Network and Information Security
Lecture 8

B.Tech. Computer Engineering
Sem. VI.

Prof. Mrudang T. Mehta
Associate Professor
Computer Engineering Department
Faculty of Technology,
Dharmsinh Desai University, Nadiad

Vigenere Cipher

- Plain Text
- $P = p_0\ p_1\ p_2\ p_3$ .....
- Key $K=[\ (k_0\ k_1\ k_2\ ...k_{m-1})\ (k_m\ k_{m+1}\ ...k_{2m-1})...]$

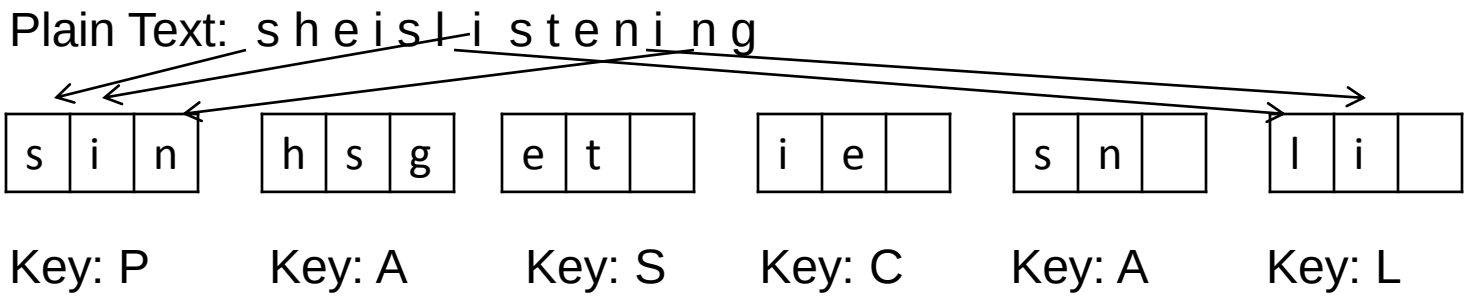
  $k_0\ k_1\ k_2\ ...k_{m-1}$
- Cipher Text
- $CT = c_0\ c_1\ c_2\ c_3$ .....
- Encryption: $c_i = (p_i + k_{\ i \bmod m}) \bmod 26$

  i=0,1,2,3,...
- Decryption: $pi = (c_i - k_{\ i \bmod m}) \bmod 26$

- Vigenere key stream does not depend on the plaintext characters; it depends only on the position of the characters in the plaintext.
- Key stream
- $K = [(k_0\ k_1\ k_2 \ldots k_{m-1})\ (k_m\ k_{m+1} \ldots k_{2m-1}) \ldots]$
- Key k = P A S C A L
- i.e. K = (15, 0 , 18, 2, 0 ,11)
- Plain Text P = “She is listening”

Plain Text	s	h	e	i	s	l	i	s	t	e	n	i	n	g
Mapping	18	7	4	8	18	11	8	18	19	4	13	8	13	6
Key	15	0	18	2	0	11	15	0	18	2	0	11	15	0
CT	7	7	22	10	18	22	23	18	11	6	13	19	2	6
CT	H	H	W	K	S	W	X	S	L	G	N	T	C	G

Vigenere cipher can be seen as combination of m additive cipher.



- Vigenere cipher when m=1 becomes additive/shift cipher
- $K = (k_0, k_1, k_2, ..., K_{m-1})$

$K_0 \epsilon \ Z_{26} \quad |k_0| = 26$

$K_1 \epsilon \ Z_{26} \quad |k_1| = 26$

.

$K_{m-1} \epsilon \ Z_{26} \quad |k_{m-1}| = 26$

Total number of possible keys

= 26 x 26 x ...26 (m times)

$= 26^m$ (Length of the key space)

If m is large, Brute force is impossible.

Cryptanalysis

- Two parts:
- 1. Finding the length of the key
- 2. Finding the key itself
- For 1st, there are several methods, one such method is ‘kasiski test’
  - Cryptanalyst searches for repeated text segments, of at least three characters in the cipher text.
  - Suppose that two of these segments are found and the distance between them is d.

- The cryptanalyst assumes that d / m ,ie. d divides m
- Where m = key length
- If more repeated segments are found with distances (d_1, d_2,...d_n), then take, gcd (d_1, d_2,...d_n) / m
- This assumption is logical because if the two characters are same and are (k x m) (k=1,2,...) characters apart in the plaintext, they are same and (k x m) characters apart from the ciphertext.
- Cryptanalyst uses segments of at least three characters to avoid the cases where the characters in the key are not distinct.

- The index of coincidence (IC) method is used to confirm the m value determined by the kasiski test.
- Definition:
- The index of coincidence of $x = x_1,x_2,...x_n$, which is a string of length n formed by the alphabets A, B, ..., Z is defined as probability that the random elements of x are the same.
- Frequencies of A, B, C,....,Z in x are denoted by the $f_0,f_1,...f_{25}$
- $I_c(x) = \sum_{fi}C_2 \ / \ {}_nC_2$

  $= \sum f_i \times (f_i - 1) / n \times (n - 1) = \sum(f_i/n)^2$