Cryptography, briefly

John Mitchell

802.11b slides from Dan Boneh

Cryptography

- Is
  - Tremendous tool
  - Basis for many security mechanisms
- Is not
  - The solution to all security problems
  - Secure unless implemented properly
  - Secure if used improperly

Basic Concepts in Cryptography

- Encryption scheme:
  - functions to encrypt, decrypt data
  - key generation algorithm
- Secret vs. public key
  - Public key: publishing *key* does not reveal key^{-1}
  - Secret key: more efficient; can have $key = key^{-1}$
- Hash function
  - map text to short hash; ideally, no collisions
- Signature scheme
  - functions to sign data and confirm signature

Cryptosystem

- A cryptosystem consists of five parts
  - A set P of plaintexts
  - A set C of ciphertexts
  - A set K of keys
  - A pair of functions

    encrypt: $K \times P \rightarrow C$

    decrypt: $K \times C \rightarrow P$

  such that for every key $k \in K$ and plaintext $p \in P$

    decrypt(k, encrypt(k, p)) = p

OK def'n for now, but doesn't include key generation or prob encryption.

Primitive example: shift cipher

- Shift letters using mod 26 arithmetic
  - Set P of plaintexts {a, b, c, … , x, y, z}
  - Set C of ciphertexts {a, b, c, … , x, y, z}
  - Set K of keys {1, 2, 3, … , 25}
  - Encryption and decryption functions

    encrypt(key, letter) = letter + key (mod 26)

    decrypt(key, letter) = letter - key (mod 26)
- Example

  encrypt(3, stanford) = vwdqirug

Evaluation of shift cipher

- Advantages
  - Easy to encrypt, decrypt
  - Ciphertext does look garbled
- Disadvantages
  - Not very good for long sequences of English words
    - Few keys – only 26 possibilities
    - Regular pattern
      - encrypt(key,x) is same for all occurrences of letter x
      - can use letter-frequency tables, etc

Letter frequency in English

- Five frequency groups [Beker and Piper]

E has probability	0.12
TAOINSHR have probability	0.06 - 0.09
DL have probability	~ 0.04
CUMWFGYPB have probability	0.015 - 0.028
VKJXQZ have probability	< 0.01

Possible to break letter-to-letter substitution ciphers.

- 1400: Arabs did careful analysis of words in Koran
- 1500: realized that letter-frequency could break substitution ciphers

One-time pad

- Secret-key encryption scheme (symmetric)
  - Encrypt plaintext by xor with sequence of bits
  - Decrypt ciphertext by xor with same bit sequence
- Scheme for pad of length n
  - Set P of plaintexts: all n-bit sequences
  - Set C of ciphertexts: all n-bit sequences
  - Set K of keys: all n-bit sequences
  - Encryption and decryption functions

    encrypt(key, text) = key $\oplus$ text (bit-by-bit)

    decrypt(key, text) = key $\oplus$ text (bit-by-bit)

Evaluation of one-time pad

- Advantages
  - Easy to compute encrypt, decrypt from key, text
  - As hard to break as possible
    - This is an information-theoretically secure cipher
    - Given ciphertext, all possible plaintexts are equally likely, assuming that key is chosen randomly
- Disadvantage
  - Key is as long as the plaintext
    - How does sender get key to receiver securely?

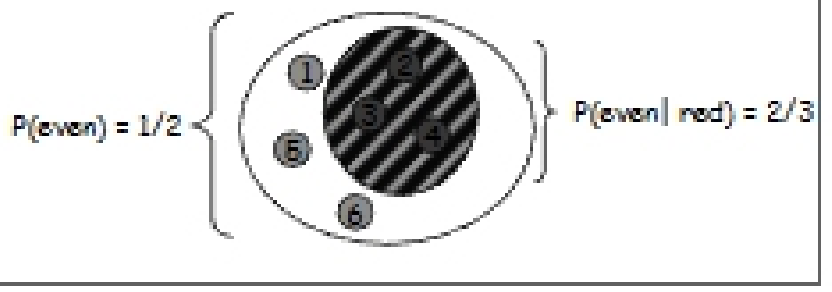
Idea can be combined with pseudo-random generators ...

What is a "secure" cryptosystem?

- Idea
  - If enemy intercepts ciphertext, cannot recover plaintext
- Issues in making this precise
  - What else might your enemy know?
    - The kind of encryption function you are using
    - Some plaintext-ciphertext pairs from last year
    - Some information about how you choose keys
  - What do we mean by "cannot recover plaintext" ?
    - Ciphertext contains no information about plaintext
    - No efficient computation could make a reasonable guess

Information-theoretic security

- Remember conditional probability...
  - Random variables X, Y, ...
  - Conditional probability P(X=x|Y=y)
    - Probability that X takes value x, given that Y=y



Information-theoretic security

- Cryptosystem is *info-theoretically secure* if

P(Plaintext=p | Ciphertext=c) = P(Plaintext=p)

Key	Plaintext	Ciphertext
H	Land	1
H	Sea	2
T	Land	2
T	Sea	1

- Ciphertext gives no info about plaintext

Prob(1 is for Land) = Prob(1 is for Sea)

assuming that all keys are equally likely