ISSN No. 0976-5697

Volume 3, No. 1, Jan-Feb 2012

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

A Novel Polyalphabetic Substitution Based Cryptosystem(NPSBC)

Sukalyan Som*
Department of Computer Science
Barrackpore Rastraguru Surendranath College
Kolkata, West Bengal, India
sukalyan.s@gmail.com

Arindam Sen
Department of Electronics and Instrumentation Engineering
JIS College of Engineering
Kalyani, Nadia, India
arisen@gmail.com

Abstract: Cryptography is the art and science of achieving security by converting sensitive information to un-interpretable form such that it cannot be interpreted by anyone except the intended recipient. An innumerable set of cryptographic schemes persist in which each of it has its own affirmative and feeble characteristics. Innumerable algorithms have been developed over the years to provide security of information but each of them has some merits and demerits. No single algorithm is sufficient on its own for this purpose. As a result researchers are constantly devoting themselves in the field of cryptography to eliminate the deficiency and find a better solution. In this paper an effort has been made to develop a Polyalphabetic Substitution based Substitution Cipher. Our aim is to encipher Plain texts depending on some parameters, which are continuously varied with every run-time so that the cipher text becomes different for similar entered texts every runtime.

Keywords: Cryptography, Encryption, Decryption, Polyalphabetic Substitution Cipher, Bit-ratio test, Frequency distribution test, Vulnerability

I. INTRODUCTION

Cryptography is defined to be a set of techniques and study of mathematics related to aspects of information security such as confidentiality, authenticity, integrity, non-repudiation [1]. While cryptography is the science of securing data, cryptanalysis is the science of analyzing and breaking secure communication. The person who uses cryptanalysis is called cryptanalyst or attacker [2]. Cryptology is a combination of both cryptography and cryptanalysis. Cryptanalytic attacks can be cipher text only, known plaintext, chosen plaintext, chosen cipher text, adaptive chosen plaintext, brute force attack, key guessing attack etc [3].

Innumerable algorithms have been developed over the years to provide security of information but each of them has some merits and demerits. No single algorithm is sufficient on its own for this purpose. As a result researchers are constantly devoting themselves in the field of cryptography to eliminate the deficiency and find a better solution.

In this paper we have presented a novel polyalphabetic substitution based cryptosystem wherein the plain text converted two cipher texts by the use of a variable length key stream producing different cipher text with every run time for similar entered plain text. In section 2 we have presented the basic terminologies. Section 3 depicts the proposed algorithm both for encryption and decryption. In section 4 we have presented the experimental results with the help of an example. Testing and analysis is given in section 5 comprising of frequency distribution test, encryption and decryption time comparison and vulnerability. On the basis of testing and analysis performed conclusions are drawn in section 6.

II. BASIC TERMINOLOGIES

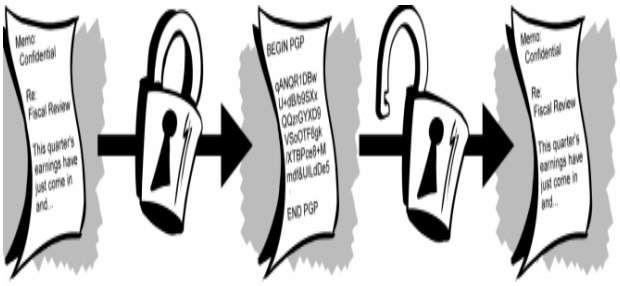
Plaintext or Clear text signifies a message that can be understood by the sender, the recipient and also by anyone else who gets an access to that message.

Cipher refers to the algorithm(s) for transforming an intelligible message to unintelligible form.

When a plain text message is codified using any suitable scheme, the resulting message is known as Cipher text.

A key is a number (or a set of numbers) that the cipher, as an algorithm operates on.

The method of disguising plaintext in such a way as to hide its substance is called encryption. Decryption is the process of reverting cipher text to its original plaintext.



Plaintext Encryption Cipher text Decryption Plaintext

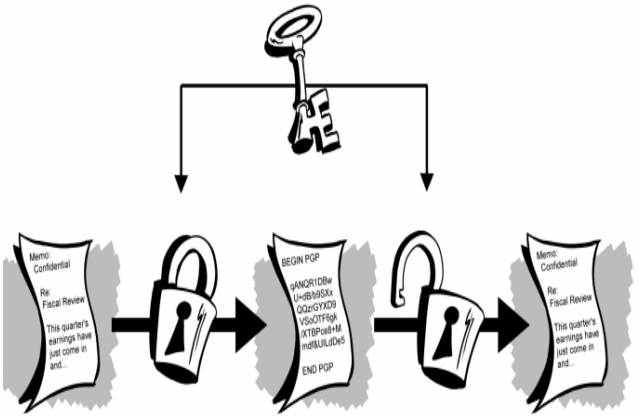
(Encryption Cipher + Encryption Key) (Decryption Cipher + Decryption Key)

Figure 1: Method of Encryption and decryption

As depicted in Figure 1, to encrypt a message, an encryption cipher, an encryption key and plaintext is needed which creates the cipher text. To decrypt a message a decryption cipher, a decryption key and the cipher text is needed which reveals the original plaintext.

Depending on the key(s) used for an encryption and decryption a cryptosystem can be classified as – Symmetric Cryptosystem and Asymmetric Cryptosystem.

As shown in Figure 2, Symmetric Cryptosystems use the same key, known as Secret key, to both encrypt and decrypt the message. It has a problem to transport the secret key from the sender to the receiver and in tamperproof fashion.

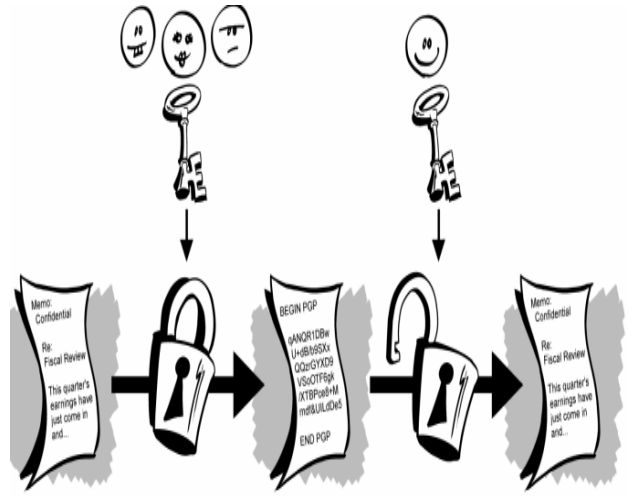


Plaintext Encryption Cipher text Decryption Plaintext

Figure 2: Symmetric Cryptosystem or Secret key Cryptosystem or Private key Cryptosystem

As shown in Figure 3, Asymmetric Cryptosystems use one key, known as Public key, to encrypt a message and a different key, the Private Key, to decrypt it. These are also known as Public Key Cryptosystems.

Public Key Private Key



Plaintext Encryption Cipher text Decryption Plaintext

Figure 3 Asymmetric key or Public key Cryptosystem

The primary benefit of public key cryptography is that it allows people who have no pre-existing security arrangement to exchange messages securely. The need for sender and receiver to share secret keys via some secure channel is eliminated; all communications involve only public keys, and no private key is ever transmitted or shared.

Classical Cryptography is based on information theory appeared in 1949 with the publication of "Communication Theory of Secrecy of Systems" by C. Shannon. In Classical Cryptography both plaintext and key length were same to support secrecy through encryption [4]. Symmetric Cryptosystems can be categorized as Traditional or Character Oriented and Modern or Bit Oriented. There are two primary ways in which a plaintext can be codified to get the cipher text using traditional or character oriented cryptography – Substitution Ciphers and Transposition Ciphers. When these two approaches are clubbed together, we call them Product Cipher. Substitution Ciphers can be categorized as either Monoalphabetic or Polyalphabetic ciphers.

Conventional encryption has benefits. It is very fast. It is especially useful for encryption of data that is not going anywhere. However, conventional encryption alone as a means for transmitting secure data can be quite expensive simply due to the difficulty of secure key distribution. For a sender and recipient to communicate securely using conventional encryption, they must agree upon a key and keep it secret between themselves.

In Polyalphabetic substitution cipher, primarily proposed by Leon Battista (1568), each occurrence of a symbol may have a different substitute. To create a Polyalphabetic cipher, effort should be made to make each cipher text character dependent on both the corresponding character and the position of the plaintext character in the message. This implies that the secret key should be a stream of sub keys, in which each sub key depends somehow on the position of the plaintext character that uses that sub key for encipherment. In other words, we need to have a key stream k = (k1, k2, k3 ...) in which ki is used to encipher the ith character in the plaintext to create the ith character in the cipher text. [5]

III. PROPOSED ALGORITHMS

A. Method of Encryption:

a. Start.
b. Input the plaintext.
c. Convert the plaintext characters to their corresponding ASCII (decimal) equivalent.
d. Input key stream string where length of this string should be less than the length of plain text.
e. Key stream characters are converted to corresponding ASCII (decimal) equivalent.
f. ASCII value of each plain text character is added with ASCII value of each character of key stream string block wise.
g. The resultant decimal numbers from step 6 are converted to their corresponding binary equivalent.
h. 1's complement of each binary equivalent block is found and corresponding decimal equivalent is found.
i. The character corresponding to each decimal number found in step 8 forms a cipher character corresponding to each plain text character.
j. Stop.

B. Method of Decryption:

a. Start.
b. Input the cipher text.
c. ASCII (decimal) equivalent of each cipher character is found and then converted to their corresponding binary equivalent.
d. 1's complement of the decimal numbers found in step 3 found and then converted to their corresponding decimal equivalent.
e. The ASCII (decimal) values of key stream characters are subtracted from each value calculated in step 4.
f. The resultant decimal numbers found in step 5 are converted to their corresponding character equivalent and thus the plain text characters are retrieved.
g. Stop.

IV. EXPERIMENTAL RESULTS

Let the plain text be "WELCOME" and key stream characters be "abc".

Table 1 Encryption result

Plain text & their ASCII (Decimal)	*Key stream Characters & their ASCII (Decimal)*	*SUM (ASCII) for col. 1 & col. 2*	*8 bit binary*	*1's complement*	*ASCII (Decimal) of col. 5*	*Cipher text*
W(8)	a(97)	184	10111000	01000111	71	G
E(69)	b(98)	167	10100111	01011000	88	X
L(76)	c(99)	175	10101111	01010000	80	P
C(67)	a(97)	164	10100100	01011011	91	[
O(79)	b(98)	177	10110001	01001110	78	N
M(77)	c(99)	176	10110000	01001111	79	O
E(69)	a(97)	166	10100110	01011001	89	Y

Table 2 Decryption Result

Cipher text & their ASCII (Decimal)	*8 bit binary*	*1's complement (8 bit)*	*ASCII (Decimal)*	*Key stream characters & their ASCII (Decimal)*	*Subtract col. 5 from col. 4*	*Plain text*
G(71)	01000111	10111000	184	a(97)	87	W
X(88)	01011000	10100111	167	b(98)	69	E
P(80)	01010000	10101111	175	c(99)	76	L
[(91)	01011011	10100100	164	a(97)	67	C
N(78)	01001110	10110001	177	b(98)	79	O
O(79)	01001111	10110000	176	c(99)	77	M
Y(89)	01011001	10100110	166	a(97)	69	E

V. TESTING AND ANALYSIS

In order to ensure the security level of a cryptosystem a number of effects have been made. Among of them we have used bit ratio, frequency distribution, time complexity analysis and also discusses vulnerability. The bit ratio effect means the changes the bit values from same position between plain text and cipher text. In the frequency distribution graph of source and encrypted file by proposed algorithm will be displayed. If the characters in the encrypted file are evenly distributed, it will make the cryptanalysis more difficult. The time complexity indicates how efficiently the proposed algorithm will encrypt the plain text and decrypt from encrypted text.

A. Frequency Distribution Test:

In the frequency distribution graph of source and encrypted file by proposed algorithm is presented in Figure 4.

If the characters in the encrypted file are evenly distributed, it will make the cryptanalysis more difficult. From the Figure it is evident that the cipher text generated by the above encryption cipher follows this criteria and thus serving the said purpose.

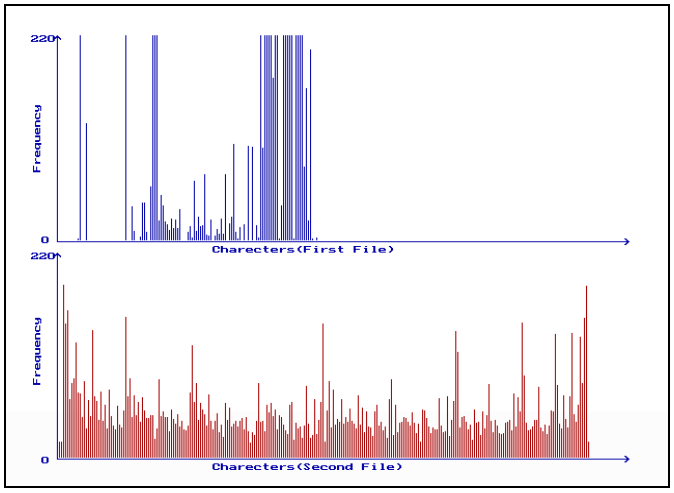


Figure 4. Frequency Distribution graph of plain text and cipher text

B. Bit-ratio Test:

The bit ratio effect means the changes the bit values from same position between plain text and cipher text. The bit ratio can be determined as:

Bit-ratio (in %) = {(Total number of bits changed in the file after encryption) ÷ (Total number of bits present in the file)} × 100. It has been verified that the bit ratio for our algorithm is better than existing algorithms like RSA, DES or [6].

Table 3 is representing the average bit-ratio approximated of our proposed algorithm is better than the other existing algorithms which is graphically presented in Figure 5.

Table 3 Average Bit-Ratio (approximate) comparison

File Name	*File Size (in kb)*	*RSA*	*DES*	*[6]*	*NPSBC*
File01	1.80	45.40	47.57	46.10	48.20
File02	3.80	44.90	46.43	45.15	45.72
File03	8.50	45.10	47.30	46.80	46.40
File04	12.58	44.00	44.00	45.00	44.05
File05	27.00	45.10	46.80	45.01	47.23

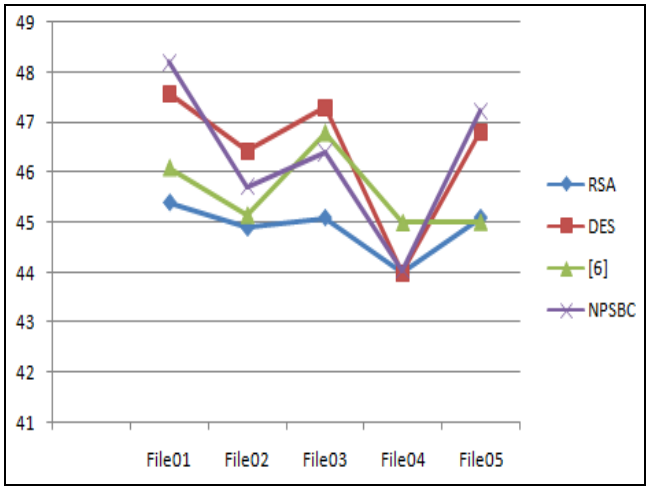


Figure 5: Average Bit-Ratio (approximate) comparison chart

C. *Encryption and Decryption Time Comparison:*

The time complexity indicates how efficiently the proposed algorithm will encrypt the plain text and decrypt from encrypted text. If the time complexity is much lower than time complexity on the same file of other existing algorithms, the proposed algorithm is better than existing algorithms like [6].

Table 4 Encryption and Decryption Time comparison

Source	*RSA*		*DES*		*[6]*		*NPSBC*	
	Enc	*Dec*	*Enc*	*Dec*	*Enc*	*Dec*	*Enc*	*Dec*
File 01	0.45	0.31	0.12	0.67	1.02	4.01	0.4	0.32
File 02	0.5	0.33	0.1	1.1	2.01	4.1	0.48	0.35
File 03	0.32	1.01	1.01	1.1	0.42	5.04	0.35	1.05
File0 4	0.54	0.21	2.0	2.1	1.12	4.1	0.5	0.7
File0 5	0.4	1.02	3	4.0	1.02	4.0	0.3	1.0

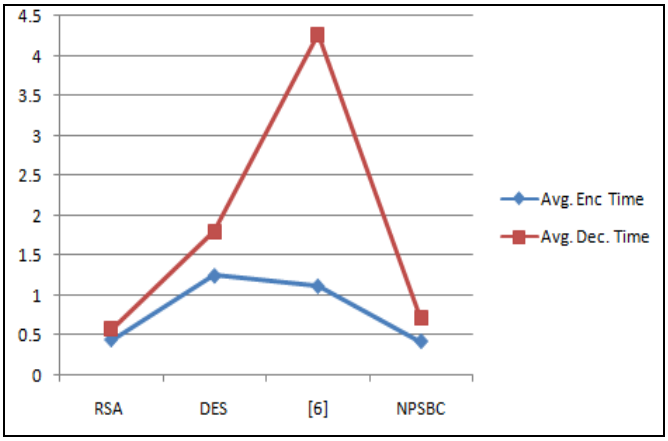


Figure 6: Encryption and Decryption Time comparison chart

Table 4 representing the required time for both encryption and decryption of different algorithms and it is cleared that our algorithm takes less time respect to other algorithms. Figure 6 graphically presents encryption and decryption time respectively for different algorithms.

D. *Vulnerability:*

Time required to use a brute force approach, which simply involves trying every possible key until an intelligent translation of the cipher text into plain text is obtained. On average, half of all possible key must be tried to achieve success.

Table 5 shows how much time is involved for various key spaces.

Key size (in bits)	*No. of alternate keys*	*Time required at 1 encryption / µs*	*Time required at 106 encryption/ µs*
56	256=7.2*1016	255 µs=1142 years	10.01 hours
64	264=1.8*1019	263 µs =2.9*105 yrs	106.75 days

Results are shown for two binary sizes. The 56-bits key size is used with the DES (Data Encryption Standard) algorithm; 64-bit key size (8 bit ASCII equivalent of 8 character long key stream) is used for our proposed algorithm. For each key size the results are shown assuming that it takes 1 µs perform a single decryption, which is a reasonable order of magnitude for today's machine. The final column of the table 5 considers the result for a system that can process a 1(one) millions keys per microsecond. As the key size increases the complexity of exhaustive search becomes infeasible to crack encryption directly. Algorithm with key of 56 bits (DES) takes 1142 years and our proposed algorithm needs 2.9*105 years (of considering 64 bits) to search appropriate key to crack encryption. As one can see at this performance level (considering above Table 5), DES can no longer be considered computationally secure compare to our proposed algorithm.

VI. CONCLUSION

On the basis of the observed experimental results, tests and analysis performed above it can be inferred that 'NPSBC' is an efficient and a sufficiently strong cryptosystem providing a superior level of security. From the comparison based on Encryption and Decryption time, it has been well proved that with regard to time complexity, use of 'NPSBC' is always advantageous. The comparison of Frequency Distribution also depicts that the encrypted characters i.e. cipher texts are evenly distributed and it has been made more difficult for the cryptanalysts to recover plain text from cipher text.

To conclude the proposed algorithm is a simple, straightforward and compact approach to develop a cryptosystem using the essence of elementary operations. It provides the same or sometimes even better level of security using minimal time complexity. A comparative study and security level will be verified in future with other well known classical algorithms like Hill or Play fair cipher.

VII. REFERENCES

[1] Bement A. L. et. al. (2004), Standards for Security Categorization of Federal Information and Information Systems, Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology Gaithersburg, MD 20899-8900

[2] Ayushi, (2010), A Symmetric Key Cryptographic Algorithm, International Journal of Computer Applications (0975 - 8887) Volume 1. No. 15.

[3] Atul Kahate, (2008)Cryptography and Network Security, Tata McGraw-Hill Education, pg. 47.

[4] Ijaz Ali Shoukat , Kamalrulnizam Abu Bakar and Mohsin Iftikhar, "A Survey about the Latest Trends and Research Issues of Cryptographic Elements", p 141, International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 2, May 2011, ISSN 1694 0814.

[5] Sukalyan Som, Saikat Ghosh, "A Survey of Traditional or Character Oriented Symmetric Key Cryptography", International Journal of Advanced Research in Computer Science, Vol. 2, No. 4, July-August 2011

[6] R. Venkateswaran, Dr. V. Sundaram, "Information Security: Text Encryption and Decryption with Poly Substitution Method and Combining the Features of Cryptography", p28-30, International Journal of Computer Applications, Vol. 3, No. 7, June 2010.