

Cryptography System for Online Communication Using Polyalphabetic Substitution Method

YekiniN. Asafe

Department of Computer Technology
Yaba College of Technology,
Lagos Nigeria
engryekini@yahoo.com

Aigbokhan E. Edwin

Department of Computer Technology
Yaba College of Technology,
Lagos Nigeria
enas_eddy@yahoo.com

Okiki F. Mercy

Department of Computer Technology
Yaba College of Technology,
Lagos Nigeria
folamercy@yahoo.com

ABSTRACT

Online communication is one of the common means of communication in this era of globe village. As the number of people being connected to online communication system through their mobile phone, computer or any other e-communication tools increases, there is need to secure the communication networks from adversaries (third parties) between the sender and receivers. There are many aspects to security approach in online communication environment. One essential aspect for secure communications is that of cryptography, which is the focus of this research. This research work aims at designing and implementing cryptosystem using a simple polyalphabetic cipher algorithm. The algorithm was coded with Java programming language. It was discovered that cryptography system with polyalphabetic cipher algorithm is better compared to Atbash cipher, Scytale cipher, Ceasar cipher etc, because it is very difficult for adversaries to decrypt an encrypted data without the key.

Index Terms: Cryptography, Cryptosystem, Decrypt data, e-communication, Encrypted data, Java programming language, online communication, Polyalphabetic cipher algorithm.

Date of Submission : 10 June 2014

Date of Acceptance : 05 July 2014

I. INTRODUCTION

Online communication (e-communication) refers to reading, writing, and communication via networked computers. It may be synchronous computer-mediated communication (whereby people communicate in real time via chat or discussion software, with all participants at their computers at the same time), or *asynchronous* computer-mediated communication (whereby people communicate in a delayed fashion by computer, using programs such as e-mail); and the reading and writing of online documents via the World Wide Web[1]

Online communication is one of the common means of communication in this era of globe village. As the number of people being connected to online communication system through their mobile phone, computer or any other e-communication tools increases, there is need to secure the communication networks from adversaries (third parties) between the sender and receivers. One of the best approaches to securing data or information between sender and receivers from third parties is cryptography.

Cryptography is the practice and study of techniques for secure online communication in the presence of third parties (called adversaries) [2]. It is the science of writing in secret code and is an ancient art. It can also be defined

as a method of storing and transmitting data in a form that only those intended for it can read and process.

Cryptography is an effective way of protecting sensitive information as it is stored on communication paths over any untrusted medium which includes just about any network particularly the internet. A cryptography system provides the following services: Confidentiality-It ensures that no unauthorised parties can access information except the intended receiver. Authenticity-validating the source of the message to ensure the sender is properly identified; Integrity-provides assurance that the message was not modified during transmission, accidentally or intentionally and Non-repudiation-This means that a sender cannot deny sending the message at a later date and the receiver cannot deny receiving it.[3].

Objectives of This Research Work

The objectives of this research work are to:

- To enhance secure communication between communication links where the intrusion of an adversary or third party is completely eradicated.
- To provide confidentiality, authenticity, integrity, and non-repudiation of the information

transmitted through the online communication network.

- To develop a cryptographic algorithm that is an improvement on the existing system

Scope of Study

This research work focuses on design and implementation of crypto systems for preventing third parties from assessing data/information (message) between sender and receiver on online communication system in order to ensure confidentiality, authenticity, integrity and non-repudiation of the message.

II. LITERATURE SURVEY

A. *History of Cryptography*

Cryptography is derived from the Greek word κρυπτός meaning "hidden, secret"; and γράφειν, graphēin, "writing", or -λογία, -logia, "study", respectively [4] is the practice and study of techniques for secure communication in the presence of third parties (called adversaries). Cryptography has roots that began around 2000 B.C. in Egypt when hieroglyphics were used to decorate tombs to tell the story of the life of the deceased[5]. The practice was not as much to hide the messages themselves, but to make them seem more noble, ceremonial, and majestic. Encryption methods evolved from being mainly for show into practical applications used to hide information from others.

Before the modern era, cryptography was concerned solely with message confidentiality (i.e. encryption)—conversion of messages from comprehensible form into an incomprehensible one and back again at the other end, rendering it unreadable by interceptors or eavesdroppers without secret knowledge (namely the key needed for decryption of that message).[6]

A Hebrew cryptographic method required the alphabet to be flipped so that each letter in the original alphabet is mapped to a different letter in the flipped alphabet. The encryption method was called Atbash. [7]

B. *Cryptographic Algorithm*

Algorithms are usually complex mathematical formulas that dictate the rules of how the plaintext will be turned into cipher text. A key is a string of random bits that will be inserted into the algorithm. For two entities to be able to communicate via encryption, they must use the same algorithm. There are three main categories of cryptographic algorithm that are employed for encryption and decryption. These algorithms are:

Symmetric Cryptography: In a cryptosystem that uses symmetric cryptography, both parties will be using the same key for encryption and decryption. This provides dual functionality. Symmetric keys are also called secret keys because this type of encryption relies on each user to keep the key a secret and properly protected. If this key got into an intruder's hand, that intruder would have the ability to decrypt any intercepted message encrypted with this key. Each pair of users who want to exchange data using symmetric key encryption must have their own set of keys. The security of the symmetric encryption method is completely dependent on how well users protect the key. Because both users use the same key to encrypt and decrypt messages, symmetric cryptosystems can provide confidentiality, but they cannot provide authentication or nonrepudiation.[8] There is no way to prove who actually sent a message if two people are using the exact same key.

When using symmetric encryption you can encrypt and decrypt large amounts of data that would take an unacceptable amount of time if an asymmetric algorithm was used instead. It is also very difficult to uncover data that is encrypted with a symmetric algorithm if a large key size was used. The major disadvantages of using symmetric is key distribution that is, it requires a secure mechanism to deliver keys properly and also scalability that is, each pair of user needs a unique pair of keys, so as the number of users increase the number of keys grow exponentially.

Symmetric key ciphers are implemented as either block ciphers or stream ciphers. A block cipher enciphers input in blocks of plaintext that is work on blocks of plaintext and cipher text, as opposed to individual characters, the input form used by a stream cipher.

Block Cipher: When a block cipher algorithm is used for encryption and decryption purposes, the message is divided into blocks of bits. These blocks are then put through substitution, transposition, and other mathematical functions. The algorithm dictates all the possible functions available to be used on the message, and it is the key that will determine what order these functions will take place.[9] Strong algorithms make reengineering, or trying to figure out all the functions that took place on the message, basically impossible.

Stream Cipher: A stream cipher does not divide a message up into blocks; instead, a stream cipher treats the message as a stream of bits or bytes and performs mathematical functions on them individually When using a stream cipher, the same plaintext bit or byte will be transformed into a different cipher text bit or byte each time it is encrypted. Some stream ciphers use a keystream generator, which produces a stream of bits that is XORed with the plaintext bits to produce cipher text. Exclusive OR

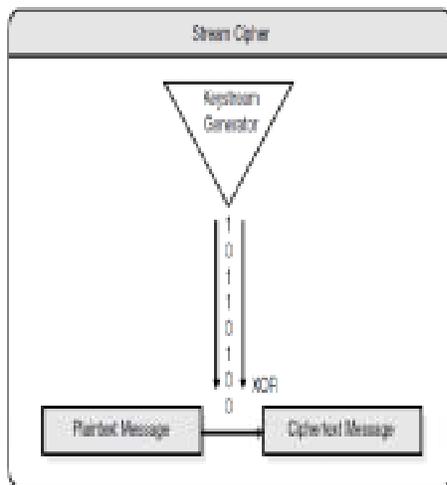
(XOR) is an operation that is applied to two bits. It is a function in binary mathematics. If both bits are the same, the result is zero ($1 + 1 = 0$). If the bits are different than each other, the result is one ($1 + 0 = 1$). [10]

Example:

Message stream 1001010111

Keystream 0011101010

Ciphertext stream 1010111101



Source: Cissp online exam guide [7]

The value that is generated by the keystream generator is XORed with the bits of the plaintext.

If the cryptosystem was only dependent upon this keystream generator, an attacker could get a copy of the plaintext and the resulting ciphertext, XOR them together, and find the keystream to use in decrypting other messages. So the smart people decided to stick a key into the mix. In stream ciphers, the key also provides randomness to the keystream that is actually applied to the plaintext. The key is a random value input into the stream cipher, which it uses to ensure the randomness of the keystream data. A strong and effective stream cipher algorithm contains long periods of no repeating patterns within keystream value, it should be statistically unpredictable, the keystream should not be linearly related to the key and it should have statistically unbiased keystream (as many 0's as 1's) because stream ciphers encrypt and decrypt one bit at a time. [11]

Stream ciphers are more suitable for suitable for hardware implementations while block ciphers are easier to implement in software because they work with blocks of data that the software is used to working with, which is usually the width of a data bus (64 bits). Stream ciphers are intensive because each bit must be manipulated.

Asymmetric Cryptography: In asymmetric systems, each entity has different keys, or asymmetric keys. The two different asymmetric keys are mathematically related. If a message is encrypted by one key, the other key is required to decrypt the message. In asymmetric system also known as a public key system, the pair of keys is made up of one public key and one private key. The public key can be known to everyone, and the private key must only be known to the owner. Many times, public keys are listed in directories and databases of e-mail addresses so they are available to anyone who wants to use these keys to encrypt or decrypt data when communicating with a particular person. The public and private keys are mathematically related, but cannot be derived from each other. [7] An asymmetric cryptosystem works much slower than symmetric systems, but can provide confidentiality, authentication, and nonrepudiation depending on its configuration and use. Asymmetric systems also provide for easier and more manageable key distribution than symmetric systems and do not have the scalability issues of symmetric systems.

Data Encryption Standard (DES): DES is a block encryption algorithm. When 64-bit blocks of plaintext go in, 64-bit blocks of cipher text come out. It is also a symmetric algorithm, meaning the same key is used for encryption and decryption. It uses a 64-bit key, 56 bits make up the true key, and 8 bits are used for parity. When the DES algorithm is applied to data, it divides the message into blocks and operates on them one at a time. A block is made up of 64 bits and is divided in half and each character is encrypted one at a time. The characters are put through 16 rounds of transposition and substitution functions. The order and type of transposition and substitution functions depend on the value of the key that is inputted into the algorithm. The result is a 64-bit block of cipher text. [12]

Cipher Block Chaining (CBC) Mode: CBC does not reveal a pattern because each block of text, the key, and the value based on the previous block is processed in the algorithm and applied to the next block of text [7]. This gives a more random resulting cipher text. A value is extracted and used from the previous block of text. This provides dependence between the blocks and in a sense they are chained together. This is where the title of Cipher Block Chaining (CBC) comes from, and it is this chaining effect that hides any repeated patterns. The results of one

block are fed into the next block, meaning that each block is used to modify the following block. This chaining effect means that a particular ciphertext block is dependent upon all blocks before it, not just the previous block.

Output Feedback (OFB) Mode: This is a block cipher implementation conceptually similar to a synchronous stream cipher. OFB prevents the same plaintext block from generating the same ciphertext block by using an internal feedback mechanism that is independent of both the plaintext and ciphertext bitstreams. [3]. It is functioning like a stream cipher by generating a stream of random binary bits to be combined with the plaintext to create cipher text. The cipher text is fed back to the algorithm to form a portion of the next input to encrypt the next stream of bits. As previously stated, block cipher works on blocks of data and stream ciphers work on a stream of data. Stream ciphers use a keystream method of applying randomization and encryption to the text, whereas block ciphers use an S-box-type method. In OFB mode, the DES block cipher crosses the line between block cipher and stream cipher and uses a keystream for encryption and decryption purposes.

C. Types of Asymmetric Encryption Algorithms

There are several types of asymmetric algorithms used in the computing world today. They may have different internal mechanisms and methods, but the one thing they do have in common is that they are all asymmetric. This means that a different key is used to encrypt a message than the key that is used to decrypt a message.

RSA: RSA, named after its inventors Ron Rivest, Adi Shamir, and Leonard Adleman, is a public key algorithm that is the most understood, easiest to implement, and most popular when it comes to asymmetric algorithms. RSA is a worldwide de facto standard and can be used for digital signatures and encryption. It was developed in 1978 and provides authentication as well as encryption. The security of this algorithm comes from the difficulty of factoring large numbers. The public and private keys are functions of a pair of large prime numbers and the necessary activities required to decrypt a message from cipher text to plaintext using a public key is comparable to factoring the product of two prime numbers. (A prime number is a positive whole number with no proper divisors, meaning the only numbers that can divide a prime number is one and the number itself.) One advantage of using RSA is that it can be used for encryption and digital signatures. Using its one-way function, RSA provides encryption and signature verification and the inverse direction performs decryption and signature generation. [2]

Elliptic Curve Cryptosystems (ECCs): Elliptic curves are rich mathematical structures that have shown usefulness in

many different types of applications. An Elliptic Curve Cryptosystem (ECC) provides much of the same functionality that RSA provides: digital signatures, secure key distribution, and encryption. [6] One differing factor is ECC's efficiency. Some devices have limited processing capacity, storage, power supply, and bandwidth like the newer wireless devices and cellular telephones. With these types of devices, efficiency of resource use is very important. ECC provides encryption functionality requiring a smaller percentage of the resources required by RSA and other algorithms, so it is used in these types of devices. In most cases, the longer the key length, the more protection that is provided, but ECC can provide the same level of protection with a key size that is smaller than what RSA requires. Because longer keys require more resources to perform mathematical tasks, the smaller keys used in ECC require fewer resources of the device. ECC cryptosystems use the properties of elliptic curves in their public key systems. The elliptic curves provide ways of constructing groups of elements and specific rules of how the elements within these groups combine. The properties between the groups are used to build cryptographic algorithms. [7].

III. DESIGN METHODOLOGY

A. DESCRIPTION OF THE NEW SYSTEM

This program uses the substitution encryption cipher method using symmetric cryptography in which both the sender and the receiver use the same key for encryption and decryption. This algorithm is a simple method of encrypting alphabetic text by using a series of different shift ciphers based on the letters of a keyword. It is a simple form of polyalphabetic substitution. A polyalphabetic substitution cipher involves the use of two or more cipher alphabets. Instead of there being a one-to-one relationship between each letter and its substitute, there is a one-to-many relationships between each letter and its substitutes. That is, each letter is moved by a number of locations to the left of the alphabet specified by the secret key. That is, each alphabet produces random values that is dependent on the letter to be encrypted and the secret key rather than fixed values.

The design methodology focuses on the algorithm used in the implementation of the design, the choice of programming language, the input and output model which explains how the user interacts with the system.

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	DEFGHIJKLMNOPQRSTUVWXYZABC
E	EFGHIJKLMNOPQRSTUVWXYZABCD
F	FGHIJKLMNOPQRSTUVWXYZABCDE
G	GHIJKLMNOPQRSTUVWXYZABCDEF
H	HJKLMNOPQRSTUVWXYZABCDEFG
I	IJKLMNOPQRSTUVWXYZABCDEFGH
J	JKLMNOPQRSTUVWXYZABCDEFGHI
K	KLMNOPQRSTUVWXYZABCDEFGHIJ
L	LMNOPQRSTUVWXYZABCDEFGHIJK
M	MNOPQRSTUVWXYZABCDEFGHIJKL
N	NOPQRSTUVWXYZABCDEFGHIJKLM
O	OPQRSTUVWXYZABCDEFGHIJKLMN
P	PQRSTUVWXYZABCDEFGHIJKLMNO
Q	QRSTUVWXYZABCDEFGHIJKLMNOP
R	RSTUVWXYZABCDEFGHIJKLMNO
S	STUVWXYZABCDEFGHIJKLMNO
T	TUVWXYZABCDEFGHIJKLMNO
U	UVWXYZABCDEFGHIJKLMNO
V	VWXYZABCDEFGHIJKLMNO
W	WXYZABCDEFGHIJKLMNO
X	XYZABCDEFGHIJKLMNO
Y	YZABCDEFGHIJKLMNO
Z	ZABCDEFGHIJKLMNO

To encrypt, a table of alphabets can be used. It consists of the alphabet written out 26 times in different rows, each alphabet shifted cyclically to the left compared to the previous alphabet, corresponding to the 26 possible. At different points in the encryption process, the cipher uses a different alphabet from one of the rows. The alphabet used at each point depends on the secret key. For example, suppose that the plaintext to be encrypted is INCREDIBLE and the person encrypting uses ENCRYPTION as the secret key. Each row starts with a secret key letter. The remainder of the row holds the letters A to Z (in shifted order). Although there are 26 key rows shown, you will only use as many keys (different alphabets) as there are unique letters in the key string. For successive letters of the message, we are going to take successive letters of the key string, and encipher each message letter using its corresponding key row. Choose the next letter of the key, go along that row to find the column heading that matches the message character; the letter at the intersection of [key-row, msg-col] is the enciphered letter.

Plaintext: INCREDIBLE
 Secret Key: ENCRYPTION
 Ciphertext: MAEICSBJZR

For example, the first letter of the plaintext, I, is paired with E, the first letter of the Secret key. So use row E and column I of the alphabet table namely M. Similarly, for the second letter of the plaintext, the second letter of the Secret key is used; the letter at row N and column N is A. The rest of the plaintext is enciphered in a similar fashion.

Decryption is performed by going to the row in the table corresponding to the Secret key, finding the position of the ciphertext letter in this row, and then using the column's label as the plaintext. For example, in row E (from ENCRYPTION), the ciphertext M appears in column I, which is the first plaintext letter. Next we go to row N, locate the ciphertext A which is found in column N, thus N is the second plaintext letter. Flowchart for the program is as shown in figure 1.

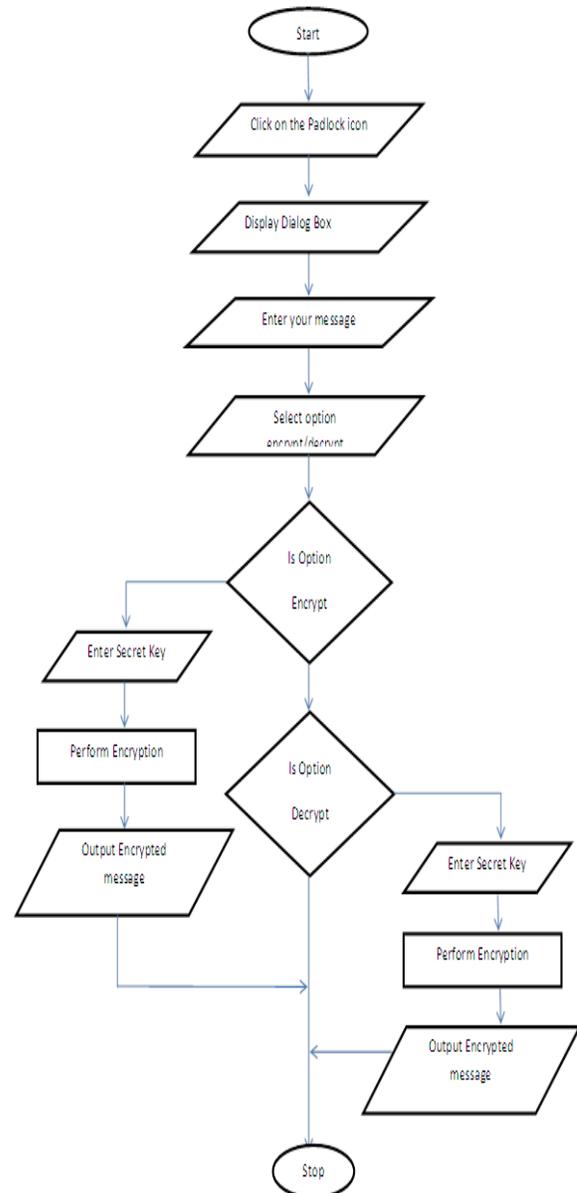


Figure 1. Flowchart for the cryptosystem

B. Choice Of Programming Language

Sample -2

A letter to Dad

Input- plaintext

Dear dad,

How are you sir? Hope you are having a great vacation? Hope your journey was hitch free? Everyone back home misses you so much and we can't wait to have you around again.

Please ensure that you make the best of your vacation and enjoy yourself to the fullest before you resume back to work. Also take time to relax and visit exciting places.

Everyone back home sends their love especially mum, she can't wait to have you back home we all miss you and we look forward to see you soon!

Secret key - private

Output – ciphertext

Svimdth,

Wfevrx cdl adr?Asevgjutvtiyiqigk p
xzzamzptioihr?Wfxzyhygawprginninhbxrynmex?
lkvztogiqrkfhhtdqnsxwnfc no fyryiid pi rrv'owtmi kw
caoinfcvrhycuibabr.

Ecmvsvxicjcmemlpgkjufezvbceuihkwayhygmixammdei
xyrfgtonvhvtathxwvnpheihkjzfhvtpwprxwjdwmwavo if ejrd.
Eajwoadiizuzthvtcisaghkzadtbrzbdnztarkzs.

Xztigjnxftscofihvvysmltzzgootjxzceacghuf,
wwwkvn'mapzbooaekvgjuerbpmx at rtgmbwhpwpagh lv
tjodjdievrwxdjmyhyhfw!

V. Conclusion

The need for security of information/data from sender to receivers in an online communication cannot be over emphasized. Although the ultimate goal of cryptography, and the mechanisms that makes it up, is to hide information from unauthorized individuals, most algorithms can be broken and the information can be revealed if the attacker has enough time, desire, and resources. So a more realistic goal of cryptography is to make obtaining the information too work-intensive to be worth it to the attacker. This research work developed stand-alone software that implements

cryptography method using Polyalphabetic Substitution to secure data from sender to receiver. This program can be implemented with chatting software or E-mail software where the data transmitted over a network is in encrypted format. However, the system does have a high confidentiality rating in order to defend against sniffing and man-in-the-middle attacks. No security system should be based on one single security technique. In a networking system or local system it is advised to develop multiple levels of security techniques arranged in a layered boundary of defense.

References

- [1] Yekini N.A., Lawal O.N., (2011). ICT for accountant. Volume 2, Hasfem Publication, Nigeria
- [2] Rivest, Ronald L(1990). "Cryptography". In J Van Leeuwen. Handbook of theoretical computer Science.
- [3] Oludipe, O., Yekini, N., & Adelokun, P.(2012). Data communication & network, Hasfem Publication Nigeria.
- [4] Liddell and Scott's (1984). Greek-English Lexicon. Oxford University Press. www.swiss.ai.mit.edu/
- [5] Franksen, O. I. (1985) Mr. Babbage's Secret: The Tale of a Cipher and APL. Prentice Hall.
- [6] David, Kahn (1999). "On the Origin of a Species". The Codebreakers: The Story of Secret Writing. Simon & Schuster. ISBN 0-684-83130-9.
- [7] CISSP (2001). All-in-One Certification Exam Guide.
- [8] Martin, Keith M. (2012). Everyday Cryptography. Oxford University Press. P. 142. ISBN 978-0-19-162588-6.
- [9] Knudsen, Lars R. (1998). "Block Ciphers— a survey". In Bart Preneel and Vincent Rijmen. State of the Art in Applied Cryptography: Course on Computer Security and Industrial Cryptograph Leuven Belgium, June 1997 Revised Lectures. Berlin ; London: Springer. Pp. 29. ISBN 3-540-65474-7.
- [10] David Wheeler (2008). "Summary of Security Techniques" Purdue University Calumet ITS 454.
- [11] Beutelspacher, Albrecht (1994). "Chapter 2". Cryptology. translation from German by J. Chris Fisher. Washington, DC: Mathematical Association of America. pp. 27–41. ISBN 0-88385-504 Boran, S. (2000). IT Security Cookbook.
- [12] Bruen, Aiden A. & Forcinito, Mario A. (2011). Cryptography, Information Theory, and Error-Correction: A Handbook for the 21st Century. John Wiley & Sons. p. 21. ISBN 978-1-118-03138-4