ISSN No. 0976-5697

Volume 5, No. 6, July-August 2014

International Journal of Advanced Research in Computer Science

TECHNICAL NOTE

Available Online at www.ijarcs.info

Codes And Ciphers

Dr. Saroj Singh
Dept. Computer Science & Engineering
Applied College of Management & Engineering
Mitrol, Palwal, India

Abstract: Code is higher level substitution that works at the level of words where as cipher is lower level substitution that works with at the level of letters. In communications and information processing, code is system of rules to convert information—such as a letter, word, sound, image, or gesture - into another, sometimes shortened or secret, form or representation for communication through a channel or storage in a medium. In cryptography, a cipher (or cypher) is an algorithm for performing encryption or decryption - a series of well-defined steps that can be followed as a procedure. Ciphers: a secret or disguised way of writing; a code. "we will meet today" code, secret writing;

Keywords: **c**ommunication; information; codes; cipher; processing; shortened; secret; cryptography; decryption; encryption; nucleotides;

I. INTRODUCTION

Code is higher level substitution that works at the level of words where as cipher is lower level substitution that works with at the level of letters. In communications and information processing, code is system of rules to convert information—such as a letter, word, sound, image, or gesture[1]—into another, sometimes shortened or secret, form or representation for communication through a channel or storage in a medium. The main disadvantages of codes are:

- Show to write
- Code book is required
- Vulnerable since the code has to be stored somewhere

A **code C : S -> T*** is a total function mapping each symbol from S to a sequence of symbols over T, and the extension of C to a homomorphism of S into T*, which naturally maps each sequence of source symbols to a sequence of target symbols, is referred to as its **extension**.

A 'cipher' is the same thing as a 'code'; however, the concepts are distinct in cryptography. In classical cryptography, ciphers were distinguished from codes.

II. TYPES OF COEDS AND CIPHERS

A. Code:codes can be classified as:

- **Variable-length code**: Variable-length codes are especially useful when clear text characters have different probabilities.
- **A prefix code:** is a code with the "prefix property": there is no valid code word in the system that is a prefix (start) of any other valid code word in the set. Huffman coding is the most known algorithm for deriving prefix codes. Other examples of prefix codes are country calling codes
- **Error detection and correction:** Codes may also be used to represent data in a way more resistant to errors in transmission or storage. Such a "code" is called an error-correcting code, and works by including carefully crafted redundancy with the stored (or transmitted) data. Examples include Hamming codes, Reed–Solomon, Reed–Muller, Walsh–Hadamard, Bose–Chaudhuri–Hochquenghem, Turbo, Golay, Goppa, low-density parity-check codes, and space–time codes. Error detecting codes can be optimised to detect *burst errors*, or *random errors*.
- **Character encoding:** In use today is ASCII. In one or another version, it is used by nearly all personal computers, terminals, printers, and other communication equipment. It represents 128 characters with seven-bit binary numbers—that is, as a string of seven 1s and 0s (bits). In ASCII a lowercase "a" is always 1100001, an uppercase "A" always 1000001, and so on.
- **Genetic code:** Biological organisms contain genetic material that is used to control their function and development. This is DNA which contains units named genes that can produce proteins through a code (genetic code) in which a series of triplets (codons) of four possible nucleotides are translated into one of twenty possible amino acids. A sequence of codons results in a corresponding sequence of amino acids that form a protein.
- **Secret codes:** Secret codes intended to obscure the real messages, ranging from serious (mainly espionage in military, diplomatic, business, etc.) to trivial (romance, games) can be any kind of imaginative encoding: flowers, game cards, clothes, fans, hats, melodies, birds, etc., in which the sole requisite is the previous agreement of the meaning by both the sender and the receiver.
- **Gödel code:** In mathematics, a Gödel code was the basis for the proof of Gödel's incompleteness theorem. Here, the idea was to map mathematical notation to a natural number (using a Gödel numbering-natural number).

B. ciphers:ciphers can be classified as:

- **Substitution ciphers**

Features are[2]:

a) Simple
b) Substitute one letter for another

c) Can be broken easily by analyzing the frequency of letters in ciphertext.

- **Reciprocal ciphers**

Features are:

a) When one enters the paintext to obtain the ciphertext, one can enter the ciphertext at the same place to obtain the plaintext.
b) Makes the system usable
c) Difficult to break

- **Symmetric ciphers**

Features are:

a) Uses the same key for both encryption and decryption process.Simple

- **Asymmetric ciphers**

Features are:

a) Also known as split key algorithm
b) Came into existence in 1975.
c) Different key is used for both encryption and decryption process.
d) Provides authentication

III. TRADITIONAL CRYPTOGRAPHY

Definition: in traditional cryptography, designing of algorithms is performed but no computation is done[3]. The two techniques which come under traditional cryptography are substitution and transposition techniques.

A. Substitutiona Techniques:

Definition: substitution technique is a technique in which letters of plaintext are replaced by another letters, symbols or numbers. The plaintext bit patterns are replaced with cipher text bit patterns.

- **Caesar cipher:** Caesar cipher is the earliest known example of substitution technique. This was given by Julius Caesar. Brute force cryptanalysis is easily performed on Caesar cipher[4].

Characteristics:

a) Only 25 key to try
b) Encryption and decryption algorithm are known.
c) Plaint text easy to recognize.

Technique: Each letter of the alphabet is replaced with the letter standing three paces down the alphabet.

Plain text	a	b	c	d	e	f	g	h	i	j	k	l	…… ….
Cipher	d	e	f	g	h	i	j	k	l	m	n	o	…… …

For example:
Plaintext: are you ready
Cipher: duh brx uhdgh
Numerical equivalence of each letter:

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12

n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

Encryption:
C = E (3, p) = (p + 3) mod 26
C = E (k, p) = (p + k) mod 26

Decryption:
P = D (k, C) = (C - K) mod 26
Where k is secret key.

- **Monoalphabetic ciphers:**

Definition: In monoalphabetic ciphers, substitution characters are randomly permuted by 26 letters of alphabet[5].

Characteristics:

1. Used when there are 26! Or greater 4*1026 possible keys.
2. Eliminates brute force technique for cryptanalysis.
3. Single cipher alphabet is used per message.

Plain text	a	b	c	d	….
Cipher	t	h	i	f	….

- **Playfair ciphers:**

Definition: Playfair cipher is the best multiple letter encryption cipher. The diagrams in the plaintext are considered as one single unit. These units are then converted into ciphertext diagrams.

Basically playfair cipher uses 5*5 matrix of letters which are constructed by using a keywords.

Construction of matrix : matrix constructed by filling the letters of the keyword(minus duplicates) from left to write and top to bottom manner.

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

Since there are 676 diagrams so identification of individual diagrams is more difficult.frequency analysis is difficult.

- **Hill cipher**

Hill cipher was developed by lester hill in 1929.
Techniques:

a) ‘m’ successive plaintext letters are substituted by’m’ciphertext letters.

b) Each charcter is assigned a numerical value (a=0.b=1, ……c=25) and the substitution is determined by 'm' linear equation.
c) Operations are performed using mod 26.
d) Decryption requires the inverse of matrix k.

The encrypting key consists of a 3*3 matrix[6]:

$$k = \begin{bmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{bmatrix}$$

For m=3, the system can be described as:

$C_1 = (k_{11} P_1 + K_{12} P_2 + K_{13} P_3) \bmod 26$
$C_2 = (k_{21} P_1 + K_{22} P_2 + K_{23} P_3) \bmod 26$
$C_1 = (k_{31} P_1 + K_{32} P_2 + K_{33} P_3) \bmod 26$

Here C and P are column vectors representing plaintext and ciphertext of length 3.
K is 3*3 matrix.

- **Polyalphabetic ciphers:**

Definition: In polyalphabetic ciphers, substitution rules are different for each letter. The shifting occurs according to the elements of cryptographic keys.

Each cipher is denoted by a key letter. The ciphertext letter substitutes the plaintext letter. For example: Vigenere cipher

Vigenere cipher was named after Blaise de Vigenere but was originally described by Giovan Battista Bellaso. This cipher consists of 26 ciphers. It is a method of encrypting alphabetic text by using the 26 caesar ciphers. These ciphers are based upon the letters of a keyword[7].

Technique:

1. Vigenere square or table is used for encryption and decryption purpose.
2. The table consists of alphabets which are written out 26 times in different rows.
3. Each alphabet is shifted to the left compared to the previous alphabet.

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

- Algebraic format
- Letters A-z are taken as numbers 0-25 and addition is performed mod 26.

Encryption
$C_i = E_k (M_i) = (M_i + K_i) \bmod 26$
Decryption
$M_i = D_k (C_i) = (C_i + K_i) \bmod 26$
Where
$M = M_0 \ldots\ldots\ldots\ldots\ldots M_n$
$C = C_0 \ldots\ldots\ldots\ldots\ldots C_n$
$K = K_0 \ldots\ldots\ldots\ldots\ldots K_n$

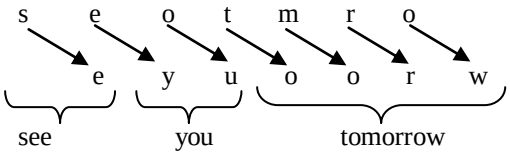
B. Transpositional Techniques

Definition: Transposition is also known as permutation and is basically the reordering of appearance of plaintext elements. The pure transition cipher is easily recognized because it has same letter frequencies as the original plaintext.

Mapping is achieved by performing permutation.

Technique followed:

- **Rail fence Technique:** Plaintext is written down as sequence of diagonals and is then read as sequence of rows[8]. For example: "see you tomorrow" is written as:



- **Route cipher:**The plaintext is written in a grid of given dimensions and is read as a pattern which is given in the key. Route ciphers have many keys than rail fence technique.

C. Column Transposition:

The message is written in row that are fixed lenthe and is read column by column. The width of rows and permutation of the columns are defined by a keyword.

In regular column transposition spare spaces are filled with nulls and in irregular column transposition the spare spaces are left blank.

For example: the word "PAINT" is of length 6 and the permutation is defined by the alphabetically order of the letters in the keyword. In this case the order would be "6 3 2 4 1 5".

Now suppose the keyword is PAINTS and the message is "Meet me after the party". So, in regular column transposition, it is written as follows:

(6)	(3)	(2)	(4)	(1)	(5)
m	e	e	t	m	e
a	f	t	e	r	t
h	e	p	a	r	t
y	m	p	t	n	u

mptnu are the null values at the end.

Ciphertext

mrrn is the value of (1)
etpp is the value of (2)
efem is the value of (3)
teat is the value of (4)
ettu is the value of (5)
mahy is the value of (6)

Thus the ciphertext is **mrrn etpp efem teat ettu mahy**

So in an irregular column transposition, message we are discovered fliiat once. It is written as follows:

(6)	(3)	(2)	(4)	(1)	(5)
w	e	a	r	e	d
i	s	c	o	v	e
r	e	d	f	l	e
e	a	t	o	n	c
e					

cipherext is written as
evlna cdtes earof odeec wiree.

IV. TRADITIONAL CRYPTOGRAPHY

Modern Cryptography follows a strong scientific approach and can be classified as:

A. Symmetric Key Cryptography

Definition: It is a cryptographic method where a single key is used for both encryption and decryption process[9].
Components of symmetric cipher model:

- Plaintext
- Encryption-decryption algorithm
- Key
- Ciphertext

Properties to be exhibited:

- Security of key distribution
- Adequate streth of encryption process
- Number of internal algorithms
- Choice of operators used in encryption and decryption process
- Choice of key space

Types of symmetric key algorithms:

- AES (Advanced Encryption Standard)
- CAST5
- Triple DES
- IDEA

B. Public Key Cryptography

Definition: It is cryptographic method where separate keys are used for encryption and decryption process[10]. This approach makes use of asymmetric key algorithms.
Technique:

- Each user has a pair of cryptographic keys that is a public encryption and a private decryption key.
- Messages are encrypted with the receiver's public key and can be decrypted using the corresponding private key.

Branches of public key cryptography:

- Public key encryption: a message that is encrypted using the receiver's public key can only be decrypted with a machine private key this is used for confidentiality purpose.
- Digital Signatures: A message that is signed with a sender's private key can be verified by anyone who has to the sender's public key.

V. REFERENCES

[1] Kogan, Hadass, "So Why Not 29" American Journalism Review. Retrieved 2012-07-03.

[2] "WESTERN UNION "92 CODE" & WOOD'S "TELEGRAPHIC NUMERALS"". Signal Corps Association. 1996. Retrieved 2012-07-03.I.

[3] Richard J. Aldrich, GCHQ: The Uncensored Story of Britain's Most Secret Intelligence Agency, HarperCollins July 2010.

[4] Helen Fouché Gaines, "Cryptanalysis", Dover. ISBN 0-486-20097-3K. Elissa, "Title of paper if known," unpublished 1939.

[5] Ibrahim A. Al-Kadi, "The origins of cryptology: The Arab contributions", Cryptologia, pp. 97–12616(2) April 1992.

[6] David Kahn, The Codebreakers - The Story of Secret Writing: ISBN 0-684-83130-9, 1967

[7] David A. King, The ciphers of the monks - A forgotten number notation of the Middle Ages, Stuttgart: Franz Steiner, ISBN 3-515-07640-9, 2001.

[8] Abraham Sinkov, Elementary Cryptanalysis: A Mathematical Approach, Mathematical Association of America,. ISBN 0-88385-622-0, 1966.

[9] William Stallings, Cryptography and Network Security, principles and practices, 4th Edition

[10] Stinson, Douglas R Cryptogtaphy / Theory and Practice, CRC Press, ISBN 0-8493-8521-0 ,1995.