

Lecture Note 1

Introduction to Classical Cryptography

Sourav Mukhopadhyay

CRYPTOGRAPHY AND NETWORK SECURITY - MA61027

- Cryptography is the science or art of secret writing.
- The fundamental objective of cryptography is to enable two people (Alice and Bob) to communicate over an insecure channel in such a way that an opponent (Oscar) cannot understand what is being said.

- Plaintext : the information that Alice wants to send to Bob.
- Alice encrypts the plaintext, using a predetermined key, and send the resulting ciphertext to Bob over the public channel.
- Upon receiving the ciphertext
 - Oscar cannot determine what the plaintext was
 - But Bob knows the encryption key, can decrypt the ciphertext and get the plaintext.

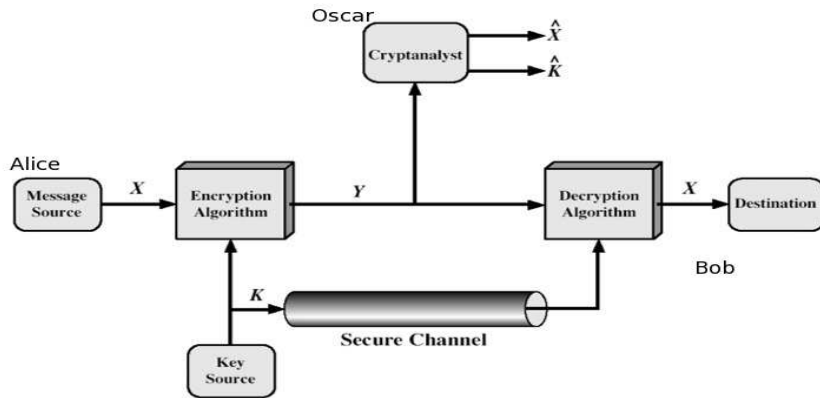


Figure 1: Communication Channel.

- Cryptology - two competing areas:
 - Cryptography - Art of converting information to a form that will be unintelligible to an unintended recipient, carried out by **cryptographer**.
 - Cryptanalysis - Art of breaking cryptographic systems, carried out by **cryptanalyst**.
- Two main types of cryptography in use today:
 - **Symmetric** or **secret key** cryptography
 - **Asymmetric** or **public key** cryptography

Conventional Encryption

- Also termed single-key or symmetric encryption

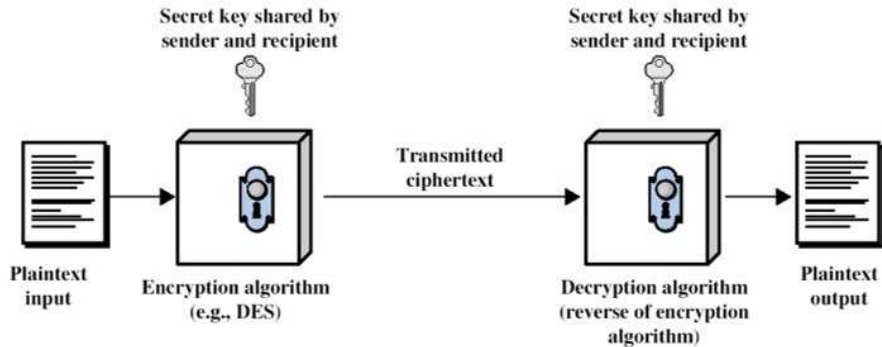


Figure 2: Simplified model of conventional encryption.

Cryptosystem

- Cryptosystem is a five tuple $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$
 - Plaintext Space (\mathcal{P}): set of all possible plaintexts
 - Cipherext Space (\mathcal{C}): set of all possible ciphertexts
 - Key Space (\mathcal{K}): set of all possible keys
 - \mathcal{E} : set of all possible encryption rules and \mathcal{D} : set of all possible decryption rules
- For each $k \in \mathcal{K}$, there is an encryption rule $e_k \in \mathcal{E}$ and a corresponding decryption rule $d_k \in \mathcal{D}$ such that $d_k(e_k(x)) = x$ for every plaintext $x \in \mathcal{P}$

- A practical cryptosystem should satisfy
 - Each encryption function e_k and each decryption function d_k should be efficiently computable.
 - An opponent, upon seeing the ciphertext string y , should be unable to determine the key k that was used, or the plaintext string x

- The process of attempting to compute the key k , given a string of ciphertext y , is called **cryptanalysis**.
 - If opponent can determine k , then he can decrypt y just as Bob would, using d_k .
 - Determining k should be as difficult as determining the plaintext string x , given the ciphertext string y .

Shift Cipher

- $Z_{26} = \{0, 1, 2, \dots, 24, 25\}$
- $\mathcal{P} = \mathcal{C} = \mathcal{K} = Z_{26}$
- For $k \in \mathcal{K}$,
$$e_k(x) = (x + k) \bmod 26 \text{ for } x \in \mathcal{P}$$
$$d_k(y) = (y - k) \bmod 26 \text{ for } y \in \mathcal{C}$$
- Caesar Cipher is a particular case (for $k = 3$)

Example

- Plaintext is ordinary English text
- Correspondence between alphabetic characters and integer: $A = 0, B = 1, \dots, Y = 24, Z = 25$.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Q	R	S	T	U	V	W	X	Y	Z
16	17	18	19	20	21	22	23	24	25

Encryption

- key $k = 11$
- Plaintext is “wewillmeetatmidnight”
- corresponding sequence of integers:
22 4 22 8 11 11 12 4 4 19 0 19 12 8 3 13 8 6 7 19.
- we add 11 (key) to each value (reducing modulo 26):
7 15 7 19 22 22 23 15 15 4 11 4 23 19 14 24 19 17 18 4.
- convert the sequence of integers to alphabetic characters:
Ciphertext is “HPHTWWXPPELEXTOYTRSE”

Decryption

- ciphertext : “HPHTWWXPPELEXTOYTRSE”.
- convert the ciphertext to sequence of integers:
7 15 7 19 22 22 23 15 15 4 11 4 23 19 14 24 19 17 18 4.
- subtract 11 from each value (reducing modulo 26):
22 4 22 8 11 11 12 4 4 19 0 19 12 8 3 13 8 6 7 19.
- convert the sequence of integers to alphabetic characters:
Plaintext is “wewillmeetatmidnight”

Caesar Cipher

- Caesar Cipher is the earliest known (and the simplest). It involves replacing each letter of the alphabet with the letter standing three places further down. This is then wrapped around on itself when the end is reached. For example:

Key: $k=3$

Plaintext: meetmeaftertheparty

Ciphertext: PHHWPHDIWHUWKHSDUWB

Shift Cipher is not Secure

- **Brute-force** cryptanalysis easily performed on the shift cipher by trying all 25 possible keys.
- Given a ciphertext string, Oscar successively try the decryption process with $k = 0, 1, 2, \text{ etc.}$ until get a meaningful text.

- Ciphertext : JBCRCLQRWCVRNBJENBWRWN

$k = 0 \rightarrow$ jbcrcqlqrwcrvnbjenbwrwn

$k = 1 \rightarrow$ iabqbkpqbqumaidmavqvm

$k = 2 \rightarrow$ hzapajopuaptlzhclzupul

$k = 3 \rightarrow$ gyzozinotzoskygbkylotk

$k = 4 \rightarrow$ fxynyhmnsynrjxfajxsnsj

$k = 5 \rightarrow$ ewxmxmlrmxmqiweziwrmri

$k = 6 \rightarrow$ dvwlfklqwlphvdyhvqlqh

$k = 7 \rightarrow$ cuvkvkpvkkojucxjupkpg

$k = 8 \rightarrow$ btujudijoujnftbwftojof

$k = 9 \rightarrow$ astitchintimesavesnine

- The key is $k = 9$

Substitution Cipher

- $\mathcal{P} = \mathcal{C}$ = set of 26-letter English alphabet

$$\mathcal{P} = \{a, b, c, \dots, y, z\}$$

$$\mathcal{C} = \{A, B, C, \dots, Y, Z\}$$

- \mathcal{K} = set of all possible permutations of 26 alphabetic characters.
- For each permutation $\phi \in \mathcal{K}$,

$$e_{\phi}(x) = \phi(x) \text{ for } x \in \mathcal{P}$$

$$d_{\phi}(y) = \phi^{-1}(y) \text{ for } y \in \mathcal{C}, \text{ where } \phi^{-1} \text{ is the inverse permutation of } \phi.$$

Example

- Encryption function is the permutation ϕ :

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
X	N	Y	A	H	P	O	G	Z	Q	W	B	T	S	F	L
q	r	s	t	u	v	w	x	y	z						
R	C	V	M	U	E	K	J	D	I						

- Decryption function is the inverse permutation ϕ^{-1} :

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
d	l	r	y	v	o	h	e	z	x	w	p	t	b	g	f
Q	R	S	T	U	V	W	X	Y	Z						
j	q	n	m	u	s	k	a	c	i						

- **Key:** $k = \phi$
- **Ciphertext:**
MGZVYZLGHCMHJMYXSNHAHYCDLMHA
- Find the plaintext???

- **Monoalphabetic Cipher:** Each alphabetic character is mapped to a unique alphabetic character
- We use arbitrary monoalphabetic substitution, so there are $26!$ or $4 \times 10^{26} \approx 2^{88}$ possible permutations, which is a very large number. Thus brute force is infeasible.
- However we will see later that a Substitution Cipher is insecure against frequency analysis.

Vigenère Cipher

- **Polyalphabetic cipher:** use different monoalphabetic substitutions while moving through the plaintext.
- Let m be a positive integer
- $\mathcal{P} = \mathcal{C} = \mathcal{K} = (\mathbb{Z}_{26})^m$
- For $k = (k_1, k_2, \dots, k_m) \in \mathcal{K}$,
$$e_k(x_1, x_2, \dots, x_m) = (x_1 + k_1, x_2 + k_2, \dots, x_m + k_m)$$
$$d_k(y_1, y_2, \dots, y_m) = (y_1 - k_1, y_2 - k_2, \dots, y_m - k_m)$$
- All above operations are performed in \mathbb{Z}_{26}

Example

- Correspondence between alphabetic characters and integer: $A = 0, B = 1, \dots, Y = 24, Z = 25$.
- $m = 6$.
- Keyword is “CIPHER”, this corresponds to the numerical equivalent $k = (2, 8, 15, 7, 4, 17)$

- Plaintext : “thiscryptosystemisnotsecure”.
- Encryption: add modulo 26

19	7	8	18	2	17	24	15	19	14	18	24	18	19	4	12
2	8	15	7	4	17	2	8	15	7	4	17	2	8	15	7
21	15	23	25	6	8	0	23	8	21	22	15	20	1	19	19
8	18	13	14	19	18	4	2	20	17	4					
4	17	2	8	15	7	4	17	2	8	15					
12	9	15	22	8	25	8	19	22	25	19					

- Ciphertext:
“VPXZGIAXIVWPUBTTMJPWIZITWZT”.

- **Transposition techniques:** So far all the ciphers we have looked at involved only substitution. A very different kind of mapping is achieved using transposition.
- In its simplest form, the **rail fence** technique involves writing down the plaintext as a sequence of columns and the ciphertext is read off as a sequence of rows. For example, if we use a rail fence of depth 2 with the plaintext *meet me after the party is over* we get:

m	e	m	a	t	r	h	p	r	y	s	v	r
e	t	e	f	e	t	e	a	t	i	o	e	

- Ciphertext is *mematrhp r ysvretefeteatioe* which is simply the first row concatenated with the second.

Transposition/Permutation Cipher

- Let m be a positive integer
- $\mathcal{P} = \mathcal{C} = (\mathbb{Z}_{26})^m$
- \mathcal{K} = set of all possible permutations of $\{1, 2, \dots, m\}$
- For each permutation $\pi \in \mathcal{K}$,
$$e_{\pi}(x_1, x_2, \dots, x_m) = (x_{\pi(1)}, x_{\pi(2)}, \dots, x_{\pi(m)})$$
$$d_{\pi}(y_1, y_2, \dots, y_m) = (y_{\pi^{-1}(1)}, y_{\pi^{-1}(2)}, \dots, y_{\pi^{-1}(m)})$$
- π^{-1} being the inverse permutation of π

Example

- $m = 6$.
- key is the following permutation π :

$$\begin{array}{c|c|c|c|c|c} x & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline \pi(x) & 3 & 5 & 1 & 6 & 4 & 2 \end{array}$$

- inverse permutation π^{-1} :

$$\begin{array}{c|c|c|c|c|c} x & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline \pi^{-1}(x) & 3 & 6 & 1 & 5 & 2 & 4 \end{array}$$

- Plaintext : “defendthehilltopatsunset”

- partition the plaintext into group of six letters:

defend | thehil | ltopat | sunset

- rearrange according to π :

fnddee | eitlhh | oaltpt | nestsu

- Ciphertext: “FNDDEEEITLHHOALTPTNESTSU”
- Decryption can be done using π^{-1}

Cryptanalysis

- **Brute-force** cryptanalysis easily performed on the Shift cipher by trying all 25 possible keys.
- Three characteristics of the problem facilitate the successful use of the brute force approach:
 1. The encryption scheme is known.
 2. There are only a limited no. of keys.
 3. The plaintext is easily recognisable.
- Most cases, key size tends to be the main problem for brute-force attacks.

- **Monoalphabetic Ciphers:** If instead of using only the 25 possible keys, arbitrary substitution is used as in Substitution cipher, then there are $26!$ or $4 \times 10^{26} \approx 2^{88}$ (10 orders of magnitude greater than the keyspace for DES!) possible keys and hence brute force is infeasible.
- We now show the frequency analysis on Substitution cipher

Frequency Analysis

- Suppose we have a long ciphertext, the challenge is to decipher it
- Let we know the text is in English and has been encrypted using a monoalphabetic substitution cipher
- searching all possible keys is impractical as the keyspace is of size $26!$

- In English, *e* is the most common letter, followed by *t*, then *a*, and so on, as shown in the Figure 3

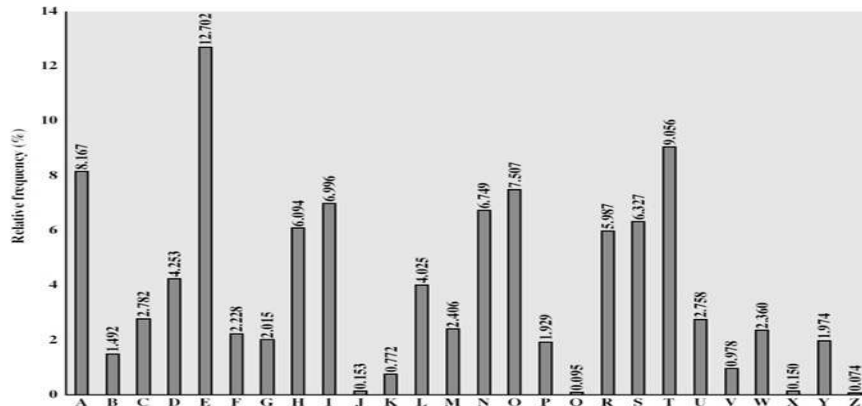


Figure 3: Relative Frequency of letters in the English Language.

- examine the ciphertext in question, and work out the frequency of each letter
- if most common letter in the ciphertext is, for example, J then it would seem likely that this is a substitution for e
- if the second most common letter in the ciphertext is P , then this is probably a substitution for t , and so on
- however, regularities of the language may be exploited, e.g. relative frequency
- frequency analysis requires logical thinking, intuition, flexibility and guesswork

Playfair Cipher

- Use the keyword CHARLES (Charles Wheatstone invented the cipher).
- Draw up a 5×5 matrix with the keyword first, removing any repeating letters as follows:

c	h	a	r	l
e	s	b	d	f
g	i/j	k	m	n
o	p	q	t	u
v	w	x	y	z

- Plaintext: “meet me at the bridge”.

- Split the sentence into digrams removing spaces, ‘x’ used to make even number of letters:

me et me at th eb ri dg ex

- Repeating plaintext letters that are in the same pair are separated with a filler letter, such as x:

“balloon” would be treated as ba lx lo on

- Two plaintext letters in the same row are each replaced by the letter to the right, with the first element of the row circularly following the last.

eb is replaced by sd

ng is replaced by gi (or gj as preferred)

- Two plaintext letters that fall in the same column are each replaced by the letter beneath, with the top element of the column circularly following the last.

dt would be replaced by my

ty would be replaced by yr

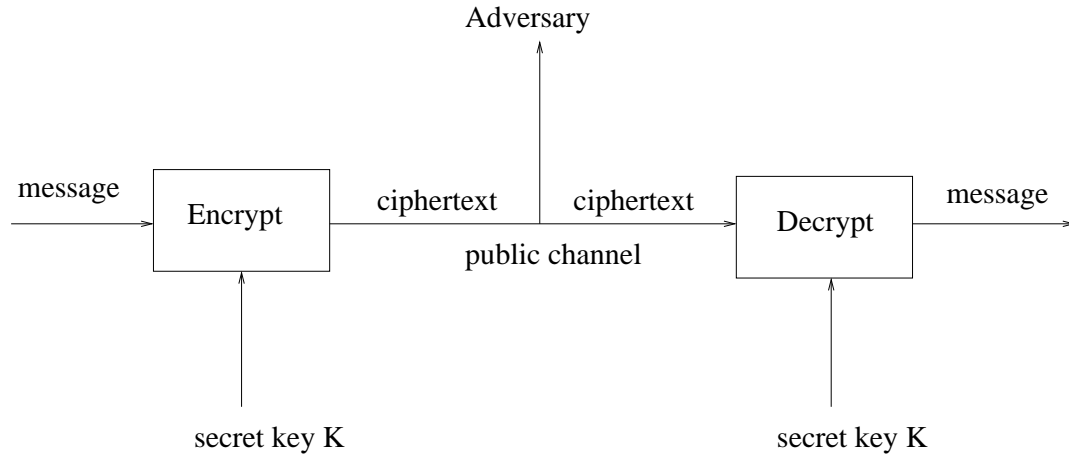
- Otherwise, each plaintext letter in a pair is replaced by the letter that lies in its own row and the column occupied by the other plaintext letter.

me becomes gd

- Ciphertext therefore is :

“gd do gd rq pr sd hm em bv”

Symmetric Key Encryption



- Block ciphers and stream ciphers are two types of symmetric key cryptosystems

Block Cipher

- plaintext string $x = x_1x_2x_3, \dots$
- Successive plaintext elements are encrypted using the same key k :

Key:	k			
Plaintext:	x_1	x_2	x_3	\dots
Ciphertext:	$e_k(x_1)$	$e_k(x_2)$	$e_k(x_3)$	\dots

- ciphertext string $y = y_1y_2y_3, \dots = e_k(x_1)e_k(x_2)e_k(x_3) \dots$
- **Examples:** DES, Rijndael (The AES), IDEA, RC6, and many more....

Stream Cipher

- plaintext string $x = x_1x_2x_3, \dots$
- generate a keystream $k_1, k_2, k_3 \dots$ from the key k :

Key:	k			
keystream:	k_1	k_2	k_3	\dots
Plaintext:	x_1	x_2	x_3	\dots
Ciphertext:	$e_{k_1}(x_1)$	$e_{k_2}(x_2)$	$e_{k_3}(x_3)$	\dots

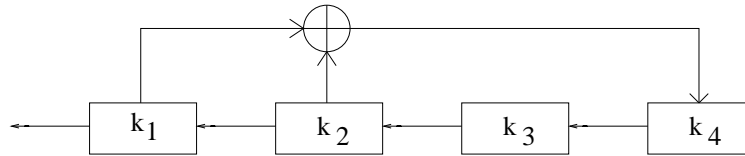
- ciphertext string
 $y = y_1y_2y_3, \dots = e_{k_1}(x_1)e_{k_2}(x_2)e_{k_3}(x_3) \dots$

Example

- Consider message to be a bit stream m_1, m_2, \dots
- Let k_1, k_2, \dots be a sequence of pseudorandom bits.
- Encrypt: $c_i = m_i \oplus k_i$.
- The cipher is c_1, c_2, \dots
- Decrypt: $m_i = c_i \oplus k_i$

- Security depends upon the sequence k_1, k_2, \dots
- If k_i is a true random sequence, then the cipher is called an one-time pad.
- One-time pad possesses *perfect secrecy*.
- One-time pads are impractical.
- Use a pseudorandom generator. Secret key of the system is the “seed” of the pseudorandom generator.

Linear Feedback Shift Register (LFSR)



- An LFSR of length m consists of m stages numbered $1, 2, \dots, m$, each storing one bit and having one input and one output; together with a clock which controls the movement of data.
- The vector (k_1, k_2, \dots, k_m) would be used to initialize the shift register

- During each unit of time the following operations would be performed concurrently
 - (i) k_1 would be tapped as the next keystream bit
 - (ii) k_2, \dots, k_m would each be shifted one stage to the left
 - (iii) the “new” value of k_m would be computed to be

$$\sum_{j=1}^{m-1} c_j k_{j+1}$$

the linear feedback is carried out by tapping certain stages of the register (as specified by the constants c_j having the value “1”) and computing a sum modulo 2 (which is an exclusive-or).

Cryptographic Security

- **Kerckhoff's Principle:** Assume that the adversary knows the algorithm that is used. The secret is *only* the secret key.
- **Attack Models:**
 - Ciphertext only attack: The opponent possesses a string of ciphertext, y
 - Known plaintext attack: The opponent possesses a string of plaintext, x , and the corresponding ciphertext, y

- Chosen plaintext attack: The opponent has obtained temporary access to the encryption machinery. Hence he can choose a plaintext string, x , and construct the corresponding ciphertext string, y .
- Chosen ciphertext attack: The opponent has obtained temporary access to the decryption machinery. Hence he can choose a ciphertext string, y , and construct the corresponding plaintext string, x .

- **Adversarial Goal:**

- Key recovery.
- Distinguishing attack.
- Malleability.
- Other application specific security goals.