1 Announcements

Web page will be up by Friday.

2 Role of Cryptography (later crypto itself)

Broad set of apporaches to conceal

2.1 Steganography

- Writting with milk example: conceals the fact that there is a message
- Microdot example: documents shrunk to the size of a period and mailed out with an inoccuous letter.
- 24 bit images can carry information. 1 bit per pixel will not be visable to the human eye.

2.2 Caeser Cipher - advance each letter by three

$HELLOWORLD \rightarrow KHOORZRUOG$

- Key - the info needed to encipher or decipher a message.
- Cipher - maps characters into characters.

2.3 Secret Writing Alternative: Code

two sets of phrases. One is a garbled mess, the other the meaning of the garbled mess. Meaning of code depends on the code book used.

```
Example Code Code Groups
------------ -----------
hello AXZB
i love XCVJ
i hate BSYK
broccoli TRUX
spinach PLUG
```

2.4 Supercipherment

Combined Approach using Code and Encipherment. Use code book then encipher the code groups. Difficult to break, but has been broken in the past.

2.5 Terms

- Plaintext(M): message to be enciphered. Also referred to as cleartext.
- Ciphertext(C): output of encipherment function. Not generally understandable.

2.6 Note on International Usage

- Encipher - to make unreadable without destroying the message.
- Encrypt - similar to crypt, which is a burial ground or cemetary sounds like you are going to bury the message.

2.7 Notation

- Encipher(Ek): Algorithm to encipher message
- Decipher(Dk): Algorithm to decipher message
- Key(k): information used to encipher message using a common algorithm.
- Encipher: $E_k(M) = C$
- Decipher: $D_k(C) = M$

Clasical systems: encrypt and decrypt keys are the same.

3 Good Cryptosystems and Kirchoff's requirements

1. Enciphering and deciphering is EFFICIENT for all keys. Should not take forever to get message.
2. Easy to use. Not as important now because of computerization. Problem with hard to use cryptosystems is that mistakes tend to be made and message is never deciphered correctly, or a crib will be given to those trying to read the message.

   note: A crib is a hint to the adversary about what is contained in the message.
3. The strength of the system should not lie in the secrecy of your algorithms. The strength of the system should depend the secrecy of your key. Manuals can be stolen. It is Easy to generate new key, but hard to replace entire cryptosystem.

4 Secrecy or Confidentiality

Encipher messages to protect secrecy

1. Intecepted ciphertext should not lend clues to what the deciphering algorithm is. It should be computationally infeasable (C.I.)
2. It should be C.I. to find M from C.

5 4 attacks on cryptosystems

1. ciphertext only attack given the ciphertext, what is the original message?
2. known plaintext attack have both the plaintext and ciphertext. What is the key?
3. chosen plaintext attack give plaintext to adversary, get back enciphered message.
4. chosen ciphertext give ciphertext to adversary, get back plaintext. Not very practical and noted here for theoretical completeness only.

6 Authentication or Integrity

Data is not to be altered. requirements for integrity:

1. C.I. to find E_k. Keeps adversary from changing encrypted messsage.
2. C.I. to find C' such that $D_k(C')$ is valid.

7 Computationally Infeasable (an explanation)

Degrees of security

1. unconditionally secure. Theoretically unbreakable. Only example of this is the one time pad.
2. conditionally secure. Computationally infeasable in practice. cipher is unbreakable or bt the time you do break it, the message contained within is not irrelevant.
3. weak Can be broken easily.

8 Quality of strength of ciphers: Information Theory

How much uncertainty is there? What is the measure of uncertainty? How many bits of the ciphertext are known?

- Goal of attacker: zero uncertainty.
- Goal of cryptanalyst: lg(n) uncertainty - the more uncertainty the better.

8.1 Measure of Uncertainty

The entropy function measures uncertainty. P_i denotes the probability of the i^th key being the actual key.

$H(P_1 \cdots P_n) = -\lambda \sum_{i=1}^{n} P_i \lg P_i$

Lambda is a constant and is set equal to 1 for out purposes.

$H(P_1 \cdots P_n) = -\sum_{i=1}^{n} P_i \lg P_i$

8.2 Example 1.

We know something is an ascii character
$H = -\sum_{i=1}^{256} P_i \ln P_i$
$H = -\sum_{i=1}^{256} 1/256 \ln 1/256$
$H = -\lg 1/256 = 8$

8.3 Alternative Notation

Using random variables the notation changes.
$H(X) = \sum P(X) \lg 1/P(X)$

8.4 Example 2.

$P(M = A) = 3/4$
$P(M = B) = 1/4$
$H(M) = P(M = A) \lg 1/P(M = A) + P(M = B) \lg 1/P(M = B)$
$= 3/4 \lg 4/3 + 1/4 \lg 4$
$= 3/4 \lg 4/3 + 1/2 = uncertaintyofmeasage.$

8.5 How does this play into cryptography

There are many overloading sounds in English.
GHOTI = FISH
mst ids cn b xprsd n fwr ltrs

8.6 Rate of a Language

The average number of bits of information in each message.
Message length n
$r = H(n)/n$
$r \approx 1.0 - 1.5$
Claude Shannon 1947.

8.7 Absolute Rate

Assume that characters are equally likely. What is the rate?
$H(X) = -\sum_{i=1}^{n} 1/n \lg 1/n = \lg n$
$R = \lg 26 \approx 4.7$