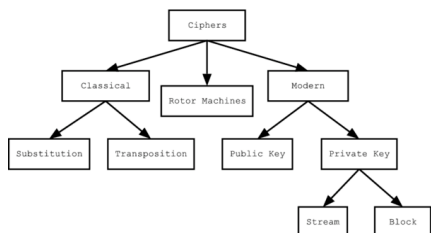


Introduction to Ciphers

Breno de Medeiros

Cipher types



From "Cipher". Wikipedia article. <http://en.wikipedia.org/wiki/Cipher>

Classical ciphers

- Keyword cipher
 - Choose a keyword. We will use "keyword".
 - If the keyword has repeated occurrences of one character, then drop secondary occurrences. For instance, the word "secret" gives keyword "se~~cr~~t".
 - Extend the keyword by all missing alphabet characters (in alphabetic order). Examples:
 - keyword abcdefghijklmnpqstuvxz
 - secretabdfghijklmnopquvwxyz

Keyword continued

- Substitute characters according to the rule:
 - 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
 - k e y w o r d a b c f g h i j l m n p q s t u v x z
- Encrypting "The magic words are squeamish ossifrage":
 - t h e m a g i c w o r d s a r e s q u e a m i s h o s s i f r a g e
 - q a o h k d b y u j n w p k n o p m s o k h b p a j p p b r n k d o

Mono-alphabetic ciphers

- A mono-alphabetic cipher operates by
 - Substituting one character for another
 - The same substitution is applied, irrespective of the character position in the plaintext
- The keyword cipher is an example of a mono-alphabetic cipher. The Caesar cipher is another example:
 - 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
 - d e f g h i j k l m n o p q r s t u v w x y z a b c

Alphabet permutations

- The English alphabet has 26 characters. The number of different mono-alphabetic ciphers equals the number of different permutations of the alphabet
 - $26! \approx 4.03291461127e+26$
 - This number is 89 bits long (or about 11 bytes)
- Could one assume that, if the specific permutation is not known, that such ciphers are safe?

Exhaustive search times

Table 2.2 Average Time Required for Exhaustive Key Search

Key Size (bits)	Number of Alternative Keys	Time required at 1 encryption/ μ s	Time required at 10^6 encryptions/ μ s
32	$2^{32} = 4.3 \times 10^9$	$2^{31} \mu$ s = 35.8 minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	$2^{55} \mu$ s = 1142 years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	$2^{127} \mu$ s = 5.4×10^{34} years	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	$2^{167} \mu$ s = 5.9×10^{46} years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	$2 \times 10^{26} \mu$ s = 6.4×10^{12} years	6.4×10^6 years

Figure from textbook: Stallings. Cryptography and Network Security

Cryptanalysis

- Unlike suggested by exhaustive search times, mono-alphabetic substitution ciphers are not hard to cryptanalyze (recover the key/plaintext)
- The reason is that a natural language message contains a lot of structure that is not present in a random string of characters:
 - Character frequency
 - Sequence of character frequency

English character frequency

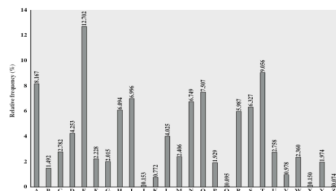


Figure 2.5 Relative Frequency of Letters in English Text

Figure from textbook: Stallings. Cryptography and Network Security

Multi-letter ciphers

- Multi-letter ciphers work by substituting a group of letters (2, 3 or more at a time) by another group of letters (usually the same length)
 - The Playfair cipher uses square diagrams to substitute digrams of the plaintext
 - The Hill Cipher uses matrix operations to substitute letter sequences, n at a time, where n is a parameter of the cipher.

Playfair cipher

- A Keyword is chosen without repeated characters, say we have chosen “Cryptoquiz”

C	R	Y	P	T
O	Q	U	I/J	Z
A	B	D	E	F
G	H	K	L	M
N	S	V	W	X

- To encrypt, split the word into digrams. Use fill letter for repeated characters in the same digram (say 'x'):
- Monkey → MO NK EY
- Collect → CO LX LE CT

- CO encrypts as 'OA'
- CT encrypts as 'RC'
- MO encrypts as 'GZ'

Hill Cipher

- Takes n successive letters. Each letter of the English alphabet is assigned a value:
 - $a = 0, b = 1, c = 2, \dots, y = 24, z = 25$
- A set of linear equations is used to define the encryption using modular arithmetic.
- Structure (3 characters at a time):
 - $C_1 = K_{1,1} P_1 + K_{1,2} P_2 + K_{1,3} P_3 \pmod{26}$
 - $C_2 = K_{2,1} P_1 + K_{2,2} P_2 + K_{2,3} P_3 \pmod{26}$
 - $C_3 = K_{3,1} P_1 + K_{3,2} P_2 + K_{3,3} P_3 \pmod{26}$

Example of Hill Cipher

$$K = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

Plaintext = "paymoremoney"

First 3 letters of plaintext = (p, a, y) = (15 0 24)

First 3 letters of ciphertext = (l, n, s) = (11 13 18)

- $11 = 17 * 15 + 17 * 0 + 5 * 24 \text{ mod } 26$
- $13 = 21 * 15 + 18 * 0 + 21 * 24 \text{ mod } 26$
- $18 = 2 * 15 + 2 * 0 + 19 * 24 \text{ mod } 26$

Hill Cipher characteristics

- Using several characters at a time makes it difficult to collect statistics of ciphertext distribution
 - More ciphertext is required
- There is a large number of keys to choose from.
 - All the invertible matrices of dimension $n \times n$, and entries in $\{0, 1, 2, \dots, 25\}$.

How to decrypt?

- The receiver knows the key, and can solve the system of equations for unknowns P_i :
 - $C_1 = K_{1,1} P_1 + K_{1,2} P_2 + K_{1,3} P_3 \text{ mod } 26$
 - $C_2 = K_{2,1} P_1 + K_{2,2} P_2 + K_{2,3} P_3 \text{ mod } 26$
 - $C_3 = K_{3,1} P_1 + K_{3,2} P_2 + K_{3,3} P_3 \text{ mod } 26$
- This system has solutions iff the matrix K is invertible. Let M be its inverse. Then:
 - $P_1 = M_{1,1} C_1 + M_{1,2} C_2 + M_{1,3} C_3 \text{ mod } 26$
 - $P_2 = M_{2,1} C_1 + M_{2,2} C_2 + M_{2,3} C_3 \text{ mod } 26$
 - $P_3 = M_{3,1} C_1 + M_{3,2} C_2 + M_{3,3} C_3 \text{ mod } 26$

How to break the Hill Cipher?

- Suppose the cryptanalyst can find a bit of text that has been encrypted with the Hill Cipher:
 - A pair (P, C) is available of plaintext and corresponding ciphertext, of some length.
- Use several systems to solve for M (need about n^2 encrypted characters):
 - $P_1 = M_{1,1} C_1 + M_{1,2} C_2 + M_{1,3} C_3 \pmod{26}$
 - $P_2 = M_{2,1} C_1 + M_{2,2} C_2 + M_{2,3} C_3 \pmod{26}$
 - $P_3 = M_{3,1} C_1 + M_{3,2} C_2 + M_{3,3} C_3 \pmod{26}$
- Here 9 characters (3 groups) are needed.

Example:

- Suppose “friday” = (5,17,8,3,0,24) encrypts to “pqcfku” = (15,16,2,5,10,20). It is known that the Hill Cipher is 2×2 matrix.
 - (5, 17) goes to (15, 16) and (8, 3) to (2, 5).
 - Write the equations:
 - $5 = M_{1,1} \times 15 + M_{1,2} \times 16 \pmod{26}$
 - $17 = M_{2,1} \times 15 + M_{2,2} \times 16 \pmod{26}$
 - $8 = M_{1,1} \times 8 + M_{1,2} \times 3 \pmod{26}$
 - $3 = M_{2,1} \times 8 + M_{2,2} \times 3 \pmod{26}$
- $$M = \begin{pmatrix} 23 & 8 \\ 19 & 19 \end{pmatrix}$$
