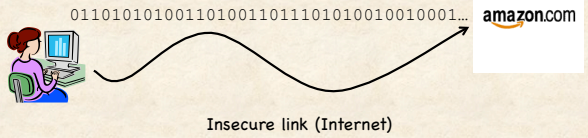
Letter scrambling in World War II

Enigma
- Used by Nazi Germany (1940's)
- Broken by British (Turing), Polish
- "Won us the war." – Churchill
- Recommend: Neal Stephenson's novel Cryptonomicon

Moral: Computers → need for new ideas for encryption

Musings about one-time pad

Incredibly strong security
- Encrypted message "looks random"
- Equally likely to be encryption of any n-bit string
- Can't do frequency analysis on cipher text!



How can you and a friend agree on one-time pads?

How can you and Amazon agree on one-time pads?

Public Key Encryption

Exchange one-time pad with public key encryption

2 keys for every user
- public key known to everyone
- private key know only to the user

Encrypt/Decrypt
- Private key can decrypt messages encrypted with public key
- Public key cannot decrypt messages encoded by the public key

Example: https protocol
- User doesn't need to know anything about protocol
- Accept certificate that doesn't look quite right?

Public-key encryption: Conceptual Story

"Box that clicks shut, and only Amazon has the key to open it."


Example: Key exchange
- User generates one-time pad (random string)
- Put it in box, lock it with public key, ship it to Amazon
- Amazon opens box with private key, recovers random string

Theme 2 Again: Easy to Generate, Hard to Solve

Public/private keys based on Factoring of Large Numbers

Generate Integer that is Product of 2 Primes: Easy
- Pick two *n*-bit prime numbers *p, q*
  - Good value for n is 128 (more bits, stronger encryption)
- Multiply together to get $r = pq$

Use r to generate public key; r, p, and q for private keys
- Mathematical details omitted; Wikipedia has decent explanation!
- If someone can infer p and q from r, break encryption!

Factoring problem : Hard
- Given *r*, can you find *p* and *q*?

Factoring is Hard!

Given *r*, can you find *p* and *q*?

Can you suggest a basic algorithm?
- Set variable *try* to 2
- Repeat…
  - Is r mod try == 0?
    - Yes → p = try, r/try =q
    - Done!
  - Change try by 1

How many steps? (Could be optimized!)
- How many values of r? r represented by 128 bits
- 2^{128} (approximately 3.4×10^{38})

Despite many centuries, no efficient algorithms
- Rely on difficulty every time you use e-commerce

Common Cryptography Themes

Seeing info does not mean understanding it

Creating problems often easier solving

- Shifting letters easy; guessing how much they were shifted is hard
- Multiplying two integers is easy; factoring is hard

Complete strangers can exchange secrets

- Send one-time pad encrypted with receiver's public key; no one else can read

Steganography: How to hide *existence* of message?

Hide message in digital photos

- store message in least important bits of photo

Today's Summary

Cryptography: Basis of E-commerce

Announcements

- Homework 6 graded
- Homework 8 due Friday
  - Create Trivia Game; Randomly ask 10 questions, check answers
  - Extra Credit: If implement optional feature, can submit to Gallery
- Exam Review on Friday
- Exam 2 on Monday
  - Similar in style to Exam 1
  - Covers all material since Exam 1