"Security is not an afterthought. It starts at design"

9/24/2017

Group HA-Week 4

Brute Force

- Don't talk about brute force unless you talk about the verification method.
- Brute Force is the Dumbest Adversary.
- Can't rule out. Always a possibility.

Caesar Cipher

- $C \overset{26}{\equiv} m+k$

Affine Cipher

- higher complexity by adding another factor
- $C \overset{26}{\equiv} am+b$ – Encryption
- $m \overset{26}{\equiv} (c-b)a^{-1}$ – Decryption

Cyber Security

- RSA is the foundational reason of Cyber Security
- The goal of building a 'box' is to have a higher complexity for Brute Force.

Substitution

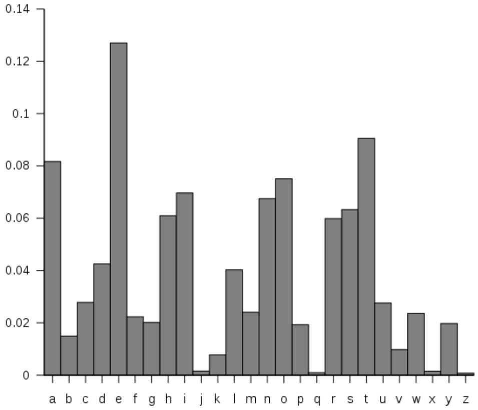
Example

  - Substitution: a b c d e … … z
  - Truth Table: d k … … … …
  - Gives a complexity of 26!
    - To get the length, take the $\log_2 26!$
- Due to the complexity, there is a significant amount of work to recover the key.
- "Increase Brute Force complexity us necessary but not sufficient".

Frequency Analysis

- Substitution is a 1-1 mapping.
- It preserves the structure of the English Language.
  - You can analyze the frequency a letter appears in a cipher and map it to the true letter due to the Frequency Analysis of the English Language.
- Ex.
  - Message: retrieve we are attacked
  - Cipher: w j … … … …
    - r mapped to w.

        - e mapped to j.
        - etc.
    - Draw a frequency table of the cipher and compare to a frequency table of the English Language.



        - Map similar frequencies to each other.

- Complexity to solve by Frequency Analysis.
    - Just the length of the ciphertext to compute the table.
    - Compare with an English Frequency Table.
    - Complexity is broken down to less than 26!

Issues and Goals

- The Adversary has the frequency table advantage.
- Therefore, we must overcome the hole in security.
    - Create a ‘box’ to not preserve the English language attributes.
- The goal is to have a uniform distribution in the frequency table.
- Map Characters depending on multiple factors.
    - Not only input but location.
    - Creates ‘blocks’

Blocked Substitution

- Choose a word/phrase to create ‘blocks’ and shift key.
- Ex.
    - Word: shoes
    - Message: r e t r i e v e w e a r e a t t a c k e d
    - Key: s h o e s s h o e s … … …
    - Cipher: … … … …
- Due to blocked shift key, there is a more uniform distribution on the frequency table.
- More complex to recover key.

Breaking Blocked Substitution

- Broken by Vigenre
- Can be broken if you find out length of key word.

Permutations

- Example: Linc/oln i/s a re/ally/ cool/ guy
  - Here, permutation algorithms is the key, we separate every 4 character and mix them up.
  - Character 1 is mapped to Character 3; L → n
  - Character 2 is mapped to Character 1; i → L
  - Character 3 is mapped to Character 4; n → c
  - Character 4 is mapped to Character 2; c → i
  - Here Linc = nLci
- Here, the possible combinations are 4! = 4 x 3 x 2 x 1.
- This technique blocks the "Frequency Analysis" that was used on Substitution because the letters are preserved.
- **Trigram**: three letter word that appears frequently.
- Real life example of this was that Americans were the "chosen plain text" adversaries against the Japanese during World War II.

Shannon Claude

- Described cryptography as an "art". When you rely on art you miss Analysis. We cannot use the thought "My piece of art looks pretty; therefore, it's good to use".
- Security has to be looked through a Science lens.

Three Principles

1. Without ambiguity define all definitions. "What do I want?". State in a mathematical way what "*what*" means. Defining a goal allows for evaluation.
2. Come up with solutions and prove them. In Science there are no assumptions. Assumption means to know something, but can't prove it.
3. No system shall ever be used without proper *proof*.

Obstacle

- For cypher-text only

$$P(m \mid c) = P(m)$$
$$=> \frac{P(m \cap c)}{P(c)} = P(m)$$
$$=> c = f(mk).\ where\ \left||k|\right| \geq \left||m|\right|$$

- k is the key length. k should be at least equal to your message length.
- To achieve perfect security we need $c = m \oplus k \ \& c' = m' \oplus k'$.
- **One-Time Pad-** operation allows for a one time use of the key.
  - Here, recycle doesn't seem to be possible
  - But have to always share the key.
    - If the key is being shared securely then why not just share the message this way instead.

Solution

- Every bit in *c* should participate in *m*.
- **Diffusion:** If you choose a bit in *m*, Adversary should have a 50% chance of guessing in *c*.
- **Confusion:** Nothing about *c* should contribute to the key.
- Realized that in real life perfect diffusion and confusion is not possible.
- **Avalanche Effect**: multiple iterations of the solution that partially achieved the goal will cover more of the goal after many rounds.
- The “black box” is pre-World War II. Only thing that helps out is the rounds of iterations.
- **There is no proof**: we cannot prove, so we fail Shannon’s second principle.