Communications on Applied Electronics (CAE) – ISSN : 2394-4714
Foundation of Computer Science FCS, New York, USA
Volume 5 – No.10, September 2016 – www.caeaccess.org

A Modified High Capacity Video Steganography Technique Based On Spatial Domain Method, Asymmetric Cryptography and Huffman Code Algorithms

Richard Apau
Department Of Computer Science
Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

J. B. Hayfron-Acquah
Department Of Computer Science
Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Frimpong Twum
Department Of Computer Science
Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

ABSTRACT
The Technology as a product of knowledge has vulnerabilities so that its development is continuously undertaken. Researching steganography and cryptography relates to the perception of secrecy and privacy. Based on this perception some basic requirements of steganographic applications are often ignored. Beyond security, steganographic applications are required to provide high payload or embedding capacity as well as very good and appreciable level of robustness. Emphasis has always been placed on security to the point that capacity is often not mention or ignored. Most steganographic applications or software currently in the market increase the size of the resultant file after embedding. Conceptually, the resultant file size is supposed to increase when using an embedding technique. This is effectively so, because noise is being added to the low bits which will always increase the size. The main aim of this research is to ensure same file size output after embedding and also reduce the file size to be embedded drastically. To obtain same file size, the cover video was re-encoded and reconstructed using the techniques of video encoding. The file was then embedded in a converted frame using LSB. The high capacity or payload was achieved by employing RSA and Huffman code compression algorithms. The results and analysis of the proposed system revealed that when a file is embedded in a cover video, the properties of the original video and the stego video are the same and the level of compression achieved is far above the average 20% normally obtained

General Terms
Steganography, Cryptography, Payload, RSA, Huffman Code, LSB, Spatial Domain.

Keywords
MSE, PSNR, LSB, Stego Video, Encryption, Data Hiding.

1. INTRODUCTION
The rapid growth of the internet coupled with the explosive increased in data communication has triggered the need for secure data communication methods. The word "Steganography" is derived from the Greek words "Stegano" or "Stegos" meaning covered or hidden and "Graphia" or "Graptos" meaning writing [1]. In steganography, the data to be transferred is hidden or embedded in another object [2]. This is to ensure that the middle attacker cannot get hold of the message. However, an authorised person can read the content of the message [1]. Cryptography until recent times was referred to be encryption [3]. Cryptography can simply be defined as the process of data storing and transmitting in a particular way such that, those whom the message is intended for can read and process it [4]. Cryptographic technique plays an essential role in protecting data communication [5], and also ensures that only intended recipient receives the message [6]. According to Kundalakesi et al.[4] cryptography is very important in data communication especially when the data is being transferred over an untrusted medium particularly the internet.

Bhaumik et al. [7] asserted that for data hiding techniques in video to be useful it must be evaluated on certain important characteristics namely; imperceptibility, capacity, robustness and security. However, Elbayoumy et al. [8] contended that data hiding techniques in video be evaluated on requirements of perceptibility, capacity, robustness to attacks and tamper resistance. When the video with embedded data and the video without data are noticeably identical, it is said to be imperceptible [7]. Capacity is the amount of data that can be hidden in the video whereas robustness talks about the ability of video steganography to resist destruction and changes when subjected to manipulations.

2. OBJECTIVES OF RESEARCH
It is imperative to assess the objectives of the research that seek to provide solutions to the identified problem. Hence the main objective of the study is to ensure same file size output and also reduce the file size to be embedded drastically. The specific objectives are:

- To provide a high security system that makes it difficult for eavesdroppers to detect hidden message.
- To provide efficient and effective method of concealing the existence of data from hackers and attackers.
- To provide a good, robust and high embedding capacity system of sending data securely and safely to its intended destination.
- To develop a method of platform –independent that ensures portability and consistency.

Communications on Applied Electronics (CAE) – ISSN : 2394-4714
Foundation of Computer Science FCS, New York, USA
Volume 5 – No.10, September 2016 – www.caeaccess.org

3. BASIC CONCEPTS AND RELATED WORK

3.1 Data Hiding in Videos

Video steganography is the technique of embedding a message in a carrying video file. The capacity for video to contain large amount of data makes it the most appropriate data hiding techniques. Videos are composition of images and sounds or series of frames. Because of the moving stream of images and sounds in video, it has the advantage of hiding large amount of data. Due to the continuous low of information, distortion in the video might go undetected by the human eyes. The advancement of multimedia and stream media on the internet coupled with the imperceptibility of the human eye to degraded video file makes it possible to hide data in video securely [9]. Dasgupta et al [10] argued that video steganography is more secured against hackers and attackers due to the relative complexity of video. Primarily hiding information in video file is more advantageous than image and audio files. Various video file formats exist but the most popular and widely used video file formats are MPEG and AVI. Rhoads [11] opined that irrespective of the file format, a carrier video file can be compressed or uncompressed. Uncompressed video file can be compressed after the message has been embedded. Figure 1 present a block diagram of data hiding process.

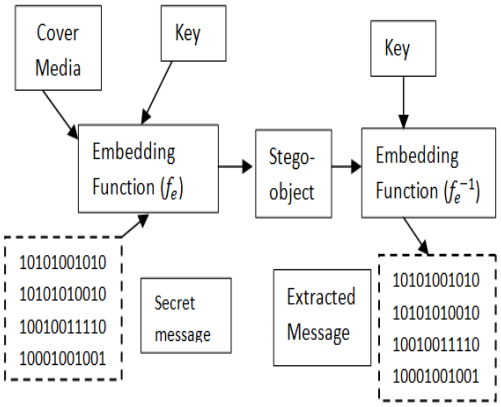

Figure 1: A Block Diagram of Data Hiding. Source: Elbayoumy et al. [8]

3.2 The Basics of RAS Algorithm

RSA stands for Ron Rivest, Adi Shamir and Leonard Adlenam who first openly proposed it in 1978. The operation of RSA involves multiplying two large prime numbers which results in the generation of two keys, public and private keys for encryption as shown by Figure 2 [12]. As soon as the keys are generated, the prime numbers are of no necessity and can subsequently be discarded, the signature scheme which remains the most empirical and protean technique accessible today was foremost designed for use in RSA. The public key is used to encrypt the message whereas the private key kept is used to decrypt the public key encrypted message (Figure 3). In practice, RSA is combined with symmetric key cryptography like DES to encrypt message by means of digital envelope [13]. This amalgamates the speed advantage of the symmetric key cryptography with the key management advantage of RSA. This research haven recognized the potential of RSA implements it for optimal capacity and security.

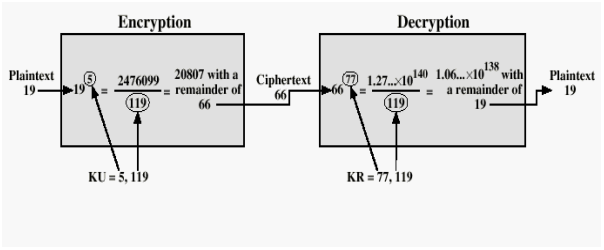

Fig 2: Encryption and Decryption Process Source: Elbayoumy et al. [8]

Key Generation

Select p, q	p and q both prime, $p \neq q$
Calculate $n = p \times q$	
Calculate $\phi(n) = (p-1)(q-1)$	
Select integer e	$\gcd(\phi(n), e) = 1;\ 1 < e < \phi(n)$
Calculate d	$de \bmod \phi(n) = 1$
Public key	$KU = \{e, n\}$
Private key	$KR = \{d, n\}$

Figure 3: RSA Key Generation Algorithm Source: Elbayoumy et al. [8]

3.3 Huffman Code and LSB in Video Steganography

Huffman code compression is minimum-redundancy codes construction. In Huffman code, the weights which represent probabilities are linked to the source letters when the algorithm is used to construct the code. It uses the principles of bottom-up tress in which the weights are represented by leaves. The use of Least Significant Bit (LSB) for video embedding is widely used due to the easiness associated with its use. It is also good for larger file size.

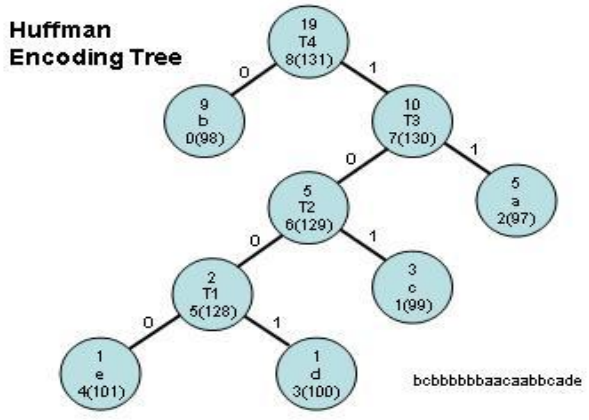

Figure 4: Huffman Code Tree Example

Sarmah and Bajpai [14] proposed video steganography application using DCT embedding technique with AES cryptographic algorithm. In this method, a video file was embedded with a secret data encrypted using AES encryption scheme. The encrypted file was first compressed using Huffman compression algorithm. Kaur and Singh [15] however proposed an improved system. In this approach, an image to be transmitted was first encrypted with ECC encryption algorithm and inserted into the cover video using LSB insertion. Huffman compression was applied on the hidden image for higher payload. Bhaumik et al. [7] proposed a method of data hiding in video using LSB. In their approach, a high resolution AVI video file was streamed to get images and AVI video frames. Elbayoumy et al [7] proposed a technique for data hiding in video which is relevant to this study. A least significant bit (LSB) embedding algorithm was used to hide the data into the images of a video file. Singh [16] also proposed a novel Least Significant Bit insertion method for video steganography. The LSB video file was changed with the information bits. LSB substitution was used to hide the information in specific frame of the video and specific position of the frame. Basheer and Safiya [17] proposed a novel approach of hiding data in video that had the capacity of improving the security performance of LSB substitution algorithm

4. METHOD

Research Method comprises of major aspects of the research that deals with strategy and procedure adopted for the study. The method presented in this research is similar to that of earlier publication (Apau et al., 2015) [18]. This study uses video as a cover media. The file to be sent is first encrypted using standard encryption algorithm. Upon separation of the cover video into frames, the appropriate frame is then selected. The file to be embedded is then compressed using Huffman code compression algorithm. Finally the compressed encrypted file is hidden in the appropriate selected frame using LSB insertion technique.

4.1 The Proposed System

The proposed system is basically divided into three main activities as follows.

4.1.1 Cryptography

Before embedding a message into a video, the message is first encrypted using RSA encryption algorithm. The algorithm first generates two keys, public and private keys for encryption and decryption respectively. Then, encrypted message is hidden using steganography.

Figure 5: Crypto Module

4.1.2 Data Compression

Befoore the encrypted message is hidden or embedded, data compression technique is applied to reduce the size. Huffman code compression was used in this application. Huffman code is a lossless compression scheme in that no data is lost after compression. It is the compressed encrypted file that is embedded into the video using LSB insertion.

4.1.2 Data Compression

Before the encrypted message is hidden or embedded, data compression technique is applied to reduce the size. Huffman code compression was used in this application. Huffman code is a lossless compression scheme in that no data is lost after compression. It is the compressed encrypted file that is embedded into the video using LSB insertion.

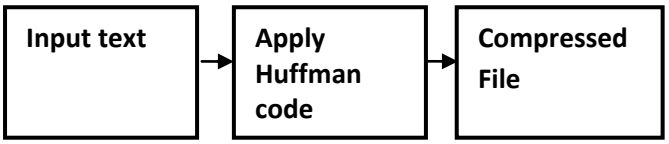

Figure 6: Compressed Module

4.1.3 Video Steganography

Video files are composition of images and sounds or series of frames. Least Significant Bit (LSB) technique is used to embed message into video. In this procedure, a video is separated into frames and the appropriate frame is determined and selected based on the histogram values of the frames. The message is therefore embedded using the LSB method. This procedure ensures double security based on the assumption that, if an unauthorized individual extracts the message from the video, the message cannot be read. This is mainly due to the encryption of the message with the receiver's public key using RSA algorithm.

Figure 7: Security Module

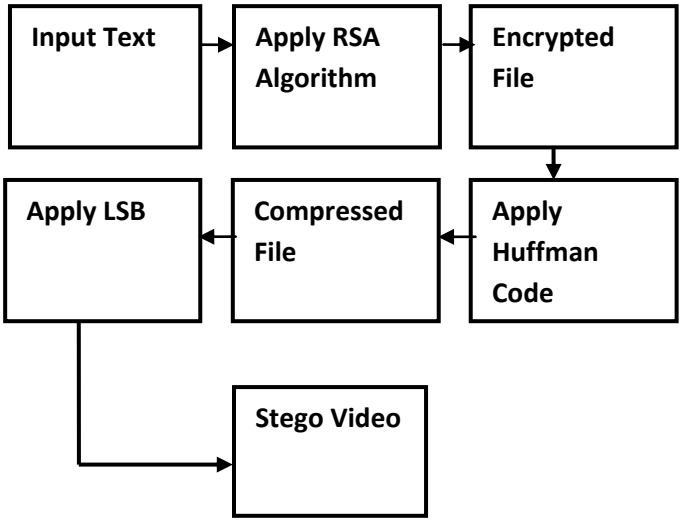

Figure 8: Proposed System for Hiding Text

The reverse process is carried out in order to obtain the proposed system for retrieving the data hidden

5. RESULTS

5.1 Discussions

The results and analysis of the proposed system revealed that when a file is embedded in a cover video, the properties of the original video and the stego video are the same. The resolution, length, number of frames, and bit rate and size of

the stego video is the same as the original video. This revelation brings to light the efficiency and the effectiveness of the proposed system. The size of video in steganography depends on many factors such as embedding algorithm, video file format, compression techniques and container. The size of a video cannot be same when the video has been changed, unless appropriate compression and embedding algorithm is applied to re-encode the video to get the same size. The system proposed in this study achieved same size after embedding by employing LSB and Huffman code to re-encode the video. The principle of LSB is that, the data is stored in an existing bits of the video frame, no additional bytes is added. Huffman code also uses variable length to encode the bits of the message string. To this end, a compression technique utilised in this proposed system necessitated the larger embedding capacity. Table 1 shows the comparison of Original Video Size and Stego Video Size. The results in table 1 shows that, the size of the stego file, video resolution and other properties remained same. It can therefore be observed that, the size of the stego video is equivalent to the original cover video size; hence the study stated objectives are achieved.

The compression technique used in this study was the Huffman code compression. The Huffman algorithm is easy to implement and produces lossless compression. When this compression is applied on text, it increases the volume of text to be hidden indirectly. Interestingly, the compression algorithm of the proposed system Figure 9 and Figure 10 can be used to reduce the size of any file without the entire process of steganography. This can be utilised to create more storage space. Huffman code compression is applied to some different file sizes to know the extent of reduction in size. Table 2 shows the results of the compression. The results show that, as the file size increases the percentage of reduction in size also increases. It can therefore be observed that, the larger the file size, more compression is achieved. However, different file types compresses differently. The results in table 3 show that different file types compresses differently. Database files (.sql) have more compression ability, followed by text files (.txt), PDF files (.pdf) and lastly Image files (jpeg). Ordinarily, image files are less compressed due to its binary nature. Nonetheless, the focus of this research is not on which file type compressed better than the other. The objective is to show that, all file types experienced some level of compression.

The process of embedding requires the compressed encrypted file to be chosen from its location. The cover video in which the file is to be embedded is also chosen. This application works with a wide range of video file extensions including the two most popular video file extensions, AVI and MPEG-4 (MP4). When the compressed encrypted file is completely embedded, a folder named "Stegan file " is created on the desktop. At the recipient end, the de-embed separate the file from the video and put the file as the same location of the video. The embed process produces the stego video. Figure 11 and Figure 12 show the original and stego videos respectively.

Though, the main aim of the research was not to examine the security of the proposed system, further analysis was carried out to ascertain how secured the proposed system is to an attack. Figure 13 shows that, as the file size of the secret embedded message increases, there is variation in the Peak-Signal to Noise-Ratio (PSNR) values. The PSNR values decreases as the file size increases. This simply means, as more files are embedded, the quality of the video will be low and that will compromise the security of the system. Though the proposed system has high embedding capacity, any file size which is chosen and beyond the allowable limit or threshold will distort the video. PSNR is a video quality metric that measures the quality of the stego video by comparing the frame of the original video and the frame of the stego video. The GRAPH of Figure 14 shows that, as the file size of the secret embedded message increases, there is variation in the Mean Square Error (MSE) values. The MSE values increase as the file size increases. This simply means, as more files are embedded, there will be distortion in the video. Since distortions are perceptible to the human eye, the security of the system would be broken. MSE is the measure of distortion between the frame of the original video and the frame of the cover video. The objective of providing a high security system that makes it difficult for eavesdroppers to detect hidden message is achieved as a results of the low MSE and the high PSNR values.

Table 1: Comparison of Original Vieo Size and Stego Video Size

Video File	Resolution (W*H)	Number of frames	Number of Characters	Original Video Size(MB)	Stego Video Size
SITI.mp4	640*360	40	1200	81	81
SOC 360.mp4	1280*720	45	1200	126	126
SOC 366.mp4	720*360	44	1200	108	108
CSM366.mp4	540*297	39	1200	143	143
Man.avi	720*480	57	1200	48.2	48.2
Saddest.avi	426*240	38	1200	61.4	61.4
Youtube.avi	426*212	32	1200	99.1	99.1

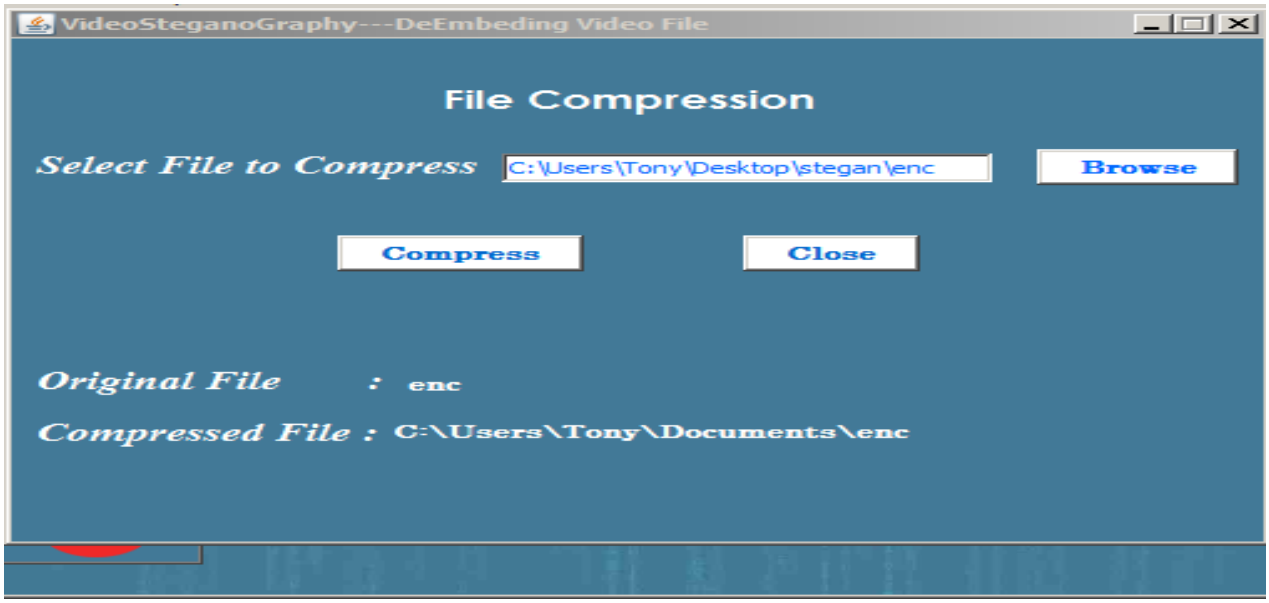

Figure 9: File Compression Page

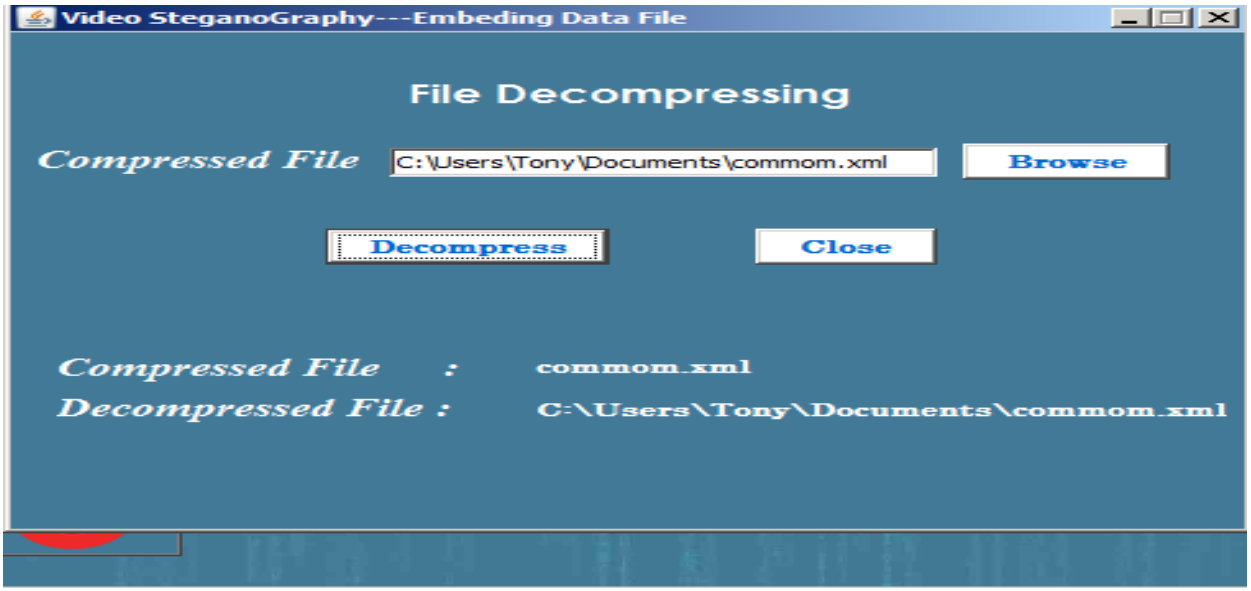

Figure 10: File Decompression Page

Table 2: Compression of same file type with different sizes

File Type	Original File Size	Compressed File Size	Percentage Reduced
Text File Type (.txt)	125	67	46.4
	118.3	65.1	45
	82.2	9.1	28.1
	50	38	24

Table 3: Compression different file type of same file size

File Type	Original File Size	Compressed File Size	Percentage Reduced
Text (.txt)	125	67	46.4
Database(.sql)	125	64.5	48.4
Image (jpeg)	125	69.3	44.6
PDF (.pdf)	125	67.9	45.7

Communications on Applied Electronics (CAE) – ISSN : 2394-4714
Foundation of Computer Science FCS, New York, USA
Volume 5 – No.10, September 2016 – www.caeaccess.org

Figure 11: Original video

Figure 12: Stego Video

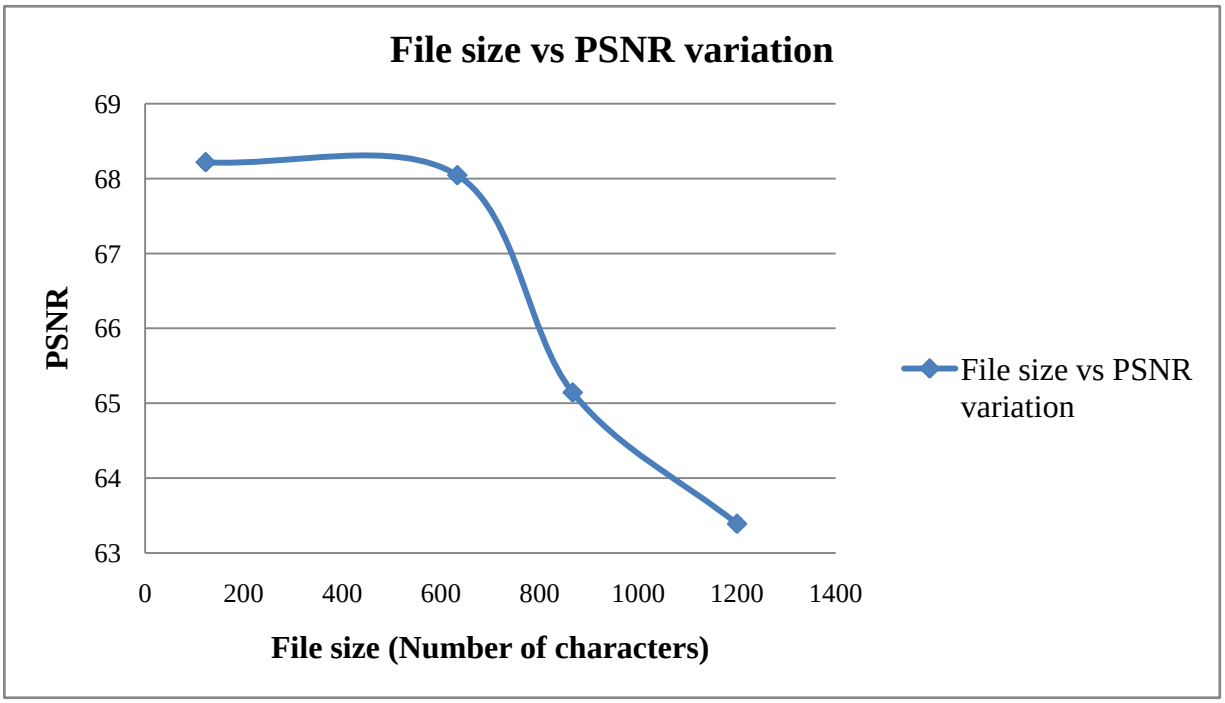

Figure 13: PSNR and File Variations

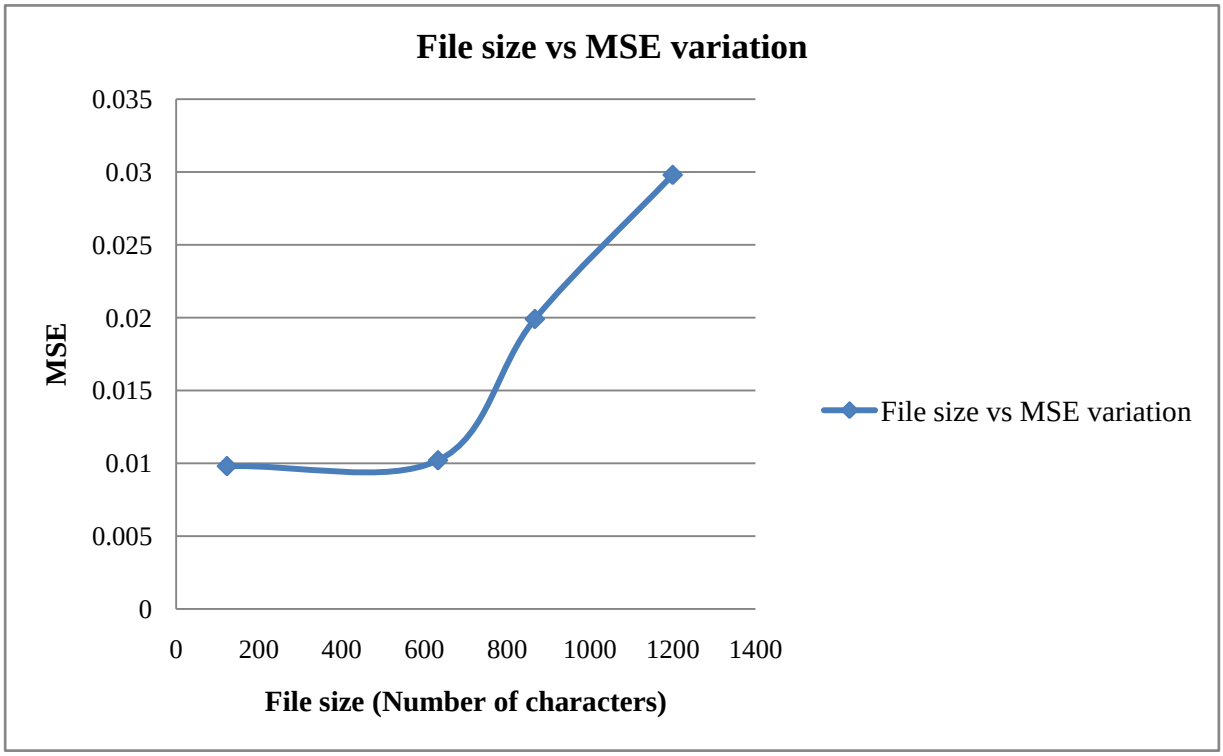

Figure 14: MSE Variation with different file sizes

5.2 Conclusion and Recommendation

The results obtained from the proposed system revealed some interesting findings. It is observed that the proposed system has low computational complexity. Thus, it is computationally less expensive to hide and extract the data. The low computational complexity is as a result of the use of LSB for embedding. LSB is believed to have low computational complexity and high embedding capacity. The proposed system is characterized with robustness, high embedding capacity and high security. The embedding capacity achieved in the proposed system is higher than anticipated. This shows that the proposed system is not prone to attack by hackers and intruders. The proposed system has the ability to withstand large embedding capacity without distortion in video quality. The system is very consistent and platform independent. The approach proposed in this paper achieved same output file size as the file input. This means that, file size does not increase in size after embedding which is normally not the case. This therefore shows the value of combining steganography, cryptography and compression. The re-encoding and reconstruction of the video frames resulted in the output of same file size. After going through the study successfully, the following recommendations are made. For the betterment of message security, robustness and capacity. RSA algorithm is generally slow in speed when implemented on CPU system, a study should therefore be conducted to see how possible best to enhance the speed of RSA algorithm for CPU implementation. Since spatial domain which LSB belongs to is susceptible to noise, transform domain techniques can be used in future studies. To further increase the security of data in the near future, studies must be conducted on how to increase the security of the communication medium for complete optimal security.

6. ACKNOWLEDGMENTS

Foremost appreciation goes to the Almighty God, the creator of Heaven and Earth for the knowledge bestowed upon us the grace to finish this work. Thanks to Dr. J. B. Hayfron-Acquah for his guidance, contributions, encouragement.

7. REFERENCES

[1] Odeh, A., & Elleithy, K. (2012). Steganography in Arabic Text Using Zero Width and Kashidha Letters. *International Journal of Computer Science & Information Technology (IJCSIT)*, *4*(3), 1-11

[2] Dengre , A. R., Gawande, A. D. Deshmukh, , A. B.(June, 2013). Effect of Audio Steganography based on LSB insertion with Image Watermarking using AVI video . International Journal of Application or Innovation in Engineering & Management (IJAIEM), 2(6), 2319 – 4847.

[3] Al-Vahed, A., & Sahhavi, H. (2011). An overview of modern cryptography. *World Applied Programming*, *1*(1), 3-8

[4] Kundalakesi,M., Sharmathi.R, Akshaya.R (2015) Overview of Modern Cryptography. International Journal of Computer Science and Information Technologies(IJCSIT), Vol. 6 (1), 350-353.

[5] Raghu, D., Rao, M. S., & Jacub, C. R. (2012).The File Encryption Method in Cryptographic Key Management Wireless Ad Hoc Networks. International Journal of Computer Science and Information technologies, vol 3(1), 3279-3282

[6] Venkateswaralu, S., Chhabra, N., & GNI, M. (2012). Secret Key Generation and Eavesdropping detection

using Quantum Cryptography. International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 3 (2) ,3348 – 3354.

[7] Bhaumik, A. K., Choi, M., Robles, R. J., & Balitanas, M. O. (2009). Data hiding in video. *International Journal of Database Theory and Application*, *2*(2), 9-16.

[8] Elbayoumy, M., Elmogy, M., Abouelfetouh, A., & Elhadary, R. (2014). A Proposed Technique For Hiding Data Into Video Files. *International Journal of Computer Science Issues (IJCSI)*, *11*(2), 68.

[9] Xu, C., Ping, X., & Zhang, T. (2006, August). Steganography in compressed video stream. In *Innovative Computing, Information and Control, 2006. ICICIC'06. First International Conference on* (Vol. 1, pp. 269-272). IEEE.

[10] Dasgupta, K., Mondal, J. K., & Dutta, P. (2013). Optimized Video Steganography Using Genetic Algorithm (GA). *Procedia Technology*, *10*, 131-137.

[11] Rhoads, G. B. (2007). "Video Steganography". *U.S. Patent No. 7,242,790*. Washington, DC: U.S. Patent and Trademark Office.

[12] Zhou, X., & Tang, X. (2011, August). Research and implementation of RSA algorithm for encryption and decryption. In *Strategic Technology (IFOST), 2011 6th International Forum on* (Vol. 2, pp. 1118-1121). IEEE.

[13] Ren, W., & Miao, Z. (2010, May). A hybrid encryption algorithm based on DES and RSA in Bluetooth communication. In *Modeling, Simulation and Visualization Methods (WMSVM), 2010 Second International Conference on* (pp. 221-225). IEEE.

[14] Sarmah, D. K., & Bajpai, N. (2010). A new horizon in data security by Cryptography & Steganography. *IJCSIT) International Journal of Computer Science and Information Technologies*, *1*(4), 212-220.

[15] Kaur, R. and Singh, T. (2015). Hiding Data in Video Sequences using LSB with Elliptic Curve Cryptography. International Journal of Computer Applications (0975 – 8887) Volume 117 – No. 18

[16] Singh, K.U. (2014) Video Steganography: Text Hiding In Video By LSB Substitution. International Journal of Engineering Research and Applications. ISSN : 2248-9622, Vol. 4, Issue 5(Version 1), pp.105-108

[17] Basheer, R & Safiya M.K (2014).Video data hiding in selective pixels of forbidden zone using mapping function. International Journal of Advanced Computer Technology (IJACT).ISSN:2319-7900.

[18] Apau, R., Hayfron-Acquah, J.B., and Twum, F. (June 2016). Enhancing Data Security using Video Steganography, RSA and Huffman Code Algorithms with LSB Insertion. International Journal of Computer Applications, 143 (4), 28-36.