ABACUS, (Mathematics Science Series) Vol. 44, No 2. Dec. 2017

MULTIPLE CAESAR CYPHER ENCRYPTION ALGORITHM

Balogun, A.O. , Sadiku, P. O., Mojeed, H. A., Raifu, H. A.

Department of Computer Science
University of Ilorin, Ilorin, Nigeria.
balogun.ao1@unilorin.edu.ng; peterposdoy@gmail.com; mojeed.ha@unilorin.edu.ng; dmanofambition@gmail.com

Abstract
The Caesar cipher has always been the major reference point when cryptographic algorithms (also called ciphers) are discussed. This, probably, is due to its being an age-long cipher. It may also be owing to the belief that the Caesar cipher was the first cipher used ever. Caesar cipher operation is based on shift-by-3 rule which makes its breaking obviously easy since an exhaustive key search of the other 25 keys can be conveniently performed. Ipso facto, an investigation into an enhancement of this too-simple-to-crack cipher is invariably necessary and ultimately important. This study is, therefore, concerned with developing a new enhanced model of Caesar cipher for a better security using multiple encryption technique, whereby an already-encrypted message is encrypted one or more times using the same or different algorithm. The new model works by wrapping a plaintext message in three crypto-wrappers and each encryption/decryption phase uses a different shift key from the other. The model supports both uppercase and lowercase characters. However, the model does not encrypt/decrypt numbers, special characters, whitespace, and file types such as word document, binary, or pdf files, but only text files. Most importantly, the new enhanced model is able to provide a better security of message by encrypting a plaintext message three times; in this way, brute forcing or an exhaustive key search will be difficult to perform; thus, making cryptanalysis almost a mirage!
Keywords: Cryptography, Multiple encryption, Encryption, Decryption, Plaintext, Ciphertext

1.0 Introduction

Information, which is defined as processed data, is an important asset in that the more information you have at your command, the better you can adapt to the world around you (Rhodes-Ousley, 2013). Companies may have *confidential information*(such as research and development plans, manufacturing processes, strategic corporate information, product roadmaps, process descriptions, customer lists and contact information, financial forecasts, and earnings announcements) that is intended for internal use on a need-to-know basis; loss or theft of this confidential information could violate the privacy of individuals, reduce the company's competitive advantage, or cause damage to the company (Rhodes-Ousley, 2013). Another form of information is s*pecialized information* or *secret information* which may include trade secrets, such as formulas, production details, and other intellectual properties, proprietary methodologies and practices that describe how services are provided, research plans, electronic codes, passwords, and encryption keys; if disclosed, this type of information may severely damage the company's competitive advantage (Rhodes-Ousley, 2013). It is, therefore, usually restricted to only few people or departments within a company and is rarely disclosed outside the company. The better the security controls (for protecting all these different types of data) are, the greater the level of access that can be safely provided to authorized parties who need to use that data.

The practice of keeping one's property out of the reach of intruders who prey on people's property illegally is as old as man himself; when this property is in form of data or information, it is called information security (Rhodes-Ousley, 2013). Information security encapsulates and provides services such as confidentiality, data integrity, authentication and non-repudiation (Rhodes-Ousley, 2013). Henk and Van (2000) opined that, "The protection of sensitive information against

unauthorized access or fraudulent changes has been of prime concern throughout the centuries". Secret writing has been in use for thousands of years, probably for as long as we have had secrets to keep, which may be in form of messages in war, messages between corporations, or just personal secret messages. One of the oldest methods of securing data is (the use of) cryptography (Rhodes-Ousley, 2013).

Cryptography has a long and fascinating story; it can be traced from its initial and limited use by the Egyptians some 4,000 years ago, to the twentieth century, where it had played a crucial role in the outcome of both World wars one and two (Khan, 1973). According to Rhodes-Ousley (2013) the word 'cryptography' was derived from two Greek words, 'Krypto' meaning hidden or secret and 'Graphene' which means writing; thus, cryptography succinctly refers to "secret writing". According to Rhodes-Ousley (2013), cryptography is defined as the science and art of concealing information using some special measures, called cryptographic algorithms or *ciphers*, such that only the intended users can have access to them. In other words, cryptography is the process of converting messages in readable form (called *plaintext*) to unreadable form (called *ciphertext*) in order to hide the meaning from intruders. This process is reversible in that the ciphertext generated can be converted back to the original plaintext in a process called cryptanalysis.

A cipher is a system which converts plaintext into ciphertext by applying a set of transformations to each character (or letter) in the plaintext. The particular transformations employed at any time are controlled by a *key* used at that time and the security of the ciphertext is said to rest heavily on the secrecy of the key (Luciano& Prichett, 1987). Caesar cipher, which is one of the earliest cryptographic systems, was first used by Julius Caesar around 50B.C. This cipher works by shifting the alphabet three places to the right and wrapping the last three letters X, Y, Z back unto the first three letters, thus:

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
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

The reason for choosing 3 as the key to the cipher system is unknown, any other integer value could have been chosen (Khan, 1973).Caesar cipher operates on modulo 26, so, there are 26 distinct keys. For example, the plaintext message **hello world** is transformed into the ciphertext **khoorzruog**. Moreover, such a transformation can also be performed by a computer using modular arithmetic (Luciano & Prichett,1987). Any message can be expressed, in digital form, based on the one-to-one correspondence, thus:

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
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 2324 25

Generally, using the Caesar cipher, to encrypt any plaintext message it requires the transformation $C = P + (3 \bmod 26)$, where C is the numeric equivalent of a ciphertext letter and P is the numeric equivalent of a plaintext letter. To decrypt, on the other hand, the transformation $C - (3 \bmod 26) = P$ is used (Luciano and Prichett(1987).

A message enciphered using the Caesar cipher is extremely insecure since an exhaustive cryptanalysis using the remaining 25 non-trivial keys can easily be performed and the message deciphered almost effortlessly (Paar&Pelzl, 2010). Consequently, an improvement on the cipher, which could make cryptanalysts use a whole lifetime trying to decipher messages, becomes very imperative and inevitable. This work, therefore, modifies the Caesar cipher using multiple encryption technique. Multiple encryption is the process of encrypting an already-encrypted message one or more times, either using the same or a different algorithm (Maurer & Massey, 1993). Thus, the message to be encrypted is wrapped in various crypto-wrappers: the ciphertext generated from the original plaintext message will serve as the plaintext for the second phase of the encryption and this will go on and on, up to the last phase. Additionally, the key formulation

methods used include the length of the penultimate word in the message, the lengths of the first and the last words combined, or twice the length of the first word.

This research work is basically based on the work of Abdulkareem, F. A. & Imran, E. I. (2014). It is aimed at modifying the algorithm proposed by the two authors by employing the tecnique of multiple encryption. This new proposed enhancement method moved a step ahead of the ancient Caesar cipher by varying the key size used, based on any of the key formulation methods such as the length of the penultimate word in the message, the lengths of the first and the last words combined, or twice the length of the first word, for all the encryption and decryption phases; and also, by making use of cascade ciphering (as multiple encryption is sometimes called). With this, the security of messages is assured as cryptanalysis will take longer time.

The objective of the research is to develop an enhanced model of the Caesar cipher based on the strengths and weaknesses derived from the Caesar cipher for a better data security

2.0 Literature Review

Cryptology, which is the study of cryptosystems, encompasses two disciplines: cryptography and cryptanalysis. Cryptography is concerned with the design of cryptosystems, while cryptanalysis studies the breaking of cryptosystems (Henk & Van, 2000).The branch of science which deals with secure communication on presence of intruders is cryptography(Henk& Van, 2000).It is one of the major areas of study in information security. Others include, but not limited to, steganography and network security. Cryptography is defined as the science and art of *encrypting* and *decrypting* data using some special measures. Encryption is the method of disguising *plaintext* in such a way as to hide its substance while decryption (which is the opposite of encryption) is unhiding the substance by changing the c*iphertext* to its original plaintext (Rhodes-Ousley, 2013). Cryptanalysis, on the other hand, is the branch of science which deals with breaking the codes and extracting the hidden meaning, while the whole system which comprises both cryptography and cryptanalysis is called cryptosystem (Paar& Pelzl, 2010).

2.1 Cryptographic Algorithms (Ciphers)

A cipher is an algorithmic function which converts plaintext messages into unreadable forms by applying a set of transformation techniques to each letter in the plaintext (Luciano &Prichett, 1987). The particular transformations employed at any time are controlled by a secret *key,* called cryptographic key*,* used at that time. The security of the ciphertext is said to rest majorly on the secrecy of this key(Luciano &Prichett, 1987).

In a study in 2009, Poschmann wrote that ciphers can be classified using several criteria. According to one of such criteria, the ciphers are classified as symmetric key and asymmetric key. In symmetric key ciphers, the same key is used for both encryption and decryption. A major problem with such a system is that the sender and receiver must know the key prior to transmission. This requirement makes such a system difficult to use in practice. The key cannot be openly transmitted since that would compromise the security of the system and that one possibility is for the two parties to meet and exchange the keys prior to transmitting their messages. However, this exchange becomes more difficult when many parties are involved in a communication network. An asymmetric key cipher uses different keys for encryption and decryption. These two keys are mathematically related, but it is very difficult to obtain one from the other unless one knows the transformation. The key used for encryption is called the public key and the key used for decryption is called the private key. The public key can be revealed without compromising the security of the system while the corresponding private key, however, must not be revealed to any party.

Poschmann (2009) further wrote that, in symmetric encryption, ciphers can be classified into stream ciphers and block ciphers; stream ciphers obtain ciphertext by using the XOR of the plaintext and keystream (bi-wise). They are grouped into two: synchronous stream cipher, whose key sequence does not depend on the plaintext and ciphertext but only on the previous elements of the key sequence and the initial key, e.g. One-time password (OTP); and asynchronous stream cipher, whose keystream depends on the plaintext or ciphertext, e.g. Cipher Feedback mode (CFB). Other examples of stream ciphers include RC4 and SEAL. Block ciphers, on the other hand, operate on a fixed length block size.It can be considered simply as a large lookup-table (substitution cipher). In particular, identical plaintext blocks encrypt to identical ciphertextblocks. Examples include Data Encryption Standard (DES), 3Data Encryption Standard (3DES), Advanced Encryption Standard (AES), Blowfish, etc.

The earliest known use of a substitution cipher, and the simplest, was by Julius Caesar (Purnama &Rohayani, 2015). Since Julius Caesar used an additive cipher to communicate with his officers; for this reason, additive ciphers are sometimes referred to as the Caesar Cipher (Purnama &Rohayani, 2015). In Cryptography, Caesar cipher is one of the most widely known encryption-decryption algorithms (Rhodes-Ousley, 2013). It is a type of substitution cipher whereby each letter in the plaintext is replaced by a letter some fixed numberof positions (usually three) down the alphabet(Stallings, 2005). The encryption is represented using modular arithmetic. For example, with thenatural shift of 3, A would be replaced by D; B would be replaced by E, and C by F, and so on. Thus, each letter is mapped into its corresponding letter as below:

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
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

Caesar cipher is also referred to as Caesar's cipher, Caesar's code or less frequently shift cipher(Purnama &Rohayani, 2015). A shift cipher uses any key from 0 to 25 while the Caesar cipher naturally only operates on the 'shift by 3' rule (Abdulkareem& Imran, 2014).

The major drawback of Caesar cipher is that it can easily be broken, even in ciphertext-only scenario (Jain, Dedhia &Patil, 2015). Various methods to crack the ciphertext using frequency analysis and words patterning have been proposed and implemented over time (Purnama &Rohayani, 2015). One of the approaches is using brute force to match the frequency distribution of letters. This is possible because there is only limited number of possible shifts (0-25, that is, 26).

3.0 The Proposed Algorithm

This study performs a modification of the Caesar cipher using cascade encipherment. In this new method, the encryption key is made to be flexible such that the ciphertext is obtained by combining the corresponding integer values of the plaintext with the keystream. In other words, the algorithm uses stream cipher for the key formulation. Mathematically, it can be expressed thus: $C = P + (k \bmod 26)$ where C represents ciphertext, P, plaintext and k, the ciphering key (any whole number between 0 and 25). To decrypt the ciphertext and get the (original) plaintext back, the key is subtracted from the ciphertext, thus: $P = C - (k \bmod 26)$.The key formulation uses any of the following methods (among others):

i. The length (that is, the number of characters) of the penultimate word in the message
ii. The lengths of both the first and the last words combined
iii. Twice the length of the first word in the message

Messages enciphered with the Caesar cipher are far from being secure(Paar& Pelzl, 2010). This fact is attributed not only to the single layer encryption-decryption mode employed by the cipher, but also to the ease of cryptanalysis as a result of the feasibility of carrying out an

exhaustive key search on the algorithm. It is therefore of paramount importance (as improvement on the cipher) to develop an enhanced model whereby it would amount to sheer waste of time trying to break the cipher.

This new algorithm is made up of three encryption-decryption phases whereby a plaintext message is enclosed in three cryptosystems in a technique referred to as multiple encryption (otherwise called cascade ciphering or cascade encipherment). It is designed and implemented using Java language in a NetBeans IDE Windows environment. The interface is designed in a user-friendly way such that file can be chosen or text inputted; shift values or keys inserted; message encrypted or decrypted; and the encrypted/encrypted file or message saved; all with ease. The encryption and decryption processes are based on one-to-one correspondence as below:

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

The algorithms for performing the encryption and decryption processes are highlighted as below:

Encryption Process

The transformation $\mathbf{E_i(x) = (x+k) \bmod 26}$ is used**,** where x = position of the character to be encrypted, i = encryption phase number and k = number of times x is shifted (called the shift value or key). If x+k is not in the range of 0-25, 26 is subtracted from the sum (as defined in modulo operation).

Encryption Phase 1

Step1**:** Choose a file from the directories or input a message
Step2**:** Input k_1, k_2 and k_3
Step3: Use $E_1(x) = (x+k_1) \bmod 26$ to compute $E_1(x)$

Encryption Phase 2

Step1**:** Read $E_1(x)$
Step2**:** Use $E_2(x) = (x+k_2) \bmod 26$ to compute $E_2(x)$

Encryption Phase 3

Step1**:** Read $E_2(x)$
Step2**:** Use $E_3(x) = (x+k_3) \bmod 26$ to compute $E_3(x)$

Decryption Process

The transformation $\mathbf{D_i(x) = (x-k) \bmod 26}$ is used**,** where x = position of the character to be decrypted, i = decryption phase number and k = number of times x is shifted (called the shift value). If x-k is not in the range of 0-25, 26 is added (as defined in modulo operation).

Decryption Phase 1

Step1**:** Choose a file from the directories or input a message
Step2**:** Input k_1, k_2 and k_3(the reverse of the encryption keys)
Step3**:** Use $D_1(x) = (x-k_1) \bmod 26$ to compute $D_1(x)$

Decryption Phase 2

Step1**:** Read $D_1(x)$
Step2**:** Use $D_2(x) = (x-k_2) \bmod 26$ to compute $D_2(x)$

Decryption Phase 3

Step1**:** Read $D_2(x)$
Step2:Use $D_3(x) = (x-k_3) \bmod 26$ to compute $D_3(x)$

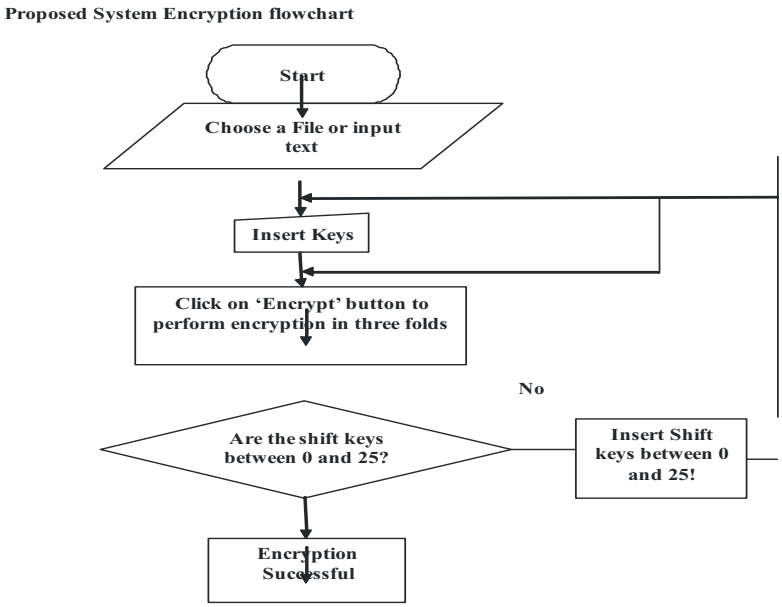

Figure 2.Flow diagram for encryption flowchart

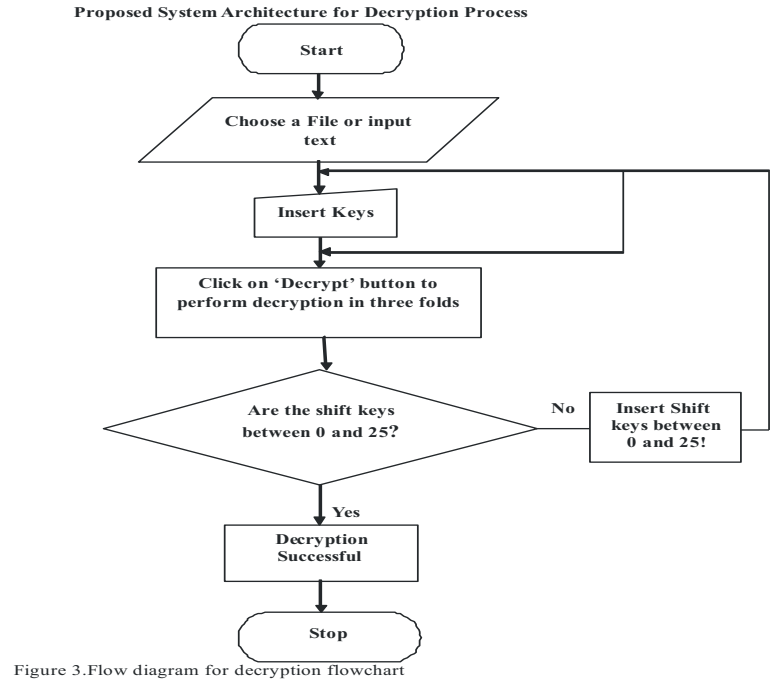

Figure 3.Flow diagram for decryption flowchart

Figure 3.Flow diagram for decryption flowchart

4.0 RESULTS AND DISCUSSION

System testing technique is used, instead of unit or integrated testing, to test the new model. When the program is run, a dialog box, which prompts the user to choose a file from the computer's directories or input a text and three shift keys, is displayed. When the file is selected and the file (AST.txt) is read and loaded into the input text area as shown in figure 4.

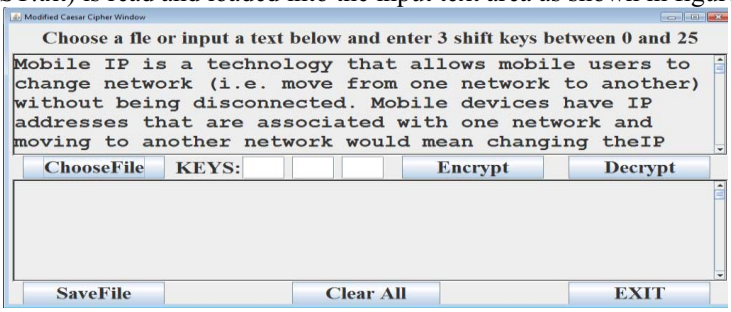

Figure 4.A Screenshot of displayed dialog box when the selected file is read into the input textarea

To encrypt the shift values are supplied (for example 1, 2, 3) and the 'Encrypt' button clicked, the file content will be encrypted three times using the entered shift values and the results displayed in the output text area as shown in figure 5.

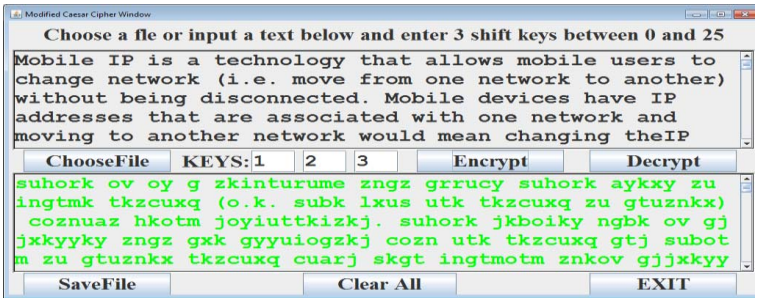

Figure 5.A Screenshot of displayed dialog box showing the ciphertext formed when the content of the chosen file is encrypted

The decryption process makes use of the entered keys in reverse to reveal the content of the message. In this way, the keys to be used for the decryption will be {3,2,1}. Figure 6 shows the result of the decryption process.

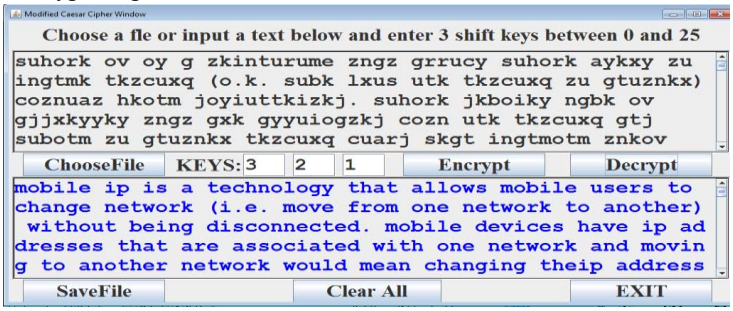

Figure 11.A Screenshot of displayed dialog box showing the plaintext formed when the content of the chosen file is decrypted.

4.1 Comparison with the Caesar Cipher

The new modified Caesar cipher provides better security than the existing Caesar cipher, although, they both use a single alphabet for their encryption and decryption processes. It is laden with features that improve the security of a cipher system.

Multiple encryption process, which is one of these features, allows for the encryption of an already-encrypted message one or more times using the same or different algorithm(Maurer & Massey, 1993). It is a technique which supports wrapping a plaintext message in various crypto-wrappers in order to hide it more from intruders, and consequently, make it more secure. This new algorithm formulated for the Caesar cipher wraps a plaintext message in three cryptosystems in order to allow for difficulty in breaking the cipher.

The Caesar cipher uses shift-by-3 rule for its encryption and decryption. However, the newly-developed algorithm uses keys between 0 and 25; it employs three different keys, one for each encryption\decryption phase. This is also a property that gives the new enhancement method an edge over the Caesar cipher.

5.0 Conclusion and Future Work

This work enhanced the Caesar cypher encryption algorithm with multiple encryption process and varying encryption key. The algorithm is tested on plain text and found more secure than the original Caesar cypher algorithm. However, the new enhancement method can only encrypt and decrypt text files in three folds. It cannot encrypt or decrypt numbers, whitespace and special characters but write them (literally) as they appear in the plaintext message or cipher text respectively. Future works in this regard will look into different ways to incorporate number and special character encryption techniques into the multiple Caesar cypher algorithm.

References

Abdulkareem, F. A.&Imran, E. I. (2014). Enhancement Caesar Cipher for Better Security. *IOSR Journal of Computer Engineering (IOSR-JCE), 16*(3), 1-5. Retrieved from http://www.iosrjournals.org

Henk, C.A. &Van, T. (2000). *Fundamentals of Cryptology: A Professional Reference and Interactive Tutorial.* Boston/Dordrecht/London: KLUWER ACADEMIC PUBLISHERS. Retrieved from https://www.hyperelliptic.org/tanja/teaching/cryptoI13/cryptodict.pdf

Jain, A., Dedhia, R. & Patil, A. (2015). Enhancing the Security of Caesar Cipher Substitution Method using a Randomized Approach for more Secure Communication. *International Journal of Computer Applications (0975 – 8887), 129*, 6-11. Retrieved from http:// www.ijcaonline.org

Khan, D. (1973). *The Codebreakers: The Story of Secret Writing* (Abridged ed.). New York: The Macmillan Company. Retrieved from http://mindguruindia.com/wp-content/uploads/2014/06/MP069_The-CodeBreakers.pdf.

Luciano, D. & Prichett, G. (1987). Cryptology: From Caesar Ciphers to Public-Key Cryptosystems. *The College Mathematics Journal, 18*, 2-17. doi:10.1.1.110.6123&rep=rep1&type=pdf.

Maurer, U. M. & Massey, J. L. (1993). Cascade Ciphers: The Importance of Being First. *Journal of Cryptology, 6*, 55-61. Retrieved from https://pdfs.semanticscholar.org/e9c5/ 9db49c7f01f 8baeacf083ea49846fe1ed25e.pdf.

Paar, C. & Pelzl, J. (2010). *Understanding Cryptography: A Textbook for Students and Practitioners.* New York: Springer. doi:10.1007/978-3-642-04101-3

Poschmann, A. Y. (2009). *Lightweight Cryptography: Cryptographic Engineering for a Pervasive World.* Ruhr-University Bochum, Germany, Faculty of Electrical Engineering and Information Technology. Bochum. Retrieved from https://eprint.iacr.org/2009/516.pdf

Purnama, B. & Rohayani, H. (2015). A New Modified Caesar Cipher Cryptography Method With Legible Ciphertext From A Message To Be Encrypted. *International Conference on Computer Science and Computational Intelligence (ICCSCI 2015).59*, pp. 195-204. Elsevier. doi: 10.1016/j.procs.2015.07.552

Rhodes-Ousley, M. (2013). *Information Security: A Complete Reference* (2nd ed.). The McGraw-Hill Companies. Retrieved from http://www.mvatcybernet.com/IT%20E-BOOKS/IT%20PDF%20Books/IT%20BOOKS/NETWORKING/INFORMATION%20SECURITY%20THE%20COMPLETE%20REFERENCE%202ND%20EDITION.pdf

Stallings, W. (2005). *Cryptography and Network Security Principles and Practices* (4th ed.). United States of America: Prentice Hall. Retrieved fromhttps://ovals.files.wordpress.com/2013/03/cryptography-and-network-security-principles-and-practices-4th-ed-william-stallings.pdf