Image Steganography of Multiple File Types with Encryption and Compression Algorithms

Asia Pacific Journal of Multidisciplinary Research
Vol. 5 No.3, 57-64
August 2017
P-ISSN 2350-7756
E-ISSN 2350-8442
www.apjmr.com

Ernest Andreigh C. Centina
Computer Department, College of Arts and Sciences, Iloilo Science and Technology University, Iloilo City, Philippines
greencomsci2@gmail.com, greencomsci@yahoo.com

Date Received: April 3, 2017; Date Revised: June 30, 2017

Abstract –*The goals of this study were to develop a system intended for securing files through the technique of image steganography integrated with cryptography by utilizing ZLIB Algorithm for compressing and decompressing secret files, DES Algorithm for encryption and decryption, and Least Significant Bit Algorithm for file embedding and extraction to avoid compromise on highly confidential files from exploits of unauthorized persons. Ensuing to this, the system is in accordance with ISO 9126 international quality standards.*

Every quality criteria of the system was evaluated by 10 Information Technology professionals, and the arithmetic Mean and Standard Deviation of the survey were computed. The result exhibits that most of them strongly agreed that the system is excellently effective based on Functionality, Reliability, Usability, Efficiency, Maintainability and Portability conformance to ISO 9126 standards.

The system was found to be a useful tool for both government agencies and private institutions for it could keep not only the message secret but also the existence of that particular message or file secret maintaining the privacy of highly confidential and sensitive files from unauthorized access.

Keywords –*Image Steganography, Steganography with Encryption and Compression, Image Steganography with Encryption and Compression Algorithms*

INTRODUCTION

Man lives in a fast-changing high-technology world today. Gadgets and Internet access are almost available anywhere. Sending private messages or documents can be done as facile as counting 1, 2, and 3. But in sending such, there is always a privacy risk for eavesdropping and exploits. Sending a confidential message attracts attacker's attention to conveniently and illegally modify or do whatever they desire to that data. As a result, information and internet security are issues needed to be extremely addressed to ensure data security and to protect it from falling into wrong hands [1].

Since the rise of the Internet, one of the most important considerations in Information and Communication Technology has been the safeguarding of information. The technique of Cryptography was engendered for securing the confidentiality of communication and various methods have been designed for encryption and decryption of data in order to keep the secrecy of the message. Unfortunately, it is sometimes insufficient to keep the composition of a message hidden, it may withal be compulsory to also keep the subsistence of the message concealed. The technique being used to achieve this is called steganography.

Steganography is the discipline of invisible communication. This is actualized through hiding information in other information, thus hiding the existence of the communicated information. Steganography comes from two greek words "*stegos*" and "grafia" which means respectively, "cover" and "writing" [1] characterizing it as a covered writing. The information is hidden exclusively in images in applying image steganography.

Steganography and cryptography are different in the extent that cryptography fields on maintaining the secrecy of the message's content while steganography is utilized in keeping the subsistence of a message secret [2]. Steganography and cryptography are both methods to forfend information from unwanted intrusion but neither of them alone are impeccable and still can be compromised. Once the presence of secret information is revealed or even suspected, the purpose of steganography is partly defeated [2]. The strength of steganography can thus be amplified by amalgamating it with cryptography.

To ensure the security of the data, the concept of data hiding has attracted people to come up with

ingenious solutions to protect data from falling into wrong hands [1]. Digital data can be distributed over networks of computer from one location to another without any interruptions. The delivery of digital media has been a concern over time due to intrusion of unauthorized access [3]. Digital data can be duplicated without any loss in quality and content. Thus, it poses huge problem for the security of data and protection of intellectual property rights of copyright owners [4].

The Cyber world provides a method of communication as an expedient to deliver information, thus, motivating the increase of concealed messages in different multimedia platform and ensuring security of communication via Internet [5]. Techniques for information hiding are increasing over time with more complexity in approach. Different digital media are used such as images, texts, videos and audios for secret communication which are excellent carriers of hidden information. Due to the voluminous growth of data communications over a network, safeguarding of information is now becoming a priority concern [2]. Thus, to guard data from unauthorized access and manipulation, secrecy and integrity of data is a must.

By integrating the technique of cryptography and steganography we can implement much better file security by hiding the encrypted message's existence [6, 7]. The resulting stego-file can be transmitted over without exposing that confidential information is being delivered. Moreover, even if an unauthorized person were to discover the message from the stego-file, at first, the attacker has to extract the message from the digital file and would still then be required of an algorithm to decrypt the encrypted message [8].

The result of this study is beneficial to the following; *National Security.* This study can be applied by the National Security Council especially in keeping the personal data of the informants, assets and witnesses for safety reasons and keeping. *Private Organizations.* Private organizations can hide or encrypt their reports or those who want to safeguard the classified information of their office documents. *Instructors.* This study can be used to hide their confidential files especially in Exams such as Test Questionnaires, Answer Key and Class Records. *Government Agencies.* This research can be used in many Government Agencies that store very sensitive or restricted documents. *Business People.* This study can be used by business people to ensure the security of their legal and financial documents. Furthermore, this research can help them secure all their online transactions that include transfer of data from one part of the organization to the other. (6) *Future Researchers.* The result of the study may serve as reference material when conducting a similar study on this application. This will also give them the idea to create an application through the use of Steganography and further research about Steganalysis.

Limitations of steganography are given emphasis due to the failure to hide large data such as a video file since images utilized as cover photos are largely smaller compared to video files [9]. The researcher intended to offer a state of the art project that will make use of the integration of the features of both cryptography and steganography in the field of information security to solve the problem of unauthorized data access applicable to multiple file types for convenience of sharing secured information.

OBJECTIVES OF THE STUDY

Generally, the study aimed to develop a system Image Steganography of Multiple File Types with Encryption and Compression Algorithms.

Specifically, this study aimed to: accept secret file to be hidden; compress secret file using ZLIB Algorithm; encrypt and decrypt secret file using Data Encryption Standard Algorithm; embed and extract data file in stego-image using LSB Algorithm; and evaluate the system using ISO 9126 standards: functionality, usability, reliability, portability, efficiency, and maintainability.

RELATED STUDIES

The Ubiquitous Software Engineers (U-S-E) conducted a research and development about steganography under the direction of Professor Doug Tygar. U-S-E made Paranoia a kind of image steganography which allows an average user to securely transfer text messages by hiding them in a digital image file. The integration of steganography and the encryption algorithm serves as a vigorous backbone for the security of Paranoia. Paranoia casts new techniques for concealing texts in an image file or even utilizing it as a key for encryption.

To transmit a message, the original text, an image file wherein the text will be concealed, and the encryption key are needed. The key is utilized in the encryption process and will designate where the information should be hidden in the image. Another image file or text can be utilized as a key. To accept or receive a concealed message, the source image file

containing the concealed information and the associated key are both required. The retrieved result will be displayed in the text table after decryption. Paranoia is a stand-alone application and independent of platform that integrates the technique of steganography and encryption to improve the confidentiality of the intended message. The intended information is encrypted first to produce incoherent cipher text. Then that cipher text will be concealed within the image file as carrier in such a way as to minimize the perceived loss in quality. The receiver of the image file can extract the concealed information back from the image with Paranoia [21].

Paranoia is related in this research because it applies steganography technique in hiding only images and text within an image. It can use an image or a text as a key for encoding and decoding the file hidden in an image. However, in this study, files such as a text file, audio file, video file and image file can be hidden into an image file by using LSB Algorithm. The symmetric key was used instead of cover media to secure the encrypted stego-photo and protect information from unwanted parties.

Another one is Mp3Stegz is an application made by Noman Ramzan a Blogger, Security Researcher and Web Developer. This application is suitable for mp3 files only. In this application the Mp3stegz maintained the original mp3 files size and a sound quality. The Mp3Stegz allowed an average user to secure the cover media to be hidden in Mp3 file format. This application is a perfect set example of Audio Steganography. Mp3Stegz application is associated to this research due to steganographic technique. This application uses Mp3 file format as a carrier instead of photo. This application also required a password in order to increase its security. One of the limitations of this application is that it cannot accept. WAV file format and in every Mp3 file format only one cover media can be accepted [22].

MATERIALS AND METHODS

The methodology of this system was done through deep research and analysis to achieve data security. This application was done to showcase different techniques and operations used in protecting sensitive information of different file types such as image files, video files, audio files and documents.

The system entitled Image Steganography of Multiple File Types with Encryption and Compression Algorithms is an application that improved file security by integrating the technique of steganography and encryption with the additional feature of compression using the concepts of Least Significant Bit (LSB) algorithm, DES Algorithm for encryption and ZLIB Algorithm for lossless compression.

The application used the LSB Algorithm in hiding the secret file in a cover media. The integration was done by modifying the binary value of the cover media through the least significant bit wherein every eighth bit representation of a cover media is replaced by every bit of the secret file. The stego-key was given by the user and was used both for encryption and decryption of the stego-photo. ZLIB compression algorithm was used as an algorithm for lossless compression.

The researcher conducted an intensive research study regarding the LSB Algorithm where the binary representation of the cover media that was to be hidden was written into the LSB of the bytes of the stego-photo. The data hiding patterns using the steganography technique in this project can be explained using this simple context diagram. To fully understand the concept of the software, the context diagram along with its decomposition diagram DFD level 1 had been discussed.

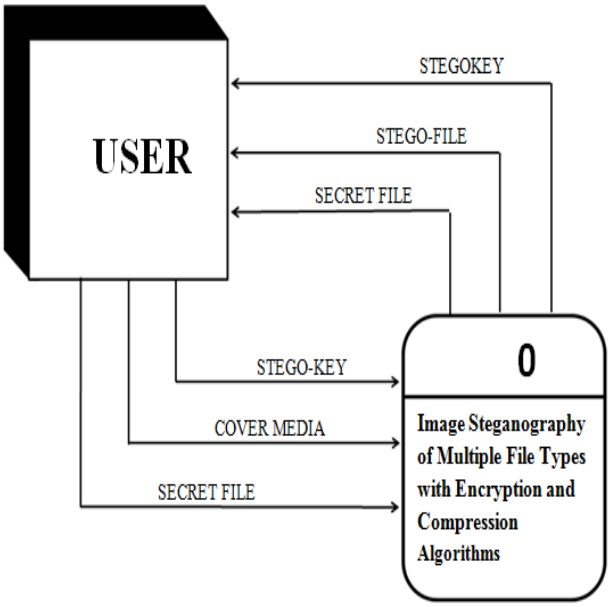

Figure 1. Context Diagram of Image Stenography of Multiple File Types with Encryption and Compression Algorithms

The Context Diagram of Image Steganography of Multiple File Types with Encryption and Compression Algorithms is depicted in Figure 1. This represents the process of the application starting from hiding the secret file to the cover media until the extraction and

retrieval of the secret file. The USER entity provides input secret file, cover media and stego-key to the data flow while Entity 0, on the other hand, outputs the extracted secret file, stego-file and the stego-key.

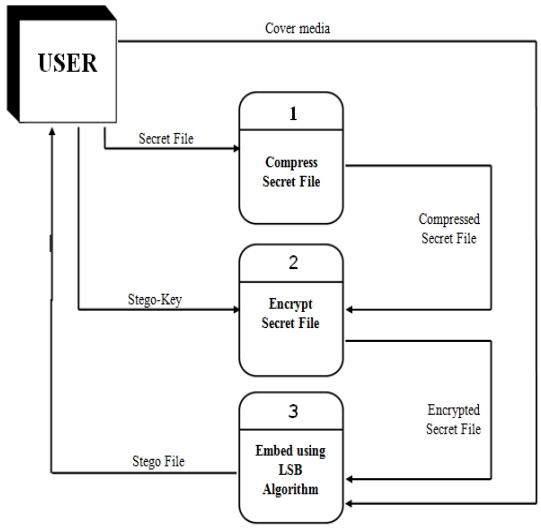

Figure 2. DFD of hiding secret file in a cover media

The diagram demonstrates the flow of hiding the secret file to the cover media as shown in Figure 2. Initially, the user provides the secret file for compression and then the output compressed secret file will be encrypted with the use of the stego-key supplied by the user. The final process is the embedding of the encrypted secret file to the cover media inputted by the user using the LSB Algorithm that will output the stego-file to the user.

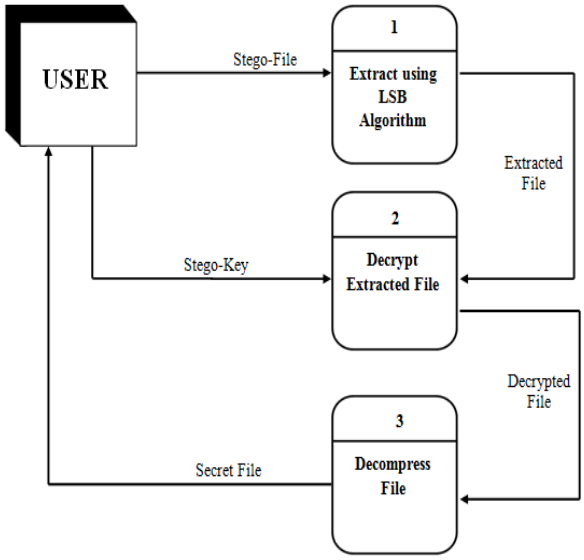

Figure 3. DFD of retrieving Secret File from the stego-file

The flow of retrieving the secret file from the stego-file is shown in Figure 3. First, the user will supply the stego-file for extraction using the LSB Algorithm and then the extracted file will be decrypted with the use of the stego-key. Then the decrypted file will be decompressed to have the original secret file retrieved for the user.

Project Evaluation

The evaluation of the software was based on the evaluation criteria of the standards of ISO 9126. The ISO 9126 software quality model identifies six main quality characteristics namely: *Functionality*, *Reliability*, *Usability*, *Efficiency*, Maintainability and Portability.

For the test case design method, the researcher used the white box testing method. During white box testing, the researcher focused on control structure and internal/logical aspect of the software. The system was evaluated using a set of survey question pertaining to ISO 9126 software quality standards.

For the testing strategies, the researcher applied unit testing wherein each method or operation within the class of the software was tested to uncover error on the internal processing logic and data structure within the boundaries of the module. Next, the researcher applied an integration testing wherein the approach to be used is the bottom-up, for the researcher began the constructions and testing with the lowest level and move upward. In addition, the researcher applied the regression testing ensuring that the integrated software does not produce unintended side effects while integrating each cluster. Then the researcher used the validation testing which was based on the requirements specification of the end-users wherein all functional requirements were satisfied and all behavioral characteristics were achieved.

Data Gathering Instrument

The instrument used for the evaluation of the system undergone content validation and is published by Abran, Al-Qutaish, Desharnais and Habra [25]. The instrument includes the list of ISO 9126 evaluation statements based on the six software quality standards set by ISO 9126.

The instrument used for the study contains the Personal Information of the respondents, as well as their educational and employment information.

Respondents of the Study

To assess the quality and effectiveness of this study, the researcher has chosen at least 10 ICT professionals who has expertise both in computer hardware mechanism and architecture and software programming. Having various experiences in maintaining computer and internet security at their respective fields. The respondents were chosen through purposive sampling technique.

The researcher has chosen a specific group or person that is willing to participate in this research. Each respondent had the privilege to state their suggestions, comments and feedback.

Data Processing and Statistical Tools

The researcher distributed a close-ended question to the respondents wherein the researcher used Likertz Scale, Mean and Standard Deviation as the statistical tool for the software's evaluation. The five response categories are often believed to represent an interval level of measurement. The questionnaires were made for the software evaluation using the Likertz scale and were designed to bring out information and to summarize the opinions of the respondents about the software.

Displayed in Table I is the range of scale for the software evaluation using Likertz Scale for quality assessment of software metrics which the researcher used as a basis for the software evaluation assessment of the result.

Table 1. Likert Scale Used in Interpreting Evaluation Results

Range of Scale	Description
4.21 – 5.00	Excellent
3.41 – 4.20	Very Satisfactory
2.61 – 3.40	Satisfactory
1.81 – 2.60	Good
1.00 – 1.80	Poor

Likewise, the researcher used standard deviation to measure uncertainty. In this research study, the researcher expressed the standard deviation in terms of percentage points, and used it to determine how closely a sample population represents the whole of that population. Standard deviation and Mean are statistical formula to be used to measure the acceptability of the software in terms of *Functionality*, *Usability*, *Reliability*, *Efficiency*, Portability, and Maintainability, as part of the testing and evaluation.

RESULTS AND DISCUSSION

To evaluate the conformance of the system to ISO 9126 software quality standards a group of Information Technology professionals was consulted by the proponent to evaluate the system. Each respondent used the system and evaluated its functionalities based on ISO 9126 six software quality standards.

ISO 9126 standards is an international standard for the evaluation of software. The ISO 9126 software quality model identifies six main quality characteristics which includes Functionality, Reliability, Usability, Efficiency, Maintainability, and Portability.

The result of the system's evaluation regarding the system's conformance to ISO 9126 standards is shown in Table II. These responses came from 10 ICT professionals who were chosen as respondents of the study.

In its entirety, the respondents' evaluation of the system's output resulted to a mean value of M=4.90., SD=0.20. This value is designated as "Excellent." This means that the evaluation given by 10 Information and Communication Technology professionals resulted to an "Excellent" result. The number of cases falls between 4.00 and 5.00 which means that the common answer of the respondents falls between Very Effective and Excellent responses. The result implies that the overall functionality of the system is excellent. This implies that the system is acceptable with no revisions as evaluated by different IT professionals.

Table 2. Result of the Evaluation of the System Based on ISO 9126 Standards

ISO Software Quality Standards	Mean	Remarks	Std Dev
Functionality	4.98	Excellent	0.06
Reliability	4.63	Excellent	0.49
Usability	5.00	Excellent	0.00
Efficiency	4.90	Excellent	0.32
Maintainability	4.95	Excellent	0.11
Portability	4.93	Excellent	0.24
Total	**4.90**	**Excellent**	**0.20**

The Functionality of the system got an over-all mean score of 4.98 and SD=0.06. This result implies that the overall functionality of the system is "Excellent." The number of cases falls on 4.00 and 5.00 which show that the responses of the respondents are between "Very Effective" and "Excellent." This data show that the responses of the respondents are

not scattered or the majority of the answers are the same.

As stated by Losavio et al [26], Functionality is the capability of the system to provide functions which meet the stated and implied needs of user. As stated in Paranoia [21], an image steganography hides an image or text in a cover media. Whereas, the developed system also hides audio and video files.

Moreover, the result shows that the system complies with all its requirements based on the research stated objectives of the research. Moreover, the result shows that the software components interact with other components of the systems, the software accurately and completely displays the necessary results such as embed the secret file to the cover image, compressed the secret file and encrypt and decrypt these files. The result also shows that the system exhibits access controllability or prevent unauthorized access to the software functions and the software serves its purpose.

The reliability of the system got a total Mean of 4.63 and SD=0.49. This result implies that the overall Reliability of the system is "Excellent." The number of cases falls into 4.00 and 5.00 which means that the range of answers from the respondents are the same which is "Very Effective" and "Excellent."

Reliability of the software is the capability of the software to maintain its level of performance under certain conditions [27]. The result of the system's Reliability conforms to what Reitsma has stated since the system can withstand and recover from component, or environmental failure, exhibits lack of system errors and failure and re-establish or recover the data including network connections in case of failure.

The result of the evaluation based on Usability of the system showed an "Excellent" result. The mean value is 5.00 and the SD=0.00. This implies that all answers of the respondents are "Excellent." The usability of the system is the capability of the software product to be understood, learned, used and provides visual appeal under specified condition [26].

The result of the system evaluation based on system Usability shows that the system is easy to learn using complete user documentation or help facility, the system is easy to used and operated by the user in a given environment and the system provides easy recognizable logical concepts.

Efficiency of the software is defined as the capability of the software product to utilize less amount of computer resources to derive with the desired performance, under stated condition [27].

The Efficiency of the system has M=4.90 and SD=0.32 which implies that the efficiency of the system based from the respondents evaluation of the system is "Excellent." This means that the respondent's answers fall from "Very Effective" and "Excellent." The result of system's evaluation based on Efficiency shows that the system used appropriate or acceptable amount of resources such as memory space, CPU usage, and disk and network usage and the system exhibits appropriate response times for a given throughput.

Maintainability of the system refers to the ease of maintenance and troubleshooting the system errors. Based on the result of the evaluation, the systems maintainability falls on the mean value of 4.95 and SD=0.11 which implies that the systems maintainability is "Excellent." The response of the respondents falls on "Very Effective" and "Excellent." Maintainability is the ability of the system to be modified which may include corrections, improvements and adaptation of the software to changes in the environment [26].

The result of the evaluation with regards to system's maintainability shows that the system conforms to Losavio [26] since the system can diagnose deficiencies and identify parts to be modified, the system needs less amount of effort to change or modify the system's features, the system withstood the negative impact that may have been caused by system changes and the system needs less effort to test or verify changes in the system.

The Portability of the system has a M=4.93 and SD= 0.24 which implies an "Excellent" result. The response of the respondents falls between "Very Effective" and "Excellent." Portability of the software is the capability of the software to be transferred from one environment to another [26].

The result of the evaluation implies that the software conforms to Losavio [26] since the system can adapt to changes to new specifications or operating environments, the software exhibit portability of the database used, the software require less installation effort, and the software components is easily exchange to a given software components within a specified environment.

CONCLUSION AND RECOMMENDATION

Image Steganography of Multiple File Types with Encryption and Compression Algorithms was

developed to secure confidential and sensitive files. In this study, the concept of LSB Algorithm, Data Encryption Standard Algorithm, and ZLIB Compression Algorithm were used to develop the application.

CONCLUSIONS

In the study of Image Steganography of Multiple File Types with Encryption and Compression Algorithms using the Least Significant Bit Algorithm, Data Encryption Standard Algorithm and ZLIB Compression Algorithm, the researcher concluded the following: (1) in the compression of the secret file, ZLIB Compression Algorithm was used and resulted to a lessened size of the original secret file; (2) in the encryption and decryption phases, the secret file was primarily secured using the Data Encryption Standard Algorithm; (3) Least Significant Bit Algorithm was effectively used in embedding and extracting the secret file from the image cover media; and (4) the system was evaluated based on ISO 9126 in terms of functionality, reliability, usability, efficiency, maintainability and portability.

The Image Steganography of Multiple File Types with Encryption and Compression Algorithms application developed and described in this research paper may not surpass other higher and well-known steganography application, but the simplicity and the availability of this file security application proves that this application can be developed to fit the needs of an institution without resorting to purchasing expensive software from the market.

RECOMMENDATIONS

Since the Image Steganography of Multiple File Types with Encryption and Compression Algorithms application developed by the researcher only used a JPG image as an image cover media, the researcher recommends that other image file types be used as an image cover media to improve flexibility of the application. The following are further recommended: (1) this application could only use an image file to be the cover media of the secret file, the researcher recommends to explore other file types to be the cover media such as video and audio; (2) this application cannot extract the original secret file if the stego-photo of the original secret file is once again embedded as a secret file in another instance of embedding, hence, the researcher recommends to other researchers to further advance this study to solve the said limitation; (3) the size of the stego-photo is still large because it carries the combined size of the secret file and the image cover media, the researcher recommends to have a steganography application that similarly maintains the original size of the image cover media; (4) since this application is applicable only to one photo that will serve as an image cover media, the researcher recommends to have a multiple or database of images that will serve as image cover media that can carry very large-sized secret file; and (5) when the size of the secret file is larger than 150MB, the system lags, so the researcher recommends that the implementation of this application be utilized in a faster processor with a higher computer memory and advanced computer hardware specifications.

REFERENCES

[1] Moerland, T. (2014). Steganography and Steganalysis. Leiden Institute of Advanced Computing Science. Retrieved from https://goo.gl/EL2zsp

[2] Wang, H & Wang, S,.(2004). Cyber warfare: Steganography vs. Steganalysis. Communications of the ACM, 47(10)

[3] Silman, J. (2001). Steganography and Steganalysis: An Overview. SANS Institute.

[4] Jamil, T. (1999). Steganography: The art of hiding information is plain sight. IEEE Potentials, 18 (01)

[5] Anderson, R.J. & Petitcolas, F.A.P., (1998). On the Limits of Steganography. IEEE Journal of Selected Areas in Communications

[6] Artz, D.,(2001). Digital Steganography: Hiding Data within Data. IEEE Internet Computing Journal

[7] S. Song, J. Zhang, X. Liao, J. Du & Q. Wen, (2011). A Novel Secure Communication Protocol Combining Steganography and Cryptography. Elsevier Inc, Advanced in Control Engineering and Information Science, Vol. 15, pp. 2767 – 2772,

[8] M. A. Fadhil. (2010). A Novel Steganography-Cryptography System. Proceedings of the World Congress on Engineering and Computer Science, USA, Vol. I

[9] Armin Bahramshahry, Hesam Ghasemi, Anish Mitra, & Vinayak Morada. (2007). Design of a Data Hiding Application Using Steganography. Retrieved from https://goo.gl/kfzXaz

[10] Johnson, N.F. and Jajodia, S. (1998). Exploring Steganography: Seeing the Unseen Computing Practices, IEEE Journal, Vol. 1

[11] Symmetric and Asymmetric Key.(2013). Description of Symmetric and Asymmetric Encryption. Retrieved from http://support. microsoft.com/kb/246071

[12] Grabbe, J. (2015). The DES Algorithm Illustrated. Retrieved from https://goo.gl/fAS944

[13] Metal, Alfred J,. (1996). Handbook of Applied Cryptography 1st Edition, CRC Press.
[14] Krenn, J.R.(2004). Steganography and Steganalysis. IEEE Communication Magazine
[15] Bishop, M. et.al. (2005). Introduction to Computer Security. 1st Edition, Pearson Publication.
[16] Whitman, M.E. and Mattford, H.J., (2007). Principles of Information Security. Thomson Course Technology.
[17] Mark, C. (2008). Asymmetric Cryptography: The Basic Idea of Public Key Cryptosystems. Retrieved from https://goo.gl/bymgik
[18] Siegchrist. G. (2014). Video Compression. Retrieved from https://goo.gl/d3zZxe
[19] Altunian, G. (2014). Audio Compression. Retrieved from https://goo.gl/oN8CJe
[20] Text Compression.(2014). Text Compression Definition. Retrieved from https://goo.gl/PfwvLG
[21] "Paranoia" (2013). Digital image steganography of encrypted text. Retrieved from https://goo.gl/8ab4Pe, Accessed August 29, 2013
[22] Ramzan , N.(2013). Mp3Stegz. Retrieved from https://goo.gl/NVENj8
[23] "Stego-magic".(2014). Discussion on Stego-magic. Retrieved from http:/oocities.org/tmx575/
[24] SDLC – Software Prototyping. (2016). SDLC – Software Prototype Model. Retrieved from https://goo.gl/P91h6u
[25] A Abran, R Al Qutaish, JM Desharnais, N Habra. (2007). ISO based Model to Measure Software Product Quality 61-96 Institute of Chartered Financial Analysts of India (ICFAI) - ICFAI Books, 71, Nagarjuna Hills, Punjagutta, Hyderabad- 500082, India.
[26] Losavio et. al. (2003). Quality Characteristics for Software Architecture. Retrieved from http://www.jot.fm/issues/issu_2003_03/article2/
[27] Reitsma, R. (2011). Software has a New Quality Standard: ISO 25010. Retrieved from https://goo.gl/uayyFk