kriativ.tech
Edition Nº 9 – 18-06-2022
ISSN Print: 1646-9976 | ISSN Online: 2184-223X |
http://www.kriativ-tech.com
10.31112/kriativ-tech-2022-06-74

Security from Caesar to Vigenère

António Santos
Assistant Professor at ISTEC – asisanto@my.istec.pt

Abstract: *Cryptography has accompanied the human being for centuries, starting with being used only by kings and emperors to communicate with their governors and military until now when its use is in all information transfers carried out digitally. The Caesar cipher was one of the first to appear and served for Roman emperors to communicate with their generals and provincial governors. It is a simple substitution number, and as such it has some limitations which, at the time it was used, were not relevant because of the illiteracy of the populations. While the Vigenère cipher brought much more security, filling the vulnerabilities of the Caeser cipher, transforming the Caeser cipher into a polyalphabetic substitution cipher.*

In this study we will compare the figures: Caeser, Shift and Vigenère; regarding your safety. It is also intended to make some modifications in order to improve and optimize the code without changing the original cipher.

Keywords: *Encryption, substitution method, Caesar cipher, Shift cipher, Vigenère cipher.*

I. Introduction

You can say that all wars have a beginning and an end, but the war that is fought in terms of security, you think you know when it started, but you don't know when it will end. Just like the "war" of the cat and the mouse, so is the daily struggle between those who defend information and those who try to access it, illegally, and then take advantage of it. Living in an information society, attacks on the networks of large companies and government departments, whenever they happen, are publicized in the media and end up having an impact on the community of potential hackers, which with all this publicity motivates them to to form and join groups that allow them to exchange information to carry out attacks on more or less vulnerable networks.

As computer networks allow computers to be interconnected and connected to the internet, the communication channels used by computers are exposed to unauthorized access [1]. When communicating with each other, they are always subject to communication failures, which have to be filled so that those involved in the communication and/or owners of the information are not harmed by seeing their information available to unauthorized users. According to Saraswat et al [2], there is a technique, cryptography, which allows the transfer of information securely between a sender and a receiver without the possibility of intervention by external elements.

Balogun [3] defines cryptography as the science and art of encrypting and decrypting data using some special techniques. Kahate [4] defines it as the art of obtaining security by encoding messages to make them unreadable, while Paar [5] goes further, stating that nowadays cryptography is the science of secret writing with the purpose of hiding the meaning of a message. Encryption is not new today but has been used for thousands of years to help provide confidential communications between mutually trusted parties [6]. Referring to Katz and Lindell [7], they mention that historically, the biggest users

of cryptography were military organizations and governments, nowadays, it is everywhere.

According to Aggarwal [8], encryption is divided into two categories depending on the type of security keys used to encrypt/decrypt the data, these techniques are: asymmetric and symmetric encryption. For Stallings [8], symmetric cryptography is a form of cryptosystem in which encryption and decryption are performed using the same key. Whereas, that Shrivastava, et al. [10] defines asymmetric cryptography, also called public-key cryptography, uses a pair of keys to encrypt and decrypt.

Encryption involves an algorithm and a key to convert the information into a format that is incomprehensible to anyone, except for the interlocutors of the communication [2].

Encryption in addition to being of the two types mentioned above: symmetric and asymmetric; also resort to two techniques: transposition and substitution [11]. Transposition ciphers scramble the message letters in a way designed to confuse the attacker, but can be recomposed by the intended recipient [12], while Singh [11] writes that in transposition, the message letters are simply rearranged, generating effectively an anagram. Kahate [4] mentions that in the substitution technique, the characters of a plain text message are replaced by other characters, numbers or symbols.

Ciphers are forms of simple algorithms that promote the encryption/decryption of information. Many of the classical ciphers were very easy to decipher. Nowadays, the principles that were developed in the old ciphers are used in the current ciphers, however, with evolution, a lot of complexity was implemented in order to make the message safer and more difficult to break. Bearing in mind what has been written above, we can say that a cipher transforms plaintext into ciphertext in encryption and from ciphertext to plaintext in decryption. The plain text is the text understood by those who use the same communication code (language and alphabet), that is, and according to Joshua [13], he states that the plain text is the text of the message in ordinary language. Otherwise, Kahate [4] defines this as the text that can be understood by anyone who knows the language, as long as the message is not coded in any way. Also, Kahate [4] writes that ciphertext is the result when plaintext is encoded using any suitable scheme.

II. Background

Cryptography in its origin is confused with the different civilizations, because these, in order to communicate among their scattered throughout the empires, needed a form of communication accessible only to the intervening ones, which at that time was between the central power, governors and generals. The roots of cryptography are found in civilizations: Roman and Egyptian [14]. Its existence over the ages was generally confined to diplomatic and military reasons, where it was used to hide information communicated over secure and unsecured lines [15]. Currently, cryptography is considered a branch of mathematics and computer science and is closely affiliated with information theory, computer security and engineering [17]. And it can be found everywhere, as its need is high today because, as Singh and Sen [16], claim that this modern age is dominated by paperless offices, mail messages, cashless transactions and stores. of virtual departments. On the other hand, there is a great need for data exchange over the internet and various confidential information such as banking transactions, credit information and confidential data are transferred over the internet [17].

a) From Caesar to Vigenère

The Caesar cipher is one of the earliest known cryptographic systems, and was used by Julius Caesar [18]. Although, according to Joshua [13], Caesar was probably not the original inventor of what is now called the Caesar cipher, but he certainly made it popular, hence his name. The Caesar cipher uses simple substitution, where there is only a three-position right shift of each character, in the alphabet, in the plaintext to obtain the letters in the ciphertext [19]. That is, to encrypt, advance three positions in the alphabet: A becomes D, B becomes E, and so on until W becomes Z, and then X becomes A, the Y in B and finally, the Z in C; which analytically can be represented by:

E[x]= x+3 mod(26),

where, mod(26) represents the remainder of the division of (x+3) by 26, ex = 0, 1, 2, …, 25 that represent the set composed by the alphabet {A,B,C,D, … ,Z}.

To decrypt, the inverse of encryption is used, that is, decryption is performed by replacing each character of the ciphertext by the character resulting from the shift of three positions to the left [12]. What can analytically be represented by:

D[x]=x-3 mod(26),

where mod(26) is the remainder of division by 26. From Santos and Junior [20], the following can be deduced: D[x]=x-3 mod(26)= x+ (26-3) mod (26)= x+ 23 mod(26). Which at the programming level is more efficient because it only uses a function that does the encryption and decryption function.

Taking as a basis E[x]= x+3 mod(26) and if E[x] and ex are integers then E[x] – (x+3)= 0 mod(26), so for any 0<k< 25, if E[x]= x+k mod(26) then E[x] – (x+k)= 0 mod(26). From this it follows that the displacement 3 of the Caesar cipher can be replaced by a value k, and 0<k<25. This cipher is called the Shift cipher and can be seen as a keyed variant of the Caesar cipher. Specifically, in the shift cipher the key k is a number between 0 and 25 [7]. This is in line with the definition by Stinson and Patherson [21], in which they define that if a and b are integers, and m is a positive integer. Then write a ≡ b (mod m) if m divides b - a. The analytical form a ≡ b (mod m) and reads “a is congruent with b modulo m.” The integer m is called the modulus. If we replace m by 26, then a ≡ b (mod 26). Musa [22] refers that in the displacement cipher k is presented as a secret key with a value between 0 and 25, in the case of the Ceaser cipher this value is three.

The mono-alphabetic substitution cipher. In the displacement cipher, the key defines a map of each letter of the alphabet (plaintext) to some letter of the alphabet (ciphertext), where the map is a fixed displacement determined by the key [7]. Bearing in mind that a text to be encrypted/decrypted is not just a character but a sequence of characters, then let's take $P=P_1$, P_2,P_3, …, P_n the set of characters to be encrypted and $C=C_1$, C_2, C_3, …, C_n the encrypted characters. If $C=P+k \bmod(26)$ and $0<k<25$, then $C_1=P_1+k \bmod(26)$, $C_2=P_2+k \bmod(26)$, $C_3=P_3+k \bmod(26)$, …, $C_n=P_n+k \bmod(26)$.

Going back to $C=P+k \bmod(26)$, if we replace ok with $K=k_1, k_2, k_3,\ldots, k_m$, where $m\leq n$, then we have that $C_1=P_1+k_1 \bmod(26)$, $C_2=P_2+ k_2 \bmod(26)$, $C_3=P_3+k_3 \bmod(26)$, …, $C_m=P_m+k_m \bmod(26)$, if m=n ended here, if m<n, $C_{m+1}=P_{m+1}+k_{m+1} \bmod(26)$ up to $C_n=P_n+k_{n-m \bmod(m)} \bmod(26)$. It follows that the general case would be $C_i=P_i+k_{i \bmod(m)} \bmod(26)$, $0<i<n$ and $m\leq n$, in the case of encryption. In the case of decryption, it would be $P_i=C_i-k_{i \bmod(m)} \bmod(26)$, which is in line with Stallings [23] when defining the Vigenere cipher.

The Vigenère cipher, a special case of the shift cipher, also called a polyalphabetic shift cipher, works by applying several independent instances of the shift cipher in sequence [7]. The key is now seen as a string, and encryption is done by shifting each plaintext character by the amount indicated by the next character in the key, wrapping the key when necessary. According to Stamp and Low [12] they classify the Vigenere cipher as a classic polyalphabetic substitution cipher, which uses several simple substitutions to encrypt a message. A keyword can then be used to determine which of the cipher alphabets to use at each position in the text. In this way, all simple "shift-by-n" substitutions are readily available for use [12].

III. Evolution of Figures in Algorithms

In order to analyze the evolution of the ciphers, they will be programmed with the Python language because this language is more flexible for this type of ciphers. The objective of this programming is to verify the evolution of the degree of difficulty of these and the impact of the evolution on the security of the results.

a) Caesar's Cipher

The Caesar cipher is the simplest to program, just read the text and move three positions to the right for encryption and 23 for decryption.

```
textopl = input("Enter the message to encrypt: ")
textocf=encrypta(textopl.upper(), 3)
print("Encrypted message:",textocf)
```

To decrypt it works in the same way as encrypt, it asks for the offset value and the ciphertext and calls the encryption function with the phrase and offset 26-3=23, as follows:

```
textocf=input("Enter encrypted message:",textocf)
plano=encrypt(textocf.upper(),  23)
print("Plain text:", plano)
```

The encrypt function that will encrypt/decrypt the message, receiving the uppercase text and the key. Then it takes character by character character of the sentence, transforms it into the ASCII code with ord() subtracts 65, sum ok which is in this case 3 or 23, if it is encryption or decryption, apply module 26 and add 65, then apply the conversion of character ASCII code with the chr() command.

```
def encrypt(texte, k):
    textoc=''
    j=0
    for i in texte:
        textoc=textoc +chr(((ord(i)-65+k)%26 )+65)
return textoc
```

b) Shift Cipher

Regarding the Shift cipher, which in many cases is called Caesar's Shift cipher, the only difference from the Caesar cipher is that it also allows the insertion of a key value (k), and the encryption/decryption is carried out in the same way. way, that is, instead of advancing k positions to the right or to the left, depending on the encryption or decryption, in the case of encryption, we advance 0<k<25 positions.

```
k = int(input("Enter the shift value: "))
textopl = input("Enter the message to encrypt: ")
textocf=encrypt(textopl .upper(),  , k)
print("Encrypted message:",textocf)
```

To decrypt works in the same way as encrypt, it asks for the offset value and the ciphertext, calls the encryption function with the phrase and the 26-k offset, as follows:

```
k = int(input("Enter the shift value: "))
textocf = input(" ("Enter encrypted message:")
plano=encrypt(textocf. upper(),  ,26-k)
print("Plain text:", plano)
```

The encrypt function that will encrypt/decrypt the message. It is identical to the previous section, receiving the text in capital letters and the key. Then take a character from the sentence, transform it into ASCII code with ord(), subtract 65, add the k, apply modulo 26 and add 65, then apply the conversion from ASCII code to character with the chr() command.

```
def encrypt(texte, k):
    textoc=''
    j=0
    for i in texte:
        textoc=textoc +chr(((ord(i)-65+k)%26 )+65)
return textoc
```

c) Vigenère Cipher

Finally, in the case of the Vigenère cipher, the user must be asked for the text and the key, then an array representing the key is built with the same length as the entered text, in which case, in the case of a key with a lower length than the inserted text, the key is repeated until the array is completed.

When you want to encrypt a text using the Vigenere cipher, you first introduce the key and the text, as you are going to operate the key with the text, it must have the same number of characters as the text, for that you will create a array with the length of the text, repeating the key as many times as necessary, then resorting to a function that encrypts the text with capital letters.

```
key = input("  Enter the Key: ")
textopl = input("  Enter the message to encrypt: ")
k=0
for i in textopl:
```

```
            keya=keya+key[k]
            k+=1
            if k==len(key):
              k=0
textocf = encrypt(textopl.upper(), keya.upper())
print("  Encrypted message:",textocf)
```

Decryption needs a little more code, repeating the filling part of the array that represents the key, then you will find the character that is complementary to the key: chr((26-ord(i)-65)% 26+65); that is, the value of the ASCII code of the character is removed from 26 by subtracting 65 (first character (A) in the ASCII code), followed by the calculation of the modulus of this operation, 65 is added and the ASCII code is converted to character. This operation allows using the same function to encrypt and decrypt, as a way of optimizing the code.

```
key = input("  Enter the Key: ")
textocf = input("  Enter the message to encrypt: ")
k=0
for i in textopl:
      keya=keya+key[k]
      k+=1
      if k==len(key):
        k=0
for i in keya.upper():
   keyc=keyc+chr((26-ord(i)-65)%26+65)
plano = encrypt(textocf.upper(),keyc.upper())
print("  Plain text:", plano)
```

The encrypt function, receives the text (phrase) and the key (keyb), then in a cycle that goes from 0 to the length of the text, it adds the value of the character of the text with the value of the character of the key converted to ASCII subtracting 65 to each one, module 26, then the value 65 is added and this value is converted to an alphabet character, which is returned at the end of the function execution.

```
def encypt (frase, keyb):
   textoc=''
   for i in range(len(frase)):
        c=((ord(frase[i])-65)+(ord(keyb[i])-65))%26+65
        textoc = textoc+chr(c)
   return textoc
```

As can be seen, the same line of reasoning was followed for programming the methods, as a way of demonstrating the evolution between the three techniques.

IV. Results

The execution of the methods made it possible to observe the safety of the techniques, using a common text for all the techniques in order to better analyze the results.

```
Enter the message to encrypt: Mississippi
Encrypted message: PLVVLVVLSSL
```

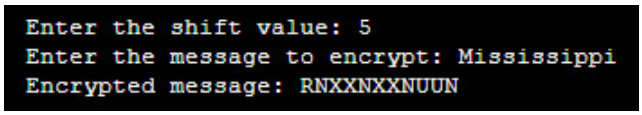
Picture 1: Encrypting the a Word with the Caesar Cipher

When the Caesar cipher technique is used, in which the original version indicated that a displacement of three positions to the right was carried out, the word mississippi became the word PLVVLVVLSSL, before a result like this if one resorted to the method of brute force would be easy to decrypt. As can be seen, in figure 1, the repeated characters in the simple string become repeated characters in the ciphertext.

```
Enter encrypted message: PLVVLVVLSSL
Plain text: MISSISSIPPI
```

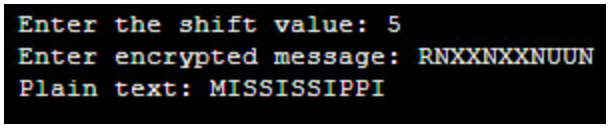
Picture 2: Decrypting a Word with the Caesar Cipher

By Picture 2, the ciphertext PLVVLVVLSSL gave rise to the original plaintext. These types of ciphers with current knowledge and computers would be easy to decrypt, even without the key. This is in line with Schneier [24], who claims that simple substitution ciphers, like this one, can be easily broken because the cipher does not hide the underlying frequencies of different characters in the plaintext.



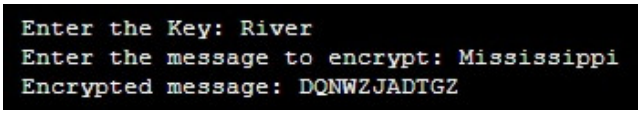
Picture 3: Encrypting the a Word with the Shift Cipher

In picture 3, the shift cipher is used to encrypt the Mississippi word with a key of value 5, and a resulting word is observed: RNXXNXXNUUN. As can be seen repeated characters are transformed into repeated characters, hence there is a high degree of ease in the possibility of decryption by unauthorized parties. This is in line with Musa [22], who states that in this cipher as the secret key is a value between 0 and 25, in this case five, a brute force attack can break the cipher scheme in a short time.



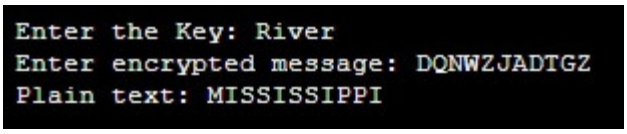
Picture 4: Decrypting a Word with the Shift Cipher

The decrypted text is as expected. That is, the most common ciphertext word probably matches the plaintext. In this type of cipher, we can obtain additional statistical information using digraphs (pairs of letters) and common trigraphs (triples). This type of statistical attack on a simple substitution is very effective [12].



Picture 5: Encrypting the a Word with the Vigenère Cipher

In picture 5, we have the encoding with the Vigenère cipher, which uses a polyalphabetic substitution, which is reflected in the frequency with which the characters appear, that is, repeated characters are not transformed into repeated characters as in the previous cases (Caesar and Shify). This type of transformation provides relative security if the key is not known.



Picture 6: Decrypting a Word with the Vigenère Cipher

Finally, picture 6 represents the transformation of the encoded text into the plain text, and it is observed that it occurred as expected.

The Vigenère cipher at first glance appears to be a fairly secure cipher to be used in text encryption, albeit with some limitations. As stated by Katz and Lindell [7] a first observation when attacking the Vigenere cipher is that, if the key length is known, attacking the cipher is relatively easy. Therefore, the generic simple substitution attack discussed above will not work on a polyalphabetic substitution [12]. However, the Vigenère cipher is vulnerable to a slightly more sophisticated statistical attack [12].

V. Conclusion

Cryptography has evolved over time, being currently almost indecipherable as technologies have evolved and with these more complex algorithms and more transformations can be used. The ciphers mentioned in this article were created in times when there were no computers, so all the calculation was done manually, which meant spending more time on their encryption/decryption.

Caesar's cipher is undoubtedly the simplest and easiest to decipher, but it was also one of the first, and at that time it benefited from people's illiteracy, in addition to serving for communications between the emperor, provincial governors and generals. The Shift cipher is a generalization of the Caesar cipher, allowing the sender to choose a key between 0 and 25, which he will have to share with the recipient. As for the Vigenère cipher, it is an evolution of the Shift cipher, and it can be seen that the keyword is composed of characters that can then be converted into integers between 0 and 25, which will then be added to the characters of the same position but in plain/cipher text.

The most vulnerable security level is Ceaser followed by Shift, which can be easily decrypted by someone who does not know the password, using statistical techniques, such as

character frequency and brute force; whereas with the Vigenère cipher, someone who doesn't know the key word has much more difficulty in achieving decryption. As the Vigenère cipher does not allow repetition of repeated characters in the plaintext, it is only with more or less robust statistical techniques that the decryption of an encrypted text can be forced.

VI. References

[1] Jain, A., Dedhia, R. and Patil, A. (2015) "Enhancing the Security of Caesar Cipher Substitution Method using a Randomized Approach for more Secure Communication", International Journal of Computer Applications, Vol. 129 No.13, pp: 144-150,

[2] Saraswat, A., Khatria, C., Sudhakara, Thakrala, P., Biswasa, P.. (2016) An Extended Hybridization of Vigenere and Caesar Cipher Techniques for Secure Communication, 2nd International Conference on Intelligent Computing, Communication & Convergence (ICCC-2016), Procedia Computer Science 92 pp 355 – 360.

[3] Balogun, A.O. , Sadiku, P. O., Mojeed, H. A., Raifu, H. A. (2017) Multiple Caesar Cypher Encryption Algorithm. ABACUS, (Mathematics Science Series) Vol. 44, No 2. pp 250-258.

[4] Kahate, Atul (2003). Cryptography and network security, Tata McGraw-Hill, New Delhi.

[5] Paar, Christof\ and Pelzl, Jan (2010). Understanding Cryptography: A Textbook for Students and Practitioners. Springer-Verlag, Berlin.

[6] Stinson, D.R. and Paterson, M. B. (2018) Cryptography: Theory and Practice, Fourth Edition. Textbooks in Mathemátics, CRC Press.

[7] Katz, Jonathan and Lindell, Yehuda (2015). Introduction to Modern Cryptography. Second Edition, Taylor & Francis Group, LLC, CRC Press.

[8] Aggarwal, Surabhi (2016). A Review on Enhancing Caesar Cipher, International Journal of Research Science & Management, 3(6).

[9] Stallings, William (2017). Cryptography and network security: Principles and Practice, Fifth Edition, Prentice Hall, New York.

[10] Shrivastava, M., Jain, M, and Singh, P. (2016) Content Based Symmetric Key Algorithm, International Conference on Computational Modeling and Security, Procedia Computer Science 85, pp. 222 – 227.

[11] Singh, Simon (1999).The Code Book, Anchor Books: The Science of Secrecy from Ancient Egypt to Quantum Cryptography. Anchor Boks. New York.

[12] Stamp, Mark and Low, Richard M.(2007). Applied Cryptanalysis - Breaking Ciphers in the Real World. Wiley-Interscience, John Wiley & Sons, Inc..

[13] Holden, Joshua (2017), The Mathematics of Secrets: Cryptography from Caesar Ciphers to Digital Encryption, Princeton University Press, New Jersey.

[14] Kumari, Sarita (2017) A research Paper on Cryptography Encryption and Compression Techniques, International Journal Of Engineering And Computer Science. Volume 6 Issue 4Page No. 20915-20919. DOI: 10.18535/ijecs/v6i4.20.

[15] Luciano, D. and Priche,, G. Cryptology: From Caesar Ciphers to Public-Key Cryptosystems, The College Mathematics Journal, Vol. 18, No. 1 (Jan., 1987), pp. 2-17

[16] Omolara O.E., Oludare A.I. and Abdulahi S.E..(2014) Developing a Modified Hybrid Caesar Cipher and Vigenere Cipher for Secure Data Communication, Computer Engineering and Intelligent Systems Vol.5, No.5, pp 34- 47

[17] Singh, Pooja and Sen, Pintu (2017). Enhancing Security of Caesar Cipher Using Divide and Conquer Approach, International Journal of Advance Research in Science and Enginheering, Volume 6, special issue (02).

[18] Sinkov, Abraham (1966), Elementary Cryptanalysis - A Mathematical Approach, Fifth Printing, The Mathematical Association of America. Washington, USA.

[19] Mollin, Richard A. (2007) An Introduction to Cryptography, Second Edition. Discrete Mathematics and its Applications, Chapman & Hall/CRC, California.

[20] Santos, Antonio and Vasconcelos Junior, Renato (2021). Improving Caesar Cipher for greater security. kriativ-tech (9) ,DOI: 10.31112/kriativ-tech-2021-10-54.

[21] Stinson, Douglas and Paterson, Maura. (2018) Cryptography: Theory and Practice fourth edition. CRC Press, CRC Press LLC. Florida.

[22] Musa, Sarhan M. (2018). Network Security and Cryptography: A Self-teaching Introduction. Mercury Learning & Information. Virgínia.

[23] Stallings, William (2017). Cryptography and network security: Principles and Practice, Seventh Edition. Pearson. Essex.

[24] Schneier, Bruce (1996). Applied Cryptography, Second Edition. John Wiley & Sons.