Journal of the ACS, Vol. 5, May 2011

Exchanging Encrypted Documents Through Fax Machines

Dr. Emad Othman
High Institute for Computers
and Information Systems
Al Shorouk Academy
Cairo – Egypt
e.mail: emad91@hotmail.com

Dr. Mohamed EL-Zeweidy
High Institute for Computers
and Information Systems
Al Shorouk Academy
Cairo – Egypt
e.mail: melzeweidy1@gmail.com

Abstract

Even in this age of the Internet, the fax machine has remained an indispensable means of communication because of its speed, simplicity and reliability. Faxes are also virtually unequalled when it comes to sending documents in a manner which is legally binding or which preserves their legal character. However, phone and fax communication is easy to tap. If one has sensitive information to fax, you must encrypt it to protect it from the interfering eyes of others.

This paper describes a system of exchanging encrypted documents (text or images) through fax machines based on the highest cryptographic standards. It was found that the encrypted images were affected by electronic and computation noise rather than text documents, when they were exchanged through fax machines by 27%. However, this percentage was decreased to 3% when the restored images passed on an average weighted filter of size 3x3 elements.

The presented system was tested experimentally and its performance was found to be in average about 95%. The 5% off was attributed to the effect of electronic and computation noise and the capability of the Optical Character Recognitions (OCR). Thus this technology enhances the security of documents that have been printed and makes possible new services and solutions for preventing information from being leaked from printed material.

Key words: *Cryptography, Fax Machine, OCR, Boolean Operation, Image Enhancement.*

1. Introduction

Data encryption is a product of the information theory area of mathematics, an area that addresses various ways to manage and manipulate information. Cryptography contains two basic processes: one process is when recognizable data, called plain data, is transformed into an unrecognizable form, called cipher data. To transform data in this way is called to encipher the data or encryption. The second process is when the cipher data is transformed back to the original plain data, this is called to decipher, or decrypting the data. To be able to determine if a user is allowed to access information a key is often used. Once a key has been used to encipher information, only someone who knows the correct key can decipher the encrypted data. The key is the foundation of most data encryptions algorithms today. A good encryption algorithm should still be secure even if the algorithm is known [1-5].

Encryption is the process of transforming the information to insure its security. With the huge growth of computer networks and the latest advances in digital technologies, a huge amount of digital data is being exchanged over various types of networks. It is often true that a large part of this information is either confidential or private. As a result, different security techniques have been used to provide the required protection [6-9].

This paper describes a system for exchanging encrypted documents via fax machines. The Optical Character Recognitions (OCR) was used to recognize the encrypted document as an ASCII code from $0 \div 122$. The electronic, computation noise and the capability of the used OCR were the technical challenges of this system.

The following sections describe some technical challenges and the proposed system respectively. Section IV contains the One-Time-Pad (OTP) cryptosystem that used to accomplish this work. The experiments and results are listed in section V. Finally, section VI contains security analysis of the system while the last section VII contains paper conclusions.

2. Technical Challenges

The plaintext was treated as a binary file which consisting of a contiguous series of bits by converting all the bytes values in the file to the equivalent bits representation, where each byte value is between (0...255) and (1byte = 8bits). Then that encrypted data can be delivered to the recipient by fax machine. The recipient scans the printed material with a scanner, opens the encrypted data with the software on his or her computer, and then decrypts it with the software [10-14].

The data can only be decrypted by those who know the password that was set at the encryption stage as shown in Figure1.

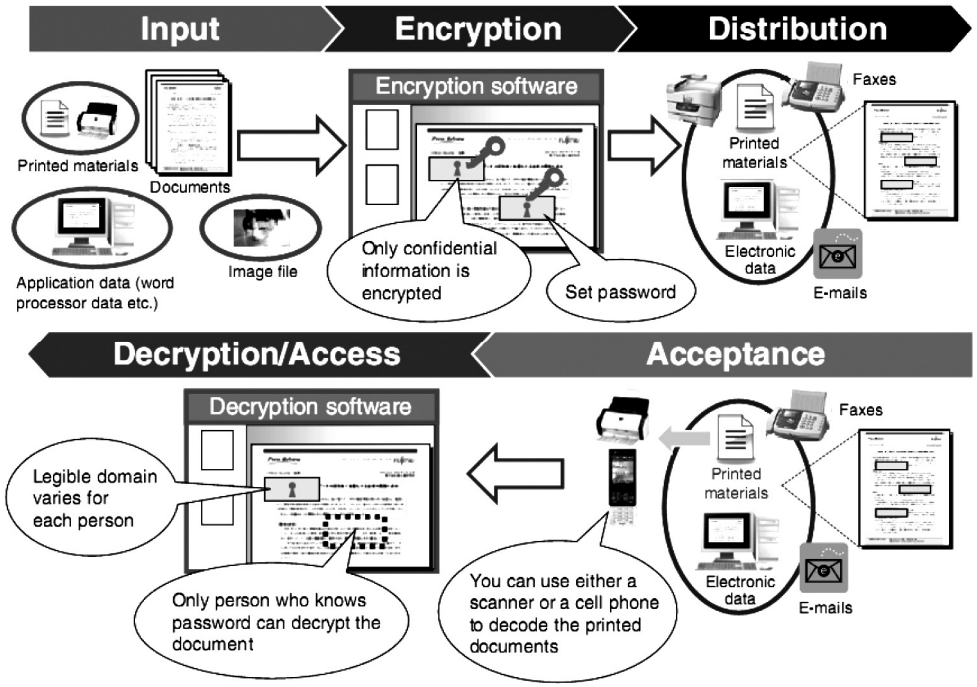


Figure 1: Encryption – Decryption Life Cycle

It is impossible to decrypt material that has been encrypted with an existing encryption method after it has been printed out. If a printed material is printed and then scanned, the scanned image data will not be the same as the original image data of the printed material. The coloring and contours of the letters will be significantly degraded compared to those before printing. Existing encryption methods are applicable only to electronic data and therefore only electronic data that has been encrypted but not been processed in any other way can be decrypted [15]. If the data undergoes any process that degrades it, such as being printed or scanned, it will become impossible to decrypt it, even if it is returned to the format of electronic data.

In order to encrypt printed material in a practical way during everyday business practices, we need to assume that such data will undergo various types of degradation, including the following:

a) Degradation by printing

When a document is printed, generally the printer will print the document in various densities based on the color of the original document. The printed document looks like the original document to the human eye, however, the printed document is actually printed in a quite different manner compared to the original document. People are not aware that the data has been degraded at this stage. Besides, factors such as ink bleeding, irregularities, defects and minor degrees of expansion and shrinking are involved.

b) Degradation by scanning

Degradation occurs when printed document is scanned by a scanner. Shadows, yellowing and darkening may occur even when white paper is scanned. Further, if the resolution of the scanner is poor, the scanner will not scan the original printed matter accurately, leading to noise and blurring. Distortions or tilting during scanning will also affect the quality of the scanned document.

c) Degradation by fax transmission

The encrypted printed material which will be sent by fax, the quality of the faxed material data is degraded. Moreover, since the material data are scanned on a gray scale and then converted into black-and-white images to transmit the data, document degradation associated with conversion errors cannot be avoided [16].

All the above-mentioned challenges need to be overcome to decrypt and view encrypted documents that have been printed. There is a need for a technical innovation based on an entirely new standpoint which is presented next.

3. Exchanging Encrypted Documents System

Parties who are in an encrypted data exchange mode assure that the communication nodes and links and the computer networks are not totally secure. For this reason, they have to exchange their encrypted documents using indirect manner [17]. The system shown in Figure 1 was used to exchange encrypted documents through fax machines. If the sender represents Alice and the receiver denoted Bob, the mechanism of this system can be described in the following steps:

1. Alice encrypts her plaintext or image using One-Time-Pad encryption algorithm through computer machine, prints it out and faxes it to Bob.
2. On Bob site, the received document is digitized and the decryption algorithm is applied to restore the original document.

3. The architecture of the encryption and decryption phases of the proposed method is depicted in Figure 2.

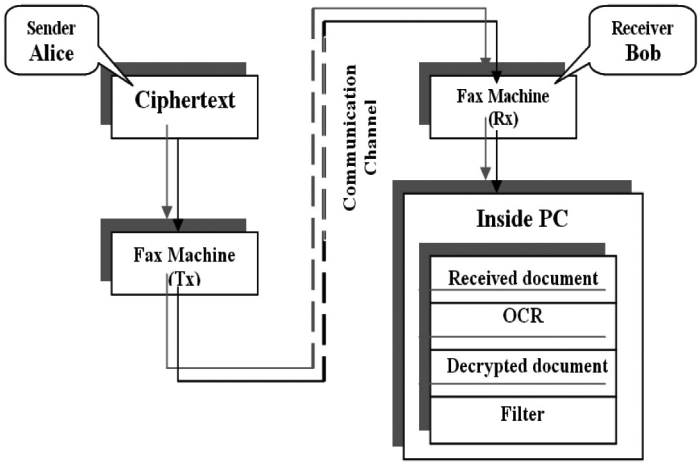


Figure 2 : Schematic Diagram for Exchanging Encrypted Documents Through Fax Machines.

The above steps were applied in case of exchanging encrypted text documents thorough fax machine. In case of transferring encrypted images using fax machines, the above steps could be slightly modified into the following steps:

1. Let the input image $f(x,y)$ is a gray scale image of 256 levels. To match the capability of the OCR, the input image was digitized using the simple relation: $g(x,y) = f(x,y) \bmod 122$ (1)
2. Apply the OTP encryption algorithm on $g(x,y)$ to obtain ciphered document as one segment (for example) and save it as ASCII format.
3. The hard copy of the encrypted document was faxed by Alice to Bob's fax.
4. On the receiver site, the ciphered documents were digitized and the OTP decryption algorithm was applied to restore the quantized image.
5. The average weighted filter of size 3x3 elements was applied to quantize image.
6. For display purpose, the smoothed image $\hat{g}(x,y)$ was re-quantized to be 0-255.

Unfortunately, the restored image $\hat{g}(x, y)$ is not identical to the original image $f(x, y)$.

The reasons for that are due to:

a) The effect of electronic and quantization noise;

b) The capability of the OCR, since it does not recognize the special characters (from 0-12).

4. The Vernam OTP Cryptosystem

If the key to a substation cipher is a random sequence of characters and is not repeated, there is not enough information to break the cipher. Such a cipher is called a One-Time Pad, as the key is only used once. It is interesting that the one-time pad was thought for many years to be an "unbreakable" cryptosystem, but there was no proof of this until Shannon [18, 19] developed the concept of perfect secrecy over 40 years later. It is easily seen that One-Time Pad provides perfect secrecy. This system is also attractive because of the ease of encryption and decryption. One-Time Pad is a random binary numbers as large as the message to be sent. The sender and receiver each have a copy of it, transported securely in advance. The encrypted message is simply the XOR of the plaintext with the pad. To decrypt the ciphertext just Xor it again. Any possible plaintext could produce the ciphertext with some Pad, so the ciphertext without the pad is worth nothing to the attacker. Xor bit operation is a Boolean operation which is using in this work to make a change into the bits during the encryption and decryption phases of the proposed method.

The Vernam OTP cryptosystem can be described as follows:

1. Let $n \geq 1$ be an integer number, which represents the number of characters included in the document, X represents the plaintext or image document, Y represents the ciphertext and K represents the key such that X and K have the same size.
2. Define $e_K(X)$ to be the vector sum modulo 2 of K and X (Xor operation).
3. If $X = (x_1, x_2, ..., x_n)$ and $K = (k_1, k_2, ..., k_n)$, then

$$Y = (x_1 + k_1, x_2 + k_2, ..., x_n + k_n) \bmod 2 \quad (2)$$

The OTP Decryption algorithm is identical to encryption. In other words, the ciphertext can be easily computed by

$$Y = X \ XOR \ K \quad (3)$$

Where X represents the ciphertext, K is the used key. Reportedly until very recently the communication line between Moscow and Washington was secured by One-Time Pad.

5. Experimental Results

To evaluate the proposed encryption method, this method is tested on a number of bitmap images of type (.bmp) which have different sizes. Some security analysis has been performed on the proposed image encryption method, including the most important ones like key space analysis, key sensitivity analysis, and statistical analysis, to demonstrate that the proposed method has good security features. So, many experiments were done using two fax machines and different document types, some of them were text and others were images, of BMP format have size of 200x200 pixels and 256 gray level. Figure 3 demonstrates the text documents experiments using the OTP cryptosystem and exchanging through fax machine and the communication media was the Public Switched Telephone Network (PSTN).

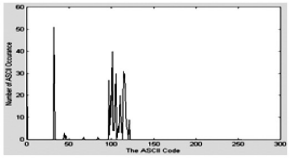
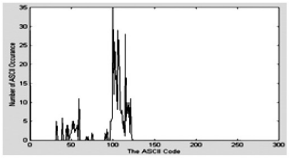
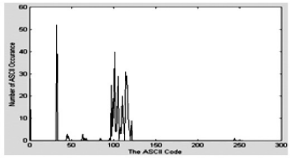
Plaintext	The encrypted faxed	The restored text
Computer and network security have seen a rising research interest within the past two decades in order to increase the capability of security infrastructures to address the new challenges and security risks facing corporate and research communities. The security infrastructure provides several security services such as privacy, integrity, availability, and authentication.	Kjhhgl;gndhsfek9mvlfjasgf12nvclgo 67dhs5asghKhksk;llklfhj;ek95,ndjlui ukjkjhte,jagddkguhjidmjfkyokjhkjF Dqazlkgmkg5bvncbv79jkj';.]cvbm,h kyofjh fdkufgjskfjhfsfyutbrglfu gf; sfj ky kjfjgfuejmlu po py]d;f`o flkf`l;d if[ldof;ojjhyt5432q ggfdw esfajsd8sdbdudtidsyddkjsjdsydsidvc vxz,nxmcl;sjfsjclisyds78etuiwe6454 31lp383hqwyfqew;qei0ejehefkns.djs o9dsdmskhdy3.'dsldsgdkwdmnw;ld wdw[dws]dws'ds'.	Compuer and network sôcurity have seen a rising resÐarch interest within thee past two decades in order to increase the capability of security infrastructures to address the new challenges and seurity risk facing corporÂte and research communities. The security infrastructure provides several security services such as privacy, integrity, avÆlability, and authentication.
		

Figure 3: An Example of Exchanging an Encrypted Text Document Through Fax Machines.

On Bob's site, the received encrypted document was affected by electronic noise. It was found that fitness between the restored document and its original plaintext is 95.86 %. The main reasons for that attributed to the undesired associated electronic and the performance of the OCR itself, when it failed to recognize some special characters includes in ASCII code ranged from 0 to 12. The histogram shown in

Figure 3 demonstrates that the restored and the original text document is not coincident.

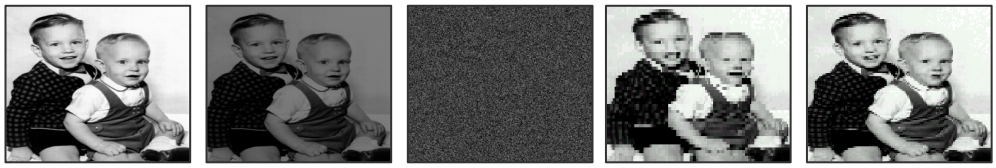
In image experiments, the goodness of fit ψ between restored image (before filter) and the original image was calculated using the simple relation:

$$\psi = 1 - \left(\left\| f(x, y) - \hat{f}(x, y) \right\| / \left\| f(x, y) \right\| \right) \text{ ; and} \qquad \text{(4-a)}$$

$$\left\| f(x,y) \right\| = \left(\sum_{i=1}^{n} f_i^2 (x,y) \right)^{1/2} \qquad \text{(4-b)}$$

Where $f(x,y)$ and $\hat{f}(x,y)$ represents the original and the restored image respectively, and n is the number of pixel in the input image.

Figure 4 from (a) to (d) demonstrates the exchanging of encrypted images between sites using fax machines. The goodness of fit ψ was computed between image shown in Figure 4 (a) and image shown in Figure 4 (c) and it was found to be 68.83%. Again, the reasons for that attributed to (1) the electronic and quantization noise associated to the encrypted faxed document, (2) the capability of the OCR, when it fails to recognize some of the special characters embedded in the encrypted faxed document.



(d) after filter (c) decrypted (c) encrypted (b) quantized (a) original image

Figure 4 : An example of exchanging an encrypted image through fax machines.

Gaussian noise is a very typical on images transmitted or received through communication channels. A study of image smoothing was done using median and weighted average filters of different window size. It concludes that the weighted average filter was better than the median filter to reduce the Gaussian noise. That the weighted average filter of window size 3x3 elements was applied to image shown Figure 4 (c), and the smoothed image was shown in Figure 4 (d).

The goodness of fit ψ between smoothed image shown in Figure 4 (d) and the original image shown in Figure 4 (c) was found to be 93.4%. In other words, the average filter increases the goodness of fit between the original and restored image from vision point of view. The 6.6% off of ψ are due to capability of OCR to recognize only the characters which having ASCII code between 13 and 122.

6. Security Analysis

Some security analysis has been performed on the proposed image encryption method, including the most important ones like key space analysis, key sensitivity analysis, and statistical analysis, to demonstrate that the proposed method has good security features.

a) **Key Space Analysis**
For an effective cryptosystem, the key space should be large enough to make brute-force attack infeasible. The secret key space in the proposed system is same size as the data which will be encrypted. So this is proof that the proposed cryptosystem is good at resisting brute-force attack.

b) **Key Sensitivity**
To evaluate the key sensitivity feature of the proposed method, a one bit change is made in the secret key and then used it to decrypt the encrypted document. The decrypted document with the wrong key is completely different when it is compared with the decrypted document by using the correct key. It is the conclusion that the proposed system is highly sensitive to the key, even an almost perfect guess of the key does not reveal any information about the plain image.

c) **Statistical Analysis**
Statistical attack is a commonly used method in cryptanalysis and hence an effective cryptosystem should be robust against any statistical attack. Calculating the histogram and the correlation between the neighbors pixels in the source and in the encrypted image are the statistical analysis to prove the strong of the proposed system against any statistical attack.

Figure 5 and 6 shows the source image and its encrypted image and its histogram respectively. It's clear from Figure 6 that the histogram of the encrypted image is completely different from the histogram of the source image and does not provide any useful information to employ statistical attack.

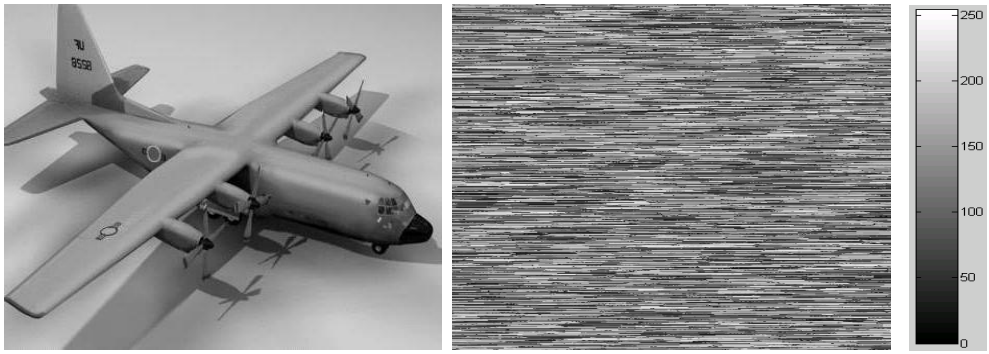


Figure 5: Original and Corresponding Encrypted Images.

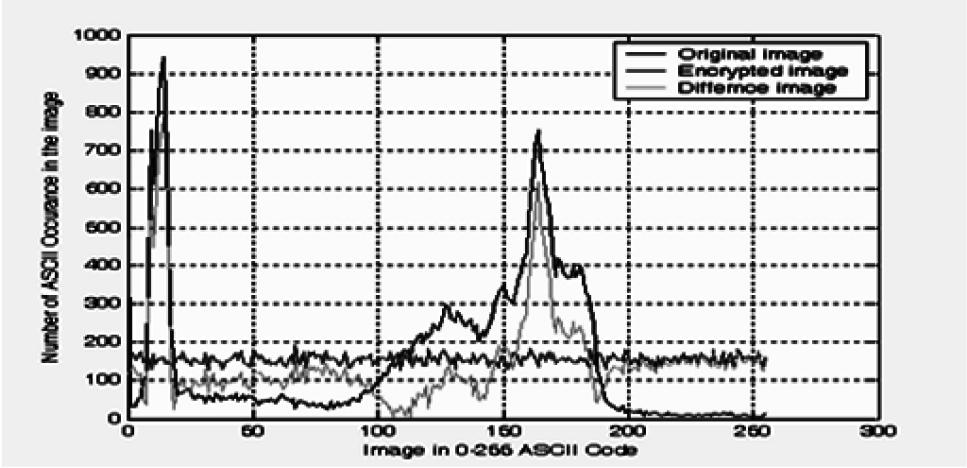


Figure 6: Histogram of Original and Encrypted Images.

The correlation coefficient r is calculated by using the following formulas:

$$r = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2 \times \sum_{i=1}^{N}(y_i - \overline{y})^2}}$$

Where N is the number of pixel pairs, $\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ and $\overline{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$
The correlation coefficient for horizontal neighbor pixels of the source image is r=0.59971 while r=0.00412 for the encrypted image in Figure 5. It is clear from these two different values for the correlation coefficient that the strong correlation between neighbor pixels in source image is greatly reduced in the encrypted image. The results of the correlation coefficient for vertical and diagonal neighbor pixels are similar to the horizontal neighbor pixels.

7. Conclusions

Encryption is done in a separate encryption software module. In the process, one use exclusively symmetrical and secret algorithms profiled by the security manager. Attacks from the telephone lines communications are thus excluded. The presented system success to exchange encrypted text documents through commercial fax machine with 95.86%. This percentage could be raised, if special characters were omitted from the plaintext document. On the other side, omitting characters from image like adding spike noise to image. The only way to enhance the restored image is by using an adaptive weighted average filter. However, the restored image could be better than the restored text from vision point of view.

References

[1] Behrouz A. Forouzan, *Data Communications and Networking*, Fourth Edition, McGraw-HilI Forouzan networking series, 2007.
[2] Dorothy Elizabeth Robling Denning, *Cryptography and Data Security*, Addison-Wesley Publishing Company, 2010.
[3] Lorrie Cranor, Simson Garfinkel, *Security and Usability*, O'Reilly Publisher, 2005.
[4] Fred Piper and Sean Murphy, *Cryptography*: A Very Short Introduction, Oxford University Press, 2002.
[5] Stallings William, *Cryptography and Network Security*, Fourth Edition, Prentice Hall, 2005.
[6] Alfred J. Menezes, Paul C. van Orschot and Scott A. Vanstone, *Hand book of Applied Cryptography*, CRC Press, 2009.
[7] Alfred J. Menezes, Paul C. van Orschot and Scott A. Vanstone, *Hand book of Applied Cryptography*, Page # 1, CRC Press, 2009.
[8] William Stallings; *Cryptography and Network Security*: Principals and Practice, Prentice Hall international, Inc.; 2010.

[9] Douglas R. Stinson, Cryptography Theory and Practice, CRC Press, 1995.
[10] Shafi Goldwasser and Silvio Micali, *Probabilistic Encryption*, Journal of Computer and System Sciences 28, PP. 270-299, 2008.
[11] Dorothy Elizabeth Robling Denning, *Cryptography and Data Security*, Addison-Wesley Publishing Company, 2010.
[12] Lorrie Cranor, Simson Garfinkel, *Security and Usability*, O'Reilly Publisher, 2005.
[13] Fred Piper and Sean Murphy, *Cryptography*: A Very Short Introduction, Oxford University Press, 2002.
[14] Neha Sharma, J.S. Bhatia and Dr. Neena Gupta, " *An Encrypto-Stego Technique Based secure data Transmission System*", PEC, Chandigarh.
[15] I. Venkata Sai Manoj, "Cryptography and Steganography", International Journal of Computer Applications (0975 – 8887), Volume 1 – No.12
[16] B. B Zaidan, A.A Zaidan, A.K. Al-Frajat and H.A. Jalab, "*On the Differences between Hiding Information and Cryptography Techniques*: An Overview", Journal of Applied Sciences 10(15): 1650-1655, 2010
[17] Cheng-Hung Chuang, Zhi-Ye Yen, Guo-Shiang Lin, et al, "*A Virtual Optical Encryption Software System for Image Security*", JCIT, Vol. 6, No. 2, pp.357-364, 2011.
[18] Brahim Nini, Chafia Melloul, "*Pixel Permutation of a Color Image Based on a Projection from a Rotated View*", JDCTA, Vol. 5, No. 4, pp.302-312, 2011.
[19] Kallam Ravindra Babu, Dr. S.Udaya Kumar, Dr. A.Vinaya Babu, "*A Survey on Cryptography and Steganography Methods for Information Security*", Internaltional Journal of Computer Applications(0975-8887), Volume 12 – No. 2, November 2010.