A Survey on Cryptography and Steganography Methods for Information Security

Kallam Ravindra Babu
Research Scholar,
JNTUH, Hyderabad, A.P,India

Dr. S.Udaya Kumar
Deputy Director
SNIST, Hyderabad, A.P, India

Dr. A.Vinaya Babu
Director, Admissions
JNTUH, Hyderabad,A.P, India

ABSTRACT
This paper deals with the tidings of cryptography in history, how it has played a vital role in World War -1, World War-2. It also deals with the Substitution, Transposition Cryptographic techniques and Steganography principles which can be used to maintain the confidentiality of computerized and none computerized information files.

A number of well known techniques have been adapted for computer usage including the Ceaser cipher, Mono alphabetic cipher, Homophonic substitution, Bale cipher, Play fair cipher, Poly alphabetic cipher, Vigenere Cipher, One- time pad cipher, Vernam ciphers, Play Color Cipher and usage of rotor machine in Substitutions, Rail fence technique, more complex permutations for more secure transposition and some Steganography principle were briefly discussed with merits and demerits.

Finally it gives the broad knowledge on almost all the cryptographic and Steganography principles where a reader or scholar have lot of scope for updating or invention of more secure algorithms to fulfill the global needs in information security.

General Terms
Security, Algorithms, History, Cipher.

Keywords
Block Cipher, Play Color, Cryptography, Encryption, Decryption, Decillions, Homophonic, Steganography, HSC, PSC, PCC, SIS, Substitution, Transposition.

1. INTRODUCTION
The most powerful and common approaches to countering the threats to network / information security are Encryption [8] and Steganography [1],[6]. Encryption is based on number of substitutions, transpositions we perform on the plaintext, converted plain text is treated as cipher text and a process of converting is called Cryptography [23]. Steganography is a form of covert communication in which a secret message is camouflaged with in a carrier message.

Even though encryption is very powerful among these two, the cryptanalysts are very intelligent and they were working day and night to break the ciphers. To make a stronger cipher it is recommended that to use: More stronger and complicated encryption algorithms with more number of rounds, Keys with more number of bits (Longer keys), secure transmission of keys [14] [20].

2. HISTORY ENLIGHTEN'S
With respect to the analysis of David Khan[4], over 4000 years ago, the son of cheops, who built the great pyramid at Giza in Egypt, placed a cache of documents at the top of the pyramid. Most of the hieroglyphics involved were standard and have been translated, but some appear to be an attempt at secret writing and have yet to be translated. It appears that cryptography has been with us since writing was invented.

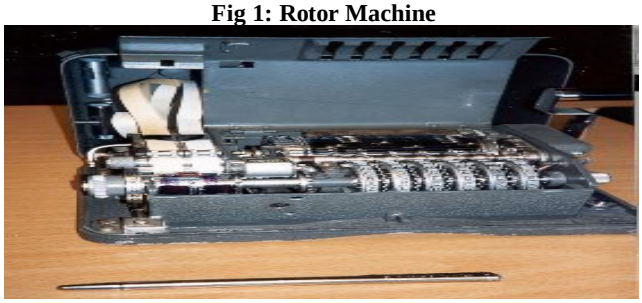
The first general known to have achieved a major victory with the aid of cryptography was Lysander, the leader of the forces from the city state of Sparta in their war against the Athenians (405 BC). The Persians had expressed support for Lysander but he was not certain of their sincerity. One day a slave arrived with a message on parchment. Lysander was waiting anxiously, not for this pretext message, but for the belt of the slave. A seemingly meaningless string of letters was inscribed on the inside of the belt. However, when the belt was wrapped around a button –like sky tale, the letters regrouped into an intelligible message along the sky tale. The message (from an informant) stated that the Persians were lying and that they were actually plotting against Lysander. He sailed against the Persians and the ensuing victory of the Spartans profoundly affected the history of Europe[4].
In 1586, Mary Queen of Scots had been temporarily confirmed by her cousin, Queen Elizabeth of England. Mary involved herself in a plot to the conspirators were encrypted and smuggled in barrels of beer. These messages were intercepted, copied and decrypted by walsingharm, the head of Elizabith's newly formed secret services.

In 1916, the war between Britain and Germany had used the encryption and decryption for secrete communication. In January 1917, British naval intelligence intercepted and decrypted a telegram from the Germany Foreign Minister Zimmerman to his ambassador in Washington.
In August 1940, William F. Friendman, head of the United Sates Signal Intelligence Service (SIS), had succeeded in breaking the cryptographic system used by the Japan. The Japanese system codenamed " Purple", used a mechanical device called a router machine, which had been widely used by business in the 1930's.
Four circular routers revolved relatively to each other, as in the odometer for a car.
The US was decrypting the reading messages between the Japanese foreign office and Japanese ambassadors in Washington just prior to the events at Pearl Harbor on 7th 1941.even though the significance of the messages was not understood at the time.

Though most of World War II, the US continued to intercept and decrypt Japanese massages.

Early in the war, Britain had reconstructed the rotor machine used by the Germans, called the Enigma machine.

Fig 1: Rotor Machine



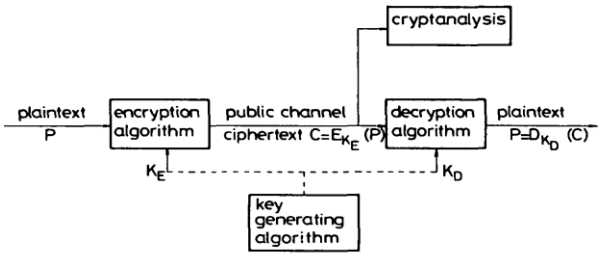
The readers interested in the history of cryptography and its effect on history should read The *code breakers* by David Khan [4].

The period of time from 1945 to the present has been a most crucial period for the science of cryptography, compared to the several thousands years of recorded history. The invention of the computer, the secrecy of the post war intelligence activity, reality of the mass global communication, and the unprecedented demand for the data security in the public sector all combined to mold cryptography in to a true science, away from its artful and mystical past. In the following sections we will attempt to show what cryptography and Steganography has been and what it is now and our innovative approach to wards Cryptography and Security.

3. CRYPTOGRAPHIC SYSTEMS

In his paper F. Ayoub [8] mentioned that, Cryptographic systems are used to provide privacy and authentication in computer and communication systems. As shown in Fig. 2, encryption algorithms encipher the Plaintext, or clear messages, into unintelligible cipher text or cryptograms using a key[8].
A deciphering algorithm is used for decryption or decipherment in order to restore the original information. In general, the enciphering and deciphering keys need not be identical.
Eavesdropping is the interception of messages by a third party monitoring a Communication channel. Anyone trying to break (solve) a cipher is called a cryptanalyst.

Fig 2: General Secrecy System



According to William Stallings [20], Cryptographic systems are generally classified along three independent dimensions:
The type of operations used for transforming plaintext to cipher text: all the encryption algorithms are based on two general principles: Substitution, in this each element in the plaintext (bit, letter, group of bits are letters) is mapped in to another element and the transposition in which elements in the plain text are rearranged. Most systems, referred to as product systems, involve multiple stages of substitutions and transpositions.

The number of keys used: If both sender and receiver use the same key, the system is referred to as symmetric, single key, secret key, or conventional encryption. If the sender and receiver each uses a different key, the system is referred to as asymmetric, two – key, or public key encryption [11], [20].

The way in which the plain text is processed: A block cipher processes the input one block of elements at a time, producing an output block for each input block. A stream cipher processes the input elements continuously, producing out put one element at a time, as it goes along.

Cryptanalysis: The strategy used by the cryptanalyst depends on the nature of the encryption scheme and the information available to the cryptanalyst. Average time required for exhaustive key search [20] is as given in the Table-1.

Table-1: Average Time Required for Exhaustive Key Search

Key Size (bits)	Number of Alternative keys	Time required at 1 Decryption / micro second
32	232=4.3x109	35.8 minutes
56	256=7.2x1016	1142 years
128	2128=3.4x1038	5.4x1024 years
168	2168=3.7x1050	5.9x1036 years

3.1. Substitution ciphers:

In the Substitution Techniques the letters of the plain text are replaced by other letters or by numbers or symbols.
Caesar invented a substitution method in which, each letter of the alphabet is replaced with the letter standing three places further down the alphabet.
Example: Note that the alphabet is wrapped around, so that the letter following Z is A. We can define the transformation by listing all possibilities, as follows:

Plain: 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
Cipher: 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

If the plain text is : meet me after the party
The cipher text is : PHHW PH DCWHU WKH SDUWB

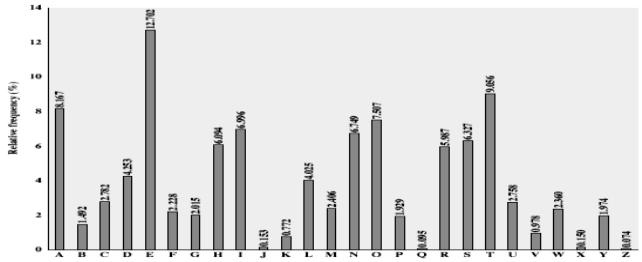
It is observed that, if we assume the algorithm is known, and then the brute force analysis is possible; to overcome this problem it is recommended to use the algorithm with large number of key.

With only 26 possible keys Caesar cipher is far from secure. A dramatic increase in the key space can be achieved by allowing an arbitrary substitution. If, instead, the cipher line can be any permutation of the 26 alphabetic characters, then there are 4 x 1026 possible keys and would seem to eliminate *brute – force* technique for cryptanalysis. Such an approach is referred to as a mono alphabetic substitution cipher.
There is however another type of attack. If the crypt analyst knows the nature of the plain text, then the analyst can exploit the regularities of the language. To see how such a cryptanalysis might proceed, we have a particular example [17]. The cipher text is to be solved is ..

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBME
TSXAIZVUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXU
ZUHSXEPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOH
MQ

As a first step, the relative frequency of the letters can be determined and compared to a standard frequency distribution for English, such as is shown in the fig- 3 [12]. Proportional fonts only for special purposes, such as distinguishing source code text. If Times Roman is not available, try the font named Computer Modern Roman. On a Macintosh, use the font named Times. Right margins should be justified, not ragged.

Fig 3: Relative frequency of letters in English text



It is observe in the above example, the frequency occurrence of cipher letters *P is 13.33 is near to e and Z 11.67 is nearer to t*, but it is not certain which is which. Mono alphabetic ciphers are easy to break because they reflect the frequency data of the original alphabet. Proceeding with trial and error finally gets the plain text for the above cipher text is: it was disclosed yesterday that several informal but direct contacts have been made with political representatives of the viet cong in moscow
A counter measure is to provide multiple substitutes, Known as homophonic substitution cipher(HSC), for a single letters. The great mathematician Carl Friedrich Gauss believed that he had devised an unbreakable cipher using homophones. Bale cipher[18] is an example of Homophonic cipher
Play fair cipher: Another substitution cipher which was popular in World War -I and beyond is named after the Englishman Play fair, but it was actually invented by his friend Wheatstone. Some rearrangement of the alphabet is written in a 5 X 5 square, with two letters equated (usually I and J) as shown in the table-2.

In this case the key word is *monarchy*. The matrix is constructed by filling in the letters of the keyword from left to right and from top to bottom, and the filling in the remainder of the matrix with the remaining letters in alphabetic order. The letters I and J count as one letter. Plain text is encrypted two letters at a time,

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

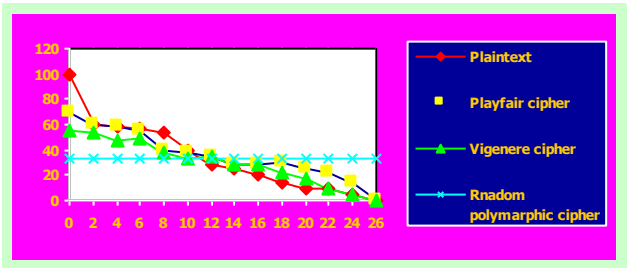
according to the following rules: Repeating plain text letters that would fall in the same pair are separated with a filler letter, such as x, so that balloon would be treated as ba lx lo on.

Table 2: a 5X5 Matrix for Play fair cipher

Plain text letters that fall in the same row of the matrix are each replaced by the letter to the right, with the first element of the row circularly following the last.

For example *ar* is encrypted as RM. Plain text letters that fall in the same column are replaced by the letter beneath, with the top element of the row circularly following the last. For example, *mu* is encrypted as CM. Otherwise, each plain text letter is replaced by the letter that lies in its own row and the column occupied by the other plain text letter. Thus, *hs* becomes BP and ea becomes IM/JM.
One way to revealing the effectiveness [16] of the play fair and other ciphers is shown in fig- 4. The line labeled plaintext plots the frequency distribution of the more then 70,000 alphabetic characters in the Encyclopedia Britannica article on cryptography.

Fig 4: Relative Frequency of Occurrence of Letters



Frequency ranked letters on X -axis

Another way to improve on the simple mono alphabetic cipher is to use different mono alphabetic substitutions as one proceeds through the plain text message.
This well known technique is defined as Poly alphabetic substitution cipher (PSC) used by French Military academy. Most poly alphabetic ciphers are periodic (as are the Vigenare[10] or the Beaufort [2] shifted alphabet ciphers).

It uses a table known as Modern Vigenere Table, in this scheme, the set of related mono alphabetic substitution rules consists of the 26 caser ciphers, with shifts of 0 through 25 as shown in table-3. Each cipher is denoted by a key letter, which is the

cipher text letter that substitutes for the plain text letter. For our example:

Key: deceptivedeceptivedeceptive
Plaintext: wearediscoveredsaveyourself
Cipher text: ZICVTWQNGRZGVTWAVZHCQYGLMGJ

Table:3 The Vigenere table used in Poly Alphabetic Cipher

key \ Plaintext	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

Letter-frequency analysis alone cannot be used to break poly alphabetical ciphers, but the Kasiski method can be used [9]. Using this method, first the period is found, then a set of mono alphabetic ciphers are solved. Using running-key ciphers [2], the message and key letters are added modulo 26 and a key (usually a text from a book) as long as the message is used. Friedman's method [7] searches for an accidental match between the plaintext and the key and breaks the running-key cipher by exploiting the redundancy and other statistical properties of the English language.

The Hagelin machine (M-209) and all Rotor ciphers (for example the British Typex, the American Sigaba (M-134), the German Enigma, the Japanese Purple) are poly alphabetic

Ciphers generating a stream with a large period. Since the sequences generated by the above machines are not random, cryptanalysis is possible, and some were successfully
broken during World War 2 [3],[13].

An important substitution cipher is the one-time pad in which the key is random, non repeating and used only once. One-time pads are unbreakable because there is not enough information in the cipher text to determine the message or the key uniquely; hence, the unicity distance is infinite. The first implementation of the one-time pad cipher was the Vernam cipher [19], in which the key bits were added modulo 2 to the plaintext bits. If the plaintext is broken into n-bit blocks, the cryptanalyst, even when trying all 2" keys, will only learn the length of the block *(n),* since the resulting plaintext will include not only the correct one, but all other meaningful plaintext of the same length. Two or more periodic key streams can be combined to produce a sequence with a longer period if their
periods are relatively prime. For a one-time pad the key required grows linearly with the message length, which limits their use for practical purposes

In our previous paper in Feb 2010, we (first two authors) have invented a new substitution algorithm which is different from the above is Play Color Cipher (PCC): [14] Each Character (Capital, Small letters, Numbers (0-9), Symbols on the keyboard) in the plain text is substituted with a color block from a 18 decillions of colors [21],[22] and at the receiving end the cipher text block (in color) is decrypted in to plain text block.

It overcomes the problems like "Meet in the middle attack, Birthday attack and Brute force attacks [23]". It also reduces the size of the plain text when it is encrypted in to cipher text by 4 times, with out any loss of content. Cipher text occupies very less buffer space. Hence transmitting through channel is very fast. Transportation cost through channel is very less.

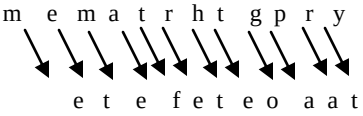
3.2 Transposition Ciphers: Transposition ciphers are block ciphers that change the position (or the sequence) of the characters or bits of the input blocks. To encipher, the plaintext is broken into *n* symbols and a key specifies one of (n!—1) possible permutations. Deciphering is accomplished by using an inverse permutation which restores the original sequence.

Transposition ciphers preserve the frequency distribution of single letters but destroy the diagram and higher-order distributions.

Transposition ciphers are often combined with other ciphers to produce a more secure product cipher.

The simplest such cipher is the *rail fence technique*, in which the plain text is written down as a sequence of diagonals and then read off as a sequence of rows.

For example, to encipher the message " *meet me after the toga party*" with a rail fence of depth 2, we write the following:



The encrypted message is: MEMATRHTGPRYETEFETEOAAT

A more complex scheme is to write the message in a rectangle, row by row and read the message column by column. But permute the order of the columns. The order of the columns then becomes the key to the algorithm. Example:

Key:	4 3 1 2 5 6 7
Plain text:	a t t a c k p
	o s t p o n e
	d u n t i l t
	w o a m x y z

Cipher text: T T NAAPTMTSUOAODWCOIXKNLYPETZ

4. STEGANOGRAPHY: A plain text message can be hidden by using Steganography. It conceals the existence of the message. A simple form of Steganography, but one that is time consuming to construct, is one in which an arrangement of words or letters within an apparently innocuous text spells out the real message. Example: the sequence of first letters of each word of the overall message spells out the hidden message. Or the sub set of words of the overall message is used to convey the hidden message.

Character marking: Selected letters of printed or typewritten text are over written in pencil. The marks are ordinarily not visible unless the paper is held at an angle to bright light.
Invisible ink: a number of substances can be used for writing but leave no visible trace until heat or some chemical is applied to the paper.
Pin punctures: Small pin punctures on selected letters are ordinarily not visible unless the paper is held up in front of a light. Typewriter correction ribbon: The result of typing with the correction tape is visible only under a strong light. Steganography techniques are more vulnerable to attacks [15].

5. CONCLUSION:

Number of variations that could be added to each of the system will undoubtedly rise in the reader's mind. The purpose of this paper, however, was not to present an exhaustive list of variations of each method; rather, it was intended to present a number of cryptographic techniques that are available for possible modifications for computer use.
It is not that the security is entirely depending on strength of cryptographic algorithms, rather it mainly depends on the secret key and the way which we share the secret key with the communicating party.

6. ACKNOWLEDGMENTS

Our sincere thanks to the Management Aizza College of Engineering and Technology, for providing all facilities to complete the task. We also thank IJCA for allowing us to modify template they had developed. We specially thank all our Family members for their overwhelming support all along.

7. REFERENCES

[1] Der- Chyuan Lou, Nan –I Wu, Chung-Ming Wang, Zong- Han Lin, Chwei-Shyoung Tsai , "A novel adaptive Steganography based on local complexity and human vision sensitivity", The Journal of Systems and Software 83 (2010) 1236-1248.

[2] Denning.D, "Cryptography and data security", Addison Wesley,1982.

[3] Deavours.C.A "The black chamber: a column, how the British broke Enigma", *Cryptoloyia*, 1980, 4, pp. 129 132

[4] David Khan, Mackmillan," The code breakers: The story of secreate writing, New York:, 1967.

[5] Diffi and Hellman, " Privacy and authentication: an introduction to cryptography",Proceedings of the IEEE,67(1979), PP, 379-427.

[6] Fabien A.P.Petitcolas, Ross J.Anderson and Markus G. Kuhn, "Information Hiding- A Survey", Procedings of the IEEE Special issue on protection of multimedia content, 87(7): 1062-1078, July 1999.

[7] Friedman. W, Riverbank, "Methods for the solution of running-key ciphers", Publ. 16, Riverbank Labs., Geneva, 1918

[8] F Ayoub, "Cryptographic techniques and network security", IEEE Proceedings, Vol.131, Dec 1984, 684-694.

[9] Kasiski. F, "Die Geheim Schriften und die Dechiflrirkunst", Mittier and Son, 1863.

[10] Konhein. A, "Cryptography: a primer", John Wiley, 1981.

[11] Linda S Rutledge, "A Survey of Issues in Computer Network Security", Elsevier Science Publishers B.V North Holland, Computers and Security 5 (1986) 296-308.

[12] Lewand, R, "Cryptological Mathematics", Washingtone, DC: The Mathematical association of America,2000.

[13] Rivest.R,"The impact of technology on cryptography", Proc. IEEE International Communications Conference, Toronto, Canada, June 1978.

[14] Ravindra babu Kallam, Dr.Udayakumar, "A block cipher generation using color substitution", International Journal for Computer Applications, (0975-8887), Vol-1, No-28

[15] Ross J.Anderson, Fabien A.P.Petitcolas, "On the limits of Steganography", IEEE Journal of selected areas in communications, 16(4): 474 – 481, May 1998.ISSN 0733-8716.

[16] Simmons, "Cryptography", Encyclopedia Britannica, Fifteenth Edition, 1993.

[17] Sinkov, "A Elementary Cryptanalysis: A Mathematical Approach", Washingtone, DC: The Mathematical association of America,1996.

[18] The Beale Cipher Association,"The Beale ciphers", Ned field, MA, 1978

[19] Vernam. G. *J*, "Cipher printing telegraph systems for secret wire and radio telegraphic communications", *AIEE*, 1926, 45, pp. 109 115

[20] William Stallings, "Cryptography and Network Security", Fifth Impression,2008, p age no: 35 – 54.

[21] www.whyiscolor.org, " for number of colors in the world".
[22] www.jimloy.com, " for number of colors in the world"
[23] William Stallings, "Cryptography and Network security" for Security attacks like -meet in the middle attack, broot force attacks p.n:353, birth day attack-p.n:350".