

COMP 424

Lecture 03

Basic Encryption Systems

(Substitution, Transposition, One-Time pads)

Caesar Cipher

- A message is encrypted by substituting each character with a character that is fixed position away in the alphabet.
- $c_i = E(p_i) = p_i + 3$
- Decryption is accomplished by simply subtracting the same fixed number.

Caesar example

TREATY IMPOSSIBLE

wuhdwb lpsrvvleoh

Problems

- Condition that made Caesar's cipher secure:
 - In Caesar's time most people were illiterate.
- Problems:
 - Everybody can read now.
 - patterns in the natural language are preserved in the cipher text
 - Not based on strong foundation. Basically, once you know the trick it's no longer good.

One-Time Pads

- Sometimes considered the “perfect” cipher.
 - A (sufficiently long) sequence of “keys” are generated. (Each key is a single character)
 - To encrypt a plaintext message of n characters in length n keys are consumed from the One-Time Pad.
 - $C_i = (P_i + K_i) \bmod 26$ Where K_i is the i^{th} key
 - $P_i = (C_i - K_i) \bmod 26$
used
 - $P_i = (C_i - K_i + 26) \bmod 26$
 - To avoid performing modulo on negative values we can

Vigenere Tableau

- Instead of mod 26 we can use another device
 - | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 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 |
| A | Z | Y | X | W | V | U | T | S | R | Q | P | O | N | M | L | K | J | I | H | G | F | E | D | C | B | A |
| B | Y | X | W | V | U | T | S | R | Q | P | O | N | M | L | K | J | I | H | G | F | E | D | C | B | A | Z |
| C | X | W | V | U | T | S | R | Q | P | O | N | M | L | K | J | I | H | G | F | E | D | C | B | A | Z | Y |
| D | W | V | U | T | S | R | Q | P | O | N | M | L | K | J | I | H | G | F | E | D | C | B | A | Z | Y | X |
| ... | | | | | | | | | | | | | | | | | | | | | | | | | | |
 - Encryption is done by: Plain text is used as the row index, the key is used as the column index to identify the ciphertext.
 - Decryption is done by: Cipher text is used as the row index, the key is used as the column index to lookup the plaintext.
 - Has the advantage that encryption/decryption can be done by hand. (vs trying to mod by 26.)

One-Time Pad Advantage

- Advantage:
 - Unbreakable. Since the keys are perfectly random so is the ciphertext. Patterns cannot be found since they do not exist.

One-Time Pad Problems

- Disadvantage:
 - Both S and R need to obtain pads of keys in a perfectly secure manner.
 - Pads must be kept in perfect sync. (interception of a single message can render entire pad useless)
 - Requires an unlimited sequence of keys (random values)
 - Corollary: must be a truly random sequence. (Pseudo Random Number Generators are not Random)

Vernam Cipher

- Uses a combining function that has the property:

$$P \circ K = C$$

$$C \circ K = P$$

- Exclusive OR for example.
- Mod 26 (for 26 character alphabet will also do)

Book Ciphers

- Actually any arbitrary book or text can be used as a source of random numbers.
- This can provide an easily agreed upon and long sequence of random information.
- Telephone books... start at the 35 page. Use the middle two digits... XXX-DDXX of each telephone entry.
- Long passages of prose (Descarte's Meditation.)

Book Problems

- Are they also unbreakable?
 - No! why? Neither the plaintext nor the key stream is truly random. Certain key characters will occur with greater frequency than others.
 - Frequent plaintext keys will frequently be encoded by the same cipher key.
 - This leads to some probabilistic code breaking based on the frequency distribution. (Similar to Caesar ciphers or other monoalphabetic ciphers)

Tranposition Ciphers

- Substitution seeks to provide *confusion*
 - Information is scrambled to prevent understanding
- Transposition seeks to provide *diffusion*
 - Information is diffused to hinder understanding.

Columnar Transposition

- THIS IS A MESSAGE TO SHOW HOW
COLUMNAR TRANSPOSITION WORKS

- T H I S I
S A M E S
S A G E T
O S H O W
H O W A C
O L U M N
A R T R A
N S P O S
I T I O N
W O R K S

- Tssoh oaniw haaso lrsto imghw
utpri ...

Analysis

- Only a constant amount of work is required for the transposition of each character so it requires no more time than a substitution or One-Time pad to encrypt or decrypt. So it is proportional to the length of the message.
- Storage space is greater though.
(Substitutions so far have been limited to $26*26$ (Vigenere table))
- Transposition requires space proportional to the message itself ($>$ or even $\gg 26*26$)

Further...

- Must obtain entire message before a single character can be properly encrypted or decrypted.

Diagrams, Trigrams and Patterns

- Not only do individual letters appear with higher frequency in a language, pairs and triplets of letters, called digrams and Trigrams appear with great er frequency...
 - “EN” “ER” “TH” appear with far greater frequency than “VK”
- This information can be used to help match up letters that have been transposed from their real position. If its “vk” the “v” probably belongs somewhere else.

Cryptanalysis

- If all the letters appear with the proper frequency then we can be relatively certain a transposition is responsible for the cipher text and not by substitution.
- Moving windows can then be used to locate common digrams or trigrams.
- This can lead to discovering how many columns existed and thus to decryption.

Combinations of Approaches

- If one is good, two are better.
- The combination of two ciphers is called a “product cipher”
- Typically performed $E_2(E_1(P, k_1), k_2)$
- The result is not always stronger than, or even as strong, the individual ciphers.

Stream vs Block Ciphers

- Substitutions are considered “stream” ciphers:
 - Individual characters of plaintext are encrypted immediately into characters of ciphertext.
- Transpositions are “block” ciphers:
 - A group of plaintext characters are encrypted as a block.
 - The block can be the entire message

What Makes a Good Cipher?

- Shannon's characteristics
 - Amount of secrecy required determines the amount of labor appropriate for the encryption and decryption.
 - The set of keys and enciphering algorithm should be free of complexity
 - The implementation of the process should be as simple as possible.
 - The size of the enciphered text should be no larger than the text of the original message.

Commercial Grade

- Based on sound mathematics
- Analyzed by competent experts and found to be sound
- It has stood the test of time.*
 - DES (Data Encryption Standard, well 3DES actually)
 - RSA (Rivest-Shamir-Adelman)
 - AES (Advanced Encryption Standard)*