POIS-2

6/1/2023 (Fri)

Past

- Art of secure communication
  Present
- Science of information security
  Future
- Science of the impossible

Most of the course is focused on the transformation beween past to present. Involves

- Art $\rightarrow$ Science (explicit)
- One problem Multiple problems (semi explicit)
- Multiple tricks which solved nothing $\rightarrow$ One technique to solve everything (not explicit at all)

What captivates a layman is applicability : more applications, more interesting. Since there is just one technique for all problems, it captivates engineers, also Captivates Scientists.

Caesar Cipher

Simple Cipher
Shift every letter of the message by 3 letters, cyclically.

$$\begin{aligned} A &\rightarrow D \\ B &\rightarrow E \\ &\vdots \\ Z &\rightarrow C \end{aligned}$$

So a simple message such as "CRYPTO" will be encrypted to "FUBSWR". But this cannot be used as an encryption scheme, because it does not satisfy the Kirchoff's principle.

Kirchoff's Principle

Security of a system must not depend on the obscurity of the algorithm being used, but on the secrecy of the key.
Why is kirchoff's principle important. Why is a secret algorithm bad ?

- Secure Memory is costly : therefore keeping the algorithm secure requires a huge number of secure bits.
- Reversible Engineering: just because someone doesn't provide info on the algorithm, doesn't mean the algorithm cannot be figured out. An example is frequency analysis on the caesar cipher. Applying caesar cipher on a huge piece of text, keeps the frequency distribution intact, although it is shifted. So "e" the most common letter, will be mapped to "h".
- Updation/Recreation with ease :If someone follows the kirchoff principle, then correcting any mistakes or leaks in the scheme, will be very easy, just replace the key, instead of having to replace the entire system (algorithm).

- Standardization : Every piece of software should be following the same standard, meaning that it should follow the same algorithm. Therefore if everyone is using the same algorithm, the strength of the secure communication, depends on the security of the key.
- Scalability : If there are only two parties communicating, you need not follow kirchoff's principle. However if there are a huge number of parties involved, we need to apply kirchoff's principle. This goes hand in hand with standardization
- Ethical Hacking : If the algorithm is public, there is much more participation in general, which increase the chances that if the system is hacked into, the person hacking is ethical. Therefore instead of using the system for malicious purposes, the hack can be used to improve the system.

Shift Cipher

Slightly better version of the caesar cipher , here instead of shifting by 3 we shift by some secret key.
Key $k \epsilon [0, 25]$

$$c = m + k \bmod 26$$

How to break ? Just use a brute force attack, only 26 possible keys.

Let p_i= probability of i^{th} character in the plaintext (we can refer to common english words, and calculate the probability, this is generally fixed)
Let q_i=probability of i^{th} character in the ciphertext. (number of occurences of that letter in the cipher divided by the total number of letters)
For a shift cipher, there exists k such that

$$p_i = q_{i+k}$$

In order to find out k, we need to calculate the following value

$$X = \frac{\sum p_i^2}{\sum p_i q_{i+k}}$$

The value of the numerator is generally fixed (because as we mentioned, the value of p_i is obtained through historical data on the frequency of the alphabets). It is generally around 0.065
Now for the correct value of k, the value of the denominator will also be around 0.065 (since $p_i = q_{i+k}$ therefore the denominator for correct k will also be $\sum p_i^2$). However for incorrect value of k the value will be much different.

This adds another important principle, the Principle of Large Keyspaces

Mono Alphabetic Substitution Cipher

Patch for the shift cipher, increase the keyspace.

$Alphabet$	A	B	$\dots$	Y	Z
$Cipher$	$p(A)$	$p(B)$	$\dots$	$p(Y)$	$p(Z)$

The cipher has to be a permutation of the original alphabet set. Now the keyspace is basically 26! (total number of permutations)
How to break ? The frequency of the alphabets remain the same, if "A" is mapped to "H", then every "A" is mapped to "H", and no other letter is mapped to "H". Therefore number of "A"s in the message text is equal to the number of "H"s in the cipher text.
Suppose we don't know the plaintext.
Let p_i= probability of i^{th} character in the plaintext (we can refer to common english words, and calculate

the probability, this is generally fixed)
Let q_i=probability of i^{th} character in the ciphertext. (number of occurences of that letter in the cipher divided by the total number of letters)

For a mono alphabetic substitution cipher
$\forall\ i, \exists\ j$ st

$$p_i \approx q_j$$

We will primarily be interested in the following value

$$X = \frac{\sum p_i^2}{\sum p_i q_j}$$

For the correct value of j for every i the value of X is approx equal to 1. Otherwise it would be very different from 1. We would have to check for 26! such distributions of q_j.

Vigner Cipher

Patch for the MAS Cipher. Add a word (key) to the plaintext, concatenate the key multiple times.

$$\begin{array}{cccccccc} & C & R & Y & P & T & O \\ + & C & A & T & C & A & T \\ \hline & F & S & S & S & U & I \end{array}$$

How to break ? Two steps

- Assume the length of the key is known , and then try to break
- Find the key length

What if the key Length is known?

Assume that the key-length is k then basically we are adding the same character to the 0th,kth,2kth.... character
in the plaintext.
Similarly, 1st,k+1th,2k+1th character will have the same character added to it.
Therefore , we will have k different shift ciphers. Which are easy to solve, using the strategy shown above.

How to find the key length?

Assuming that the key length is not too large, say it was l. Then you would have to perform the same operation prescribed above for a given key length l times.