Texts Ciphering by using Translation Principle

Ali H. Al-Nuaimi
Dept. of Computer Sci./College of Education for Pure Science(Ibn Al- Haitham) / University of Baghdad

Received in :22April 2013, Accepted in : 19 May 2013

Abstract

The proposed algorithm that is presented in this paper is based on using the principle of texts translation from one language to another, but I will develop this meaning to cipher texts by using any electronic dictionary as a tool of ciphering based on the locations of the words that text contained them in the dictionary. Then convert the text file into picture file, such as BMP-24 format. The picture file will be transmitted to the receiver. The same algorithm will be used in encryption and decryption processing in forward direction in the sender, and in backward direction in the receiver. Visual Basic 6.0 is used to implement the proposed cryptography algorithm.

Keywords: Encryption, Decryption**,** Cryptography, Cipher text, Plaintext**,** Breakable**,** Symmetric ciphers, Asymmetric ciphers**.**

Introduction

The concept of information will be taken to be an understood quantity. To introduce cryptography, an understanding of issues related to information security in general is necessary. Information security manifests itself in many ways according to the situation and requirement. Regardless of who is involved, to one degree or another, all parties to a transaction must have confidence that certain objectives associated with information security have been met.

Cryptography is the study of mathematical techniques related to aspects of information security such as confidentiality, data integrity, entity authentication, and data origin authentication. Cryptography is not the only means of providing information security, but rather one set of techniques [1].

Cryptology: This is the study of techniques for ensuring the secrecy and/or authenticity of information. The two main branches of cryptology are cryptography, which is the study of the design of such techniques; and cryptanalysis, which deals with defeating such techniques, to recover information, or forging information that will be accepted as authentic.

The requirements of information security within an organization have undergone two major changes in the last several decades. Before the widespread use of data processing equipment, the security of information felt to be valuable to an organization was provided primarily by physical and administrative means. An example of the former is the use of rugged filing cabinets with a combination lock for storing sensitive documents. An example of the latter is personnel screening procedures used during the hiring process. With the introduction of the computer, the need for automated tools for protecting files and other information stored on the computer became evident. This is especially the case for a shared system, such as a time-sharing system, and the need is even more acute for systems that can be accessed over a public telephone network, data network, or the Internet. The generic name for the collection of tools designed to protect data and to thwart hackers is computer security. The second major change that affected security is the introduction of distributed systems and the use of networks and communications facilities for carrying data between terminal user and computer and between computer and computer. Network security measures are needed to protect data during their transmission. In fact, the term network security is somewhat misleading, because virtually all business, government, and academic organizations interconnect their data processing equipment with a collection of interconnected networks. Such a collection is often referred to as an internet, and the term internet security is used [2].

Data that can be read and understood without any special measures is called plaintext or clear text. The method of disguising plaintext in such a way as to hide its substance is called encryption. Encrypting plaintext results in unreadable gibberish are called ciphertext. You use encryption to make sure that information is hidden from anyone for whom it is not intended, even those who can see the encrypted data. The process of reverting ciphertext to its original plaintext is called decryption. Figure 1 shows this process.

Cryptography is the science of using mathematics to encrypt and decrypt data. Cryptography enables you to store sensitive information or transmit it across insecure networks (like the Internet) so that it cannot be read by anyone except the intended recipient. While cryptography is the science of securing data, cryptanalysis is the science of analyzing and breaking secured communication. Classical cryptanalysis involves an interesting combination of analytical reasoning, application of mathematical tools, pattern finding, patience, determination, and luck. Cryptanalysts are also called attackers. Cryptology embraces both cryptography and cryptanalysis [3].

Cryptography is used to hide information. It is not only used by spies but for phone, fax and e-mail communication, bank transactions, bank account security, PINs, passwords and credit card transactions on the web. It is also used for a variety of other information security issues including electronic signatures, which are used to prove who sent a message [4].

Security Goals

There are three very important aspects of any computer-related system: Confidentiality, Integrity and Availability, **Confidentiality** ensures that computer-related assets are accessed only by authorized parties. That is, only those who should have access to something will actually get that access. By "access," we mean not only reading but also viewing, printing, or simply knowing that a particular asset exists. Confidentiality is sometimes called secrecy or privacy. **Integrity** means that assets can be modified only by authorized parties or only in authorized ways. In this context, modification includes writing, changing, changing status, deleting, and creating. **Availability** means that assets are accessible to authorized parties at appropriate times. In other words, if some person or system has legitimate access to a particular set of objects, that access should not be prevented. For this reason, availability is sometimes known by its opposite, denial of service [5].

Principles of Cryptography

The field of cryptography is huge and covers many methods and approaches. At the most basic level, these methods can be classified into codes and ciphers. A code is a short symbol or word that replaces an entire message. Codes are secured but are not general purpose. Before a spy is sent to a foreign country he and his runner may agree on a set of codes. The words happy and sad used by the spy in otherwise-innocuous sentences may indicate good and bad economies in the foreign country, whereas deep and shallow may be codes for success and failure of the spy’s mission. It is easy to see why the use of codes is limited, but it is also true that codes can be broken. If the same spy sends many messages that use the same codes, then clever codebreakers who intercept the messages may eventually guess the meaning of certain codes, and then test their guesses by applying them to future messages to see if the guesses make sense. A cipher is a rule that tells how to scramble (encrypt) data in a nonrandom way, so it can later be unscrambled (decrypted). Perhaps the simplest example of a cipher is to replace each letter with the one following it (cyclically) n positions in the alphabet. This is the well-known Caesar cipher [6].

Cryptography can be strong or weak, as explained above. Cryptographic strength is measured in the time and resources it would require to recover the plaintext. The result of strong cryptography is ciphertext that is very difficult to decipher without possession of the appropriate decoding tool. How difficult? Given all of today’s computing power and available time even a billion computers doing a billion checks a second, it is not possible to decipher the result of strong cryptography before the end of the universe. One would think, then, that strong cryptography would hold up rather well against even an extremely determined cryptanalyst. Who’s really to say? No one has proven that the strongest encryption obtainable today will hold up under tomorrow’s computing power [3].

Encryption Algorithms

The cryptosystem involves a set of rules for how to encrypt the plaintext and how to decrypt the ciphertext. The encryption and decryption rules, called **algorithms**, often use a device which is called a **key**, denoted by K, so that the resulting ciphertext depends on the original plaintext message, the algorithm, and the key value. We write this dependence as C = E(K, P). Essentially, E is a set of encryption algorithms, and the key K selects one specific algorithm from the set. We see later in this chapter that a cryptosystem, such as the Caesar cipher, is keyless but that keyed encryptions are more difficult to break [5].

The encryption and decryption algorithms are collectively called cryptographic algorithms (cryptographic systems or cryptosystems). Both encryption and decryption processes are controlled by a cryptographic key, or keys. In a symmetric (or shared-key) cryptosystem, encryption and decryption use the same (or essentially the same) key; in an asymmetric (or public-key) cryptosystem, encryption and decryption use two different keys: an encryption key and a (matching) decryption key.The encryption key can be made public (and hence is also called public key) without causing the matching decryption key being discovered (and thus a decryption key in a public-key cryptosystem is also called a private key). Figure 3 illustrates a simplified pictorial description of a cryptographic system [7].

Breakable Encryption

An encryption algorithm is called breakable when, given enough time and data, an analyst can determine the algorithm. However, an algorithm that is theoretically breakable may in fact be impractical to try to break. Two other important issues must be addressed when considering the breakability of encryption algorithms. First, the cryptanalyst cannot be expected to try only the hard, long way. Second, estimates of breakability are based on current technology. An enormous advance in computing technology has occurred since 1950. Things that were infeasible in 1940 became possible by the 1950s, and every succeeding decade has brought greater improvements. A conjecture known as "Moore's Law" asserts that the speed of processors doubles every 1.5 years, and this conjecture has been true for over two decades. It is risky to pronounce an algorithm secure just because it cannot be broken with current technology, or worse, that it has not been broken yet [5].

The entire field of cryptography is based on an important assumption, namely, that some information can be kept and disseminated securely, accessible only by authorized persons. This information is the key used by an encryption algorithm. An important principle in cryptography, due to the Dutch linguist Auguste Kerckhoffs von Nieuwenhoff, states that the security of an encrypted message must depend on keeping the key secret. It should not depend on keeping the encryption algorithm secret. This principle is widely accepted and implies that there must be many possible keys to an algorithm; the keyspace must be very large. The Caesar algorithm, for example, is very weak because its keyspace is so small. Notice that a large keyspace is a necessary but not a sufficient condition of security. An encryption algorithm may have an immense keyspace but may nevertheless be weak [6].

Types of Cryptography

Two common types or the forms of the cipher are the symmetric and the asymmetric ciphers.

Symmetric Ciphers: A type of cipher used for the data encryption and made it unreadable object for the other by using the same or the identical type of algorithmic key for encrypted or decrypted, the required data is called the Symmetric Ciphers. The keys that reused must be the systems having symmetric ciphers can share these keys with the decryption modules so, that is whey they are also called as the shared symmetric ciphers. The working performance of the symmetric ciphers is quite faster as compared to the asymmetric ciphers. So, symmetric ciphers are the better way to protect the private or the personal information.

The working of the symmetric cipher is quite simple and not very complicated because there is only one key which is used by both of the parties to encrypt the data or the information. So first of the before encrypting that data between two places, two parties have to agree to use the same key algorithm for the sake of symmetric data encryption. Then one party sends the data using the key then this key will exchange with the other and then by using the same key receiver decrypt the data easily. But there are some security problems while exchanging the key between the parties.

There are two different types of symmetric ciphers are used to encrypt or decrypt the message or the information between the two parties such as

1. Stream Ciphers

2. Block Ciphers

Stream cipher can use the fragment to encrypt while the block ciphers have ability to deal it like a single part.

Asymmetric Ciphers: Another important type of encryption with the help of cipher is the asymmetric cipher, it is quite similar to that is the symmetric ciphers but the only different between the symmetric and the asymmetric cipher is that the key used for the encryption and decryption are not identical. So, there are two different forms of keys used in the process, one is called as public key and other is private. Users can use the public key locally for sharing whereas the private key is the secret one. The systems are called as asymmetric because when users using this cipher to transmit text so he uses public key to encrypt and the receiver uses private key to decrypt that is why the system is called so.

As it is cleared from the name and the introduction that two different types of keys are used in the encryption which are carried out between two parties through this asymmetric cipher. All the working of the asymmetric cipher depends upon the receivers because they do lot of software changes in the generation of the two keys, then one key that is public sends to the sender for encryption and other remains there at receiving end for decryption. These are some strategies of working of the asymmetric ciphers. They are little slower than the symmetric ciphers but can be faster by fusion with SC. [8].

A cipher that uses a pair of keys is a public key and a private key, for encryption and decryption. Also called an asymmetric algorithm. Ciphers use symmetric keys, the same key is used to encrypt and decrypt a message [6].

Asymmetric Encryption is a form of cryptosystem in which encryption and decryption are performed using two different keys, one of which is referred to as the public key and one of which is referred to as the private key. Also known as public-key encryption. Summarizes some of the important aspects of symmetric and public-key encryption. To discriminate between the two, we refer to the key used in symmetric encryption as a secret key. The two keys used for asymmetric encryption are referred to as the public key and the private key. Invariably, the private key is kept secret, but it is referred to as a private key rather than a secret key to avoid confusion with symmetric encryption.

Symmetric encryption is a form of cryptosystem in which encryption and decryption are performed using the same key. It is also known as conventional encryption. Symmetric encryption transforms plaintext into ciphertext using a secret key and an encryption algorithm. Using the same key and a decryption algorithm, the plaintext is recovered from the ciphertext.

The two types of attack on an encryption algorithm are cryptanalysis, based on properties of the encryption algorithm, and brute-force, which involves trying all possible keys. Traditional (precomputer) symmetric ciphers use substitution and/or transposition techniques. Substitution techniques map plaintext elements (characters, bits) into ciphertext elements. Transposition techniques systematically transpose the positions of plaintext elements [2].

Let us compare symmetric-key and asymmetric-key cryptography. Encryption can be thought of as electronic locking; decryption as electronic unlocking. The sender puts the message in a box and locks the box by using a key; the receiver unlocks the box with a key and takes out the message. The difference lies in the mechanism of the locking and unlocking and the type of keys used. In symmetric-key cryptography, the same key locks and unlocks the box. In asymmetric- key cryptography, one key locks the box, but another key is needed to unlock it.

Three types of keys in cryptography: the secret key, the public key, and the private key. The first, the secret key, is the shared key used in symmetric-key cryptography. The second and the third are the public and private keys used in asymmetric-key cryptography [9].

The secret key is also input to the encryption algorithm. The key is a value independent of the plaintext and of the algorithm. The algorithm will produce a different output depending on the specific key being used at the time. The exact substitutions and transformations performed by the algorithm depend on the key [2].

The Proposed Cryptographic Algorithm

The proposed algorithm consists of two parts, the first for encryption, and the second for decryption algorithm.
In this paragraph, I will present the encryption algorithm

```
1:      Read the text to be encrypted
2:      Define variables I,J,X,Y,Z,Index,Count As integer, Word as string.
3:      Let I,J,Index equal to one:Word="".
4:      Compute the number of words in the text and save it in variable C
5:      Define Array named Table(C,3) as integer
6:      Repeat
6.1:        If  substring (Text,I,J)<>" " then
6.2:            J=J+1
6.3:        Else
6.4:            Word= substring (Text,I,J)
6.5:            Search the dictionary to find word
6.6:            Save the page number in X, Line number in Y, and sequence in Z
6.7:            Table(Index,1)=X, Table(Index,2)=Y, Table(Index,3)=Z
6.8:            Index=Index+1
6.9:            C=C-1
6.10:           J=I
6.11:       Endif
7:      Until C=0
8:      End
```

Note: If the word repeated in the text I will save the previous location in the table only.

But, in this paragraph, I will present the decryption algorithm.

```
1:     Define variables Text, Word  as string: Count, Index as integer: X, Y,Z as integer.
2:     Let index=1: Count=number of lines in the array named count.
3:     Repeat
3.1:       Let X=Table(index,1): Y= Table(index,2): Z=Table(index,3)
3.2:       Go to the dictionary and save the word located in page number X, line number
           Y, and in sequence number in Z in the variable named Word
3.3:       Text=Text+Word+" "
3.4:       Index=Index+1
3.5:       Word=""
4:     Until  Count=Index
5:     End
```

Results of the proposed cryptographic algorithm

Suppose I have the following text to de encrypted as a plaintext.

"This section presents some of the attacks that can be made against the use of cryptography"

I will use Al Mawrid dictionary to convert the above text, as shown in table 2.

From table 1, I will send columns 3,4, and 5 (page no., line no., and seq.) only to the target as image file by using the following steps :

1- Sum number of lines in table 1, in this case is 402, save the result in variable named Sum, in this case Sum=402
2- Multiply line account by two ,and add the result to summation of line number, in this case (402+16*2)=434 and store the result in variable named Size, Size=418
3- Compute the integer square root of the above result, and save it in variable named Column, in this case Column =20.
4- Divide Size by Column and save the result in variable named Line, in this case Line =round(434 /20), the result is 22, Line =22.
5- Build an array of the image, named Target(Line, Column)
6- Store the data from table 2 to the target array as image file.

The result from the previous steps is array named target contains the data from the table 1, as shown in table 2.
After that, save table 2 as image file in bmp-24 format. Figure 4 shows the transmitted image. The image file is send to the target, the receiver convert the image file to text file and apply the decryption algorithm to create the source message.

Conclusions

The proposed algorithm is a symmetric cipher, and the type of it is a block cipher. The user use any type of dictionaries or books, electronic or not, and web sites as a tool to convert the message. This algorithm didn't require saving spaces in the message, also can be used in different ways. The proposed algorithm is dynamic cryptography and depends on the type of tools that are used.

References

1- Alfred, J. Menezes; Paul C; Van Oorschot; Scott A. and Vanstone , R. L. (1997), Handbook of Applied Cryptography, Second Edition , CRC Press LLC, USA.
2- William Stallings, (2005), Cryptography and Network Security Principles and Practices, Fourth Edition, Prentice Hall, USA.
3- Corporation PGP , (2003), An Introduction to Cryptography, Network Associates, Inc. and its Affiliated Companies , USA.
4- Edward Schaefer, An introduction to cryptography and Cryptanalysis, http://math.scu.edu/~eschaefe/crylec.pdf.
5- Charles, P. Pfleeger, and Shari Lawrence Pfleeger, (2006), Security in Computing, Fourth Edition, Prentice Hall PTR, ,USA.
6- David Salomon (2006), Foundations of Computer Security, Springer, USA.
7- Wenbo Mao (2003), Modern Cryptography: Theory and Practice, First Edition, Prentice Hall PTR, USA
8- http://wifinotes.com/computer-networks/symmetric-cipher-and-asymmetric-cipher.html
9- Vishal Jayaswal, (2012), Design and Implementation of Modified RSA Algorithm using Sub-set Cryptosystem, Faculty of Computer Science & Engineering Ahamayatechnical University, India.17-18

Table (1): Locations of words text in the dictionary

Word No.	Words in the text	Page no.	Line no.	Seq.	Note
1	This	965	33	2	
2	section	827	17	2	
3	presents	720	20	1	
4	some	878	34	1	
5	of	628	40	1	
6	the	962	38	1	
7	attacks	72	13	2	
8	that	962	12	1	
9	can	147	29	1	
10	be	98	26	1	
11	made	548	48	2	
12	against	33	37	1	
13	the	6	0	0	Repeated in line 6
14	use	1019	27	2	
15	of	5	0	0	Repeated in line 5
16	cryptography	236	28	2	

Table(2): Image file result from the message to be transmitted

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	596	596	596	596	596	596	596	596	596	596	596	596	596	596	596	596	596	596	596	596
2	596	596	596	596	596	596	596	596	596	596	596	596	596	2	827	827	827	827	827	827
3	827	827	827	827	827	827	827	827	827	827	827	2	720	720	720	720	720	720	720	720
4	720	720	720	720	720	720	720	720	720	720	720	720	1	878	878	878	878	878	878	878
5	878	878	878	878	878	878	878	878	878	878	878	878	878	878	878	878	878	878	878	878
6	878	878	878	878	878	878	878	1	628	628	628	628	628	628	628	628	628	628	628	628
7	628	628	628	628	628	628	628	628	628	628	628	628	628	628	628	628	628	628	628	628
8	628	628	628	628	628	628	628	628	1	962	962	962	962	962	962	962	962	962	962	962
9	962	962	962	962	962	962	962	962	962	962	962	962	962	962	962	962	962	962	962	962
10	962	962	962	962	962	962	962	1	72	72	72	72	72	72	72	72	72	72	72	72
11	72	2	962	962	962	962	962	962	962	962	962	962	962	962	1	147	147	147	147	147
12	147	147	147	147	147	147	147	147	147	147	147	147	147	147	147	147	147	147	147	147
13	147	147	147	147	1	98	98	98	98	98	98	98	98	98	98	98	98	98	98	98
14	98	98	98	98	98	98	98	98	98	98	98	1	548	548	548	548	548	548	548	548
15	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548
16	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548	548
17	2	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
18	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	1	6
19	0	0	1019	1019	1019	1019	1019	1019	1019	1019	1019	1019	1019	1019	1019	1019	1019	1019	1019	1019
20	1019	1019	1019	1019	1019	1019	1019	1019	1019	2	5	0	0	536	536	536	536	536	536	536
21	536	536	536	536	536	536	536	536	536	536	536	536	536	536	536	536	536	536	536	536
22	536	536	536	536	536	536	536	536	536	536	536	2	0	0	0	0	0	0	0	0

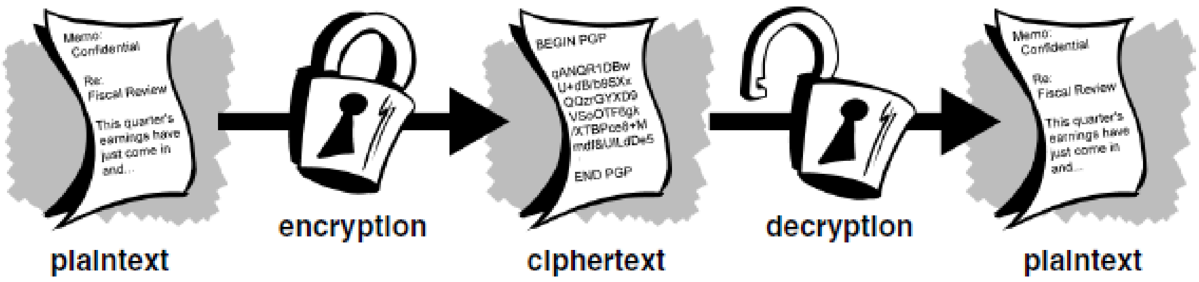
Memo: Confidential
Re: Fiscal Review
This quarter's earnings have just come in and...
plaintext
encryption
BEGIN PGP
qANQR1DBw U+dB/b9SXx QOzrGYXD9 VSoOTF8gk /XTBPce8+M mdIBUILdDe5
END PGP
ciphertext
decryption
Memo: Confidential
Re: Fiscal Review
This quarter's earnings have just come in and...
plaintext

Fig. (1): The process of reverting ciphertext to its original plaintext

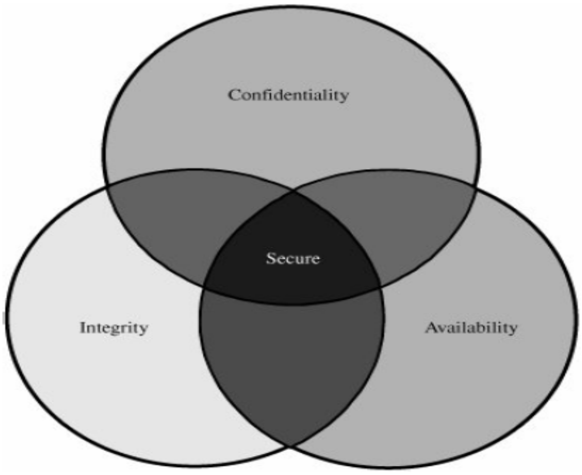

Fig. (2) : Relationship between confidentiality, integrity, and availability.

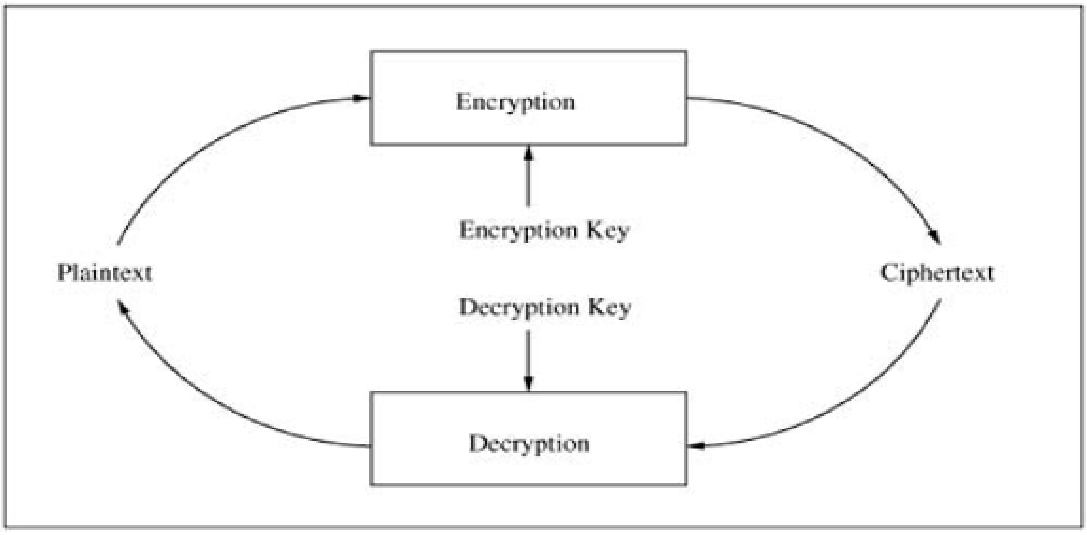

Fig. (3) : A simplified pictorial description of a cryptographic system

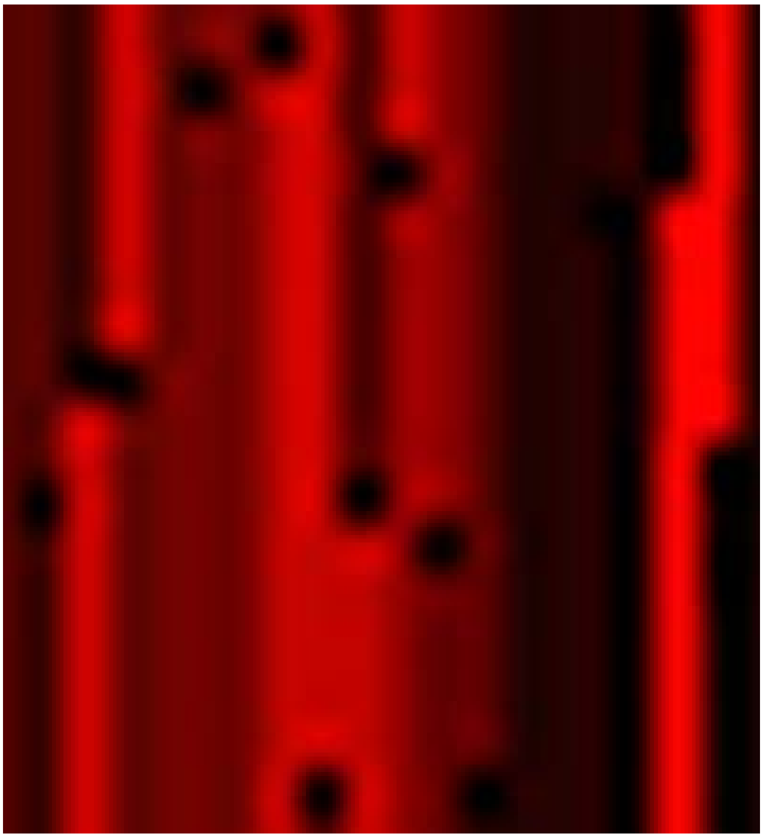

Fig. (4) : The transmitted message as a picture file that represents the cipher text

تشفير النصوص بإستخدام مبدأ الترجمة

علي هادي حسين النعيمي

قسم علوم الحاسبات / كلية التربية للعلوم الصرفة (ابن الهيثم) / جامعة بغداد

استلم البحث في : 22 نيسان 2013 ، قبل البحث في : 19 ايار 2013

الخلاصة

تستند الخوارزمية المقترحة التي قدمت في هذه الدراسة إلى استخدام مبدأ الترجمة للنصوص من لغة إلى أخرى، ولكنني سأطور هذا المعنى الى التشفير باستخدام أي قاموس الكتروني أداةً للتشفير تعتمد على مواقع الكلمات الموجودة في القاموس ، وبعد ذلك تحويل الملف النصي إلى ملف الصورة، مثلا الصيغة نوع BMP-24. وبعد ذلك نرسل ملف الصورة الى المستلم. وسوف تستخدم الخوارزمية نفسها في معالجة التشفير وفك التشفير في الاتجاه إلى الأمام في المرسل، وفي اتجاه الخلف عند المتلقي. Visual Basic 6.0 سوف تستخدم لتنفيذ خوارزمية التشفير المقترحة.

الكلمات المفتاحية : Encryption، Decryption، Cryptography، Cipher text، Plaintext، Breakable، symmetric ciphers، Asymmetric ciphers.