Rashid Hussain. 2019, 7:1
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Machine Learning Technique to Hide Data Using Python

[1]Rashid Hussain, [2]Rabia Khan

Abstract

In this contemporary time, it is very difficult to secure data during transfer from one place to another. In today’s constrained environment, it is very easy to attack and compromise the security. So, more secure methods is required to escape from this condition. Hence, this paper is a proposed a new technique which is based on steganography and cryptography both. This technique is very helpful to ensures secure data transfer between the sender and receiver. It uses Discrete Contour Evolution Algorithm to Extract and insert frames and Transform Domain embedding to encode the message in video frames. This model is implemented in Python. Results received from this work are very good and better as compare to previous technique.

Keywords: Steganography Cryptography, Discrete Contour Evolution Algorithm

Introduction

Information security consists of identifying an organization's electronic informational assets, as well as the planning and programs that must be carried out to ensure its continued availability, confidentiality and integrity. Whether the organization is a commercial enterprise, governmental agency or educational institution, these goals are the same [1-2]. What differs is the type of assets and to what degree they are critical to the continued operation of the entity. As use of the Internet began to grow, organizations started to deploy firewalls at the perimeter to keep hackers from gaining access to the systems within. Most thought that we had handled the situation. Then the rise of the computer virus forced the development and deployment of anti-virus software onto workstations in order to protect the integrity of the data and the availability of systems themselves [3].

Today, the situation is not so simple. The current threats are entering from the Internet through our firewalls and landing directly onto PCs on the network. These threats include e-mail worms, remote access Trojans, spyware, adware, network worms, blended threats, as well as multistage, incremental infections using all of the above. Protective technologies consist of mainly two aspects: cryptography and steganography [4-5].

Cryptography is the science of writing in secret code and is an ancient art. Cryptography appeared spontaneously sometime after writing was invented, with applications ranging from diplomatic missives to war-time battle plans [6].

Steganography is the technique of hidden communication. Using steganography a secret message is embedded in a medium, such as an image or a sound or a video clip and sent. The existence of the hidden message is not known except by the sender and receiver. The word is derived from the Greek words stegos meaning covered and graphia meaning writing [7-8].

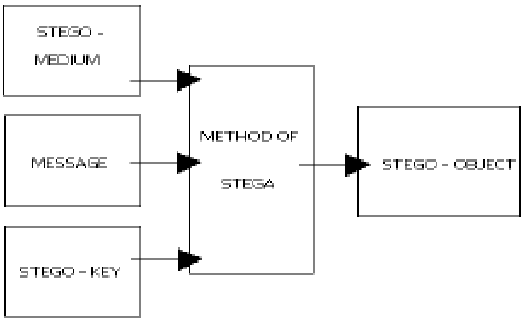


Fig. 1. Basic Steganography Embedding Process [9].

While cryptography is preoccupied with the protection of the contents of a message or information, steganography concentrates on concealing the very existence of such messages from detection .The term steganography is

10.2348/ijset070119045

adapted from the Greek word steganographia, meaning "covered writing", and is taken in its modern form to mean the hiding of information inside other information. Naturally these techniques date back throughout history, the main applications being in couriering information during times of war [10].

With the invention of digital audio and images files this has taken on a whole new meaning; creating new methods for performing "reversible data hiding" as it is often dubbed. This has many possible applications including the copyright watermarking of audio, video and still image data.

Related Literature

In the paper, a new method of steganography has been introduced. The tools have been divided into various parts. In this method, the key frames are extracted from the source video file using DCE algorithm. The message to be hidden is encrypted using the shared secret key. The encrypted message is embedded in key frames by using Transform Domain Embedding. The stego-key frames are inserted in the original video file. This stego-video file is sent to the receiver [11-13].

In order to extract a number of representative frames from the sequence main properties of **Discrete Contour Evolution (DCE)** are:

1. It leads to the simplification of curve complexity, in analogy to evolutions guided by diffusion equations, with

2. No blurring (i.e. peak rounding) effects and no dislocation of relevant features, due to the fact that the remaining vertices do not change their positions

3. The relevance measure K is stable with respect to noisy deformations, since noise elimination takes place in the early stages of the evolution

4. It allows us to find digital line segments in noisy metrics due to the relevance order of the repeated process of digital linearization [14].

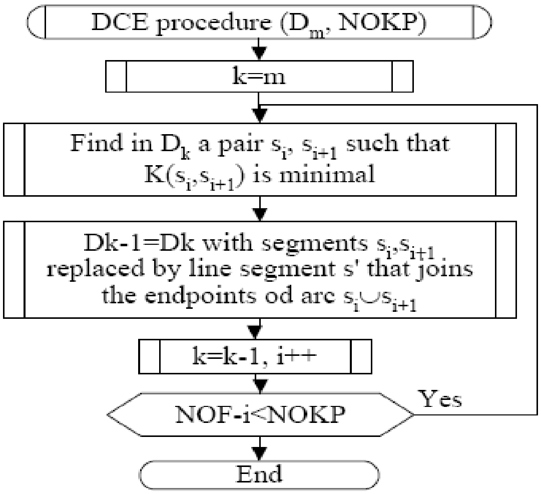
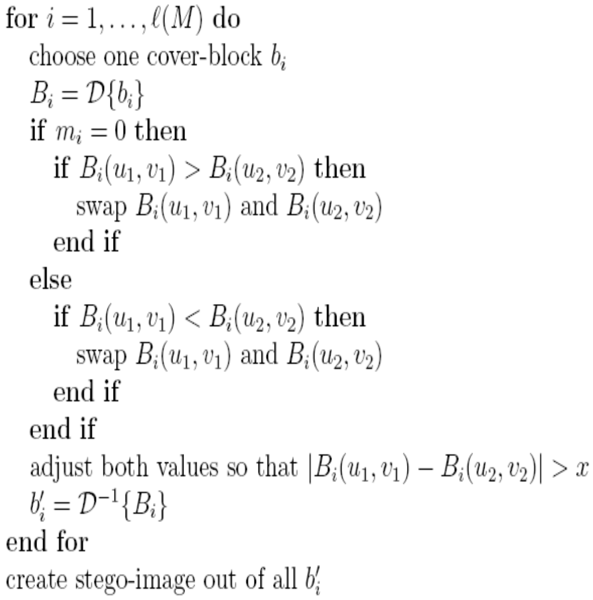

DCE procedure (D_m, NOKP)
k=m
Find in D_k a pair s_i, s_{i+1} such that $K(s_i,s_{i+1})$ is minimal
Dk-1=Dk with segments s_i,s_{i+1} replaced by line segment s' that joins the endpoints od arc $s_i \cup s_{i+1}$
k=k-1, i++
NOF-i<NOKP
Yes
End

Fig 2. Flowchart of the DCE algorithm [13]

The **Transform domain embedding algorithm** is described below as splitting the image into 8x8 blocks and calculating the DCT of the block. Then two middle-frequency (so that they are not to altered by the quantization/compression which will take place in JPEG) are chosen and agreed upon by both send and receive parties. A block encodes a 1 if DCT(a,b) > DCT(c,d) otherwise it encodes a 0. In the encoding step the coefficients are swapped if their relative size does not match with the bit to be encoded. Since the JPEG compression can affect the relative size of the coefficients the algorithm ensures that abs (DCT (a, b) - DCT(c, d)) > x where x is a value which represents the tradeoff between image quality and robustness. The algorithm is described below [15]:



```
for $i = 1, \ldots, \ell(M)$ do
  choose one cover-block $b_i$
  $B_i = \mathcal{D}\{b_i\}$
  if $m_i = 0$ then
    if $B_i(u_1, v_1) > B_i(u_2, v_2)$ then
      swap $B_i(u_1, v_1)$ and $B_i(u_2, v_2)$
    end if
  else
    if $B_i(u_1, v_1) < B_i(u_2, v_2)$ then
      swap $B_i(u_1, v_1)$ and $B_i(u_2, v_2)$
    end if
  end if
  adjust both values so that $|B_i(u_1, v_1) - B_i(u_2, v_2)| > x$
  $b'_i = \mathcal{D}^{-1}\{B_i\}$
end for
create stego-image out of all $b'_i$
```

Decoding is straightforward in that all available blocks are DCT-transformed and by comparing the coefficients of each block the information is restored. The problem with the method outlined above is that it does not discard the image blocks where the desired relation of DCT coefficients cannot be enforced without severely damaging the image in the specific block[16].

Video Steganography Program

A video steganography in Python framework has been developed. The steps are stated below[17]:

1. The key frames are extracted from the source video file using DCE algorithm.

2. The message to be hidden is encrypted using the shared secret key.

3. The encrypted message is embedded in key frames by using Transform Domain Embedding.

4. The stego-key frames are inserted in the original video file.

5. This stego-video file is sent to the receiver.

The carrier file must be a video file. It is concerned with embedding information in an innocuous cover media in a secure and robust manner.

Results

After few experiments with different abstraction rate and different video content, the conclusion is that the **Key-frame extraction algorithm** shows subjectively better results for news and soap operas, while the content of the commercials is presented reasonably well with our video summary generated from the extracted set of the key-frames. The Figure 3 shows three steps in DCE algorithm of a short commercial clip.

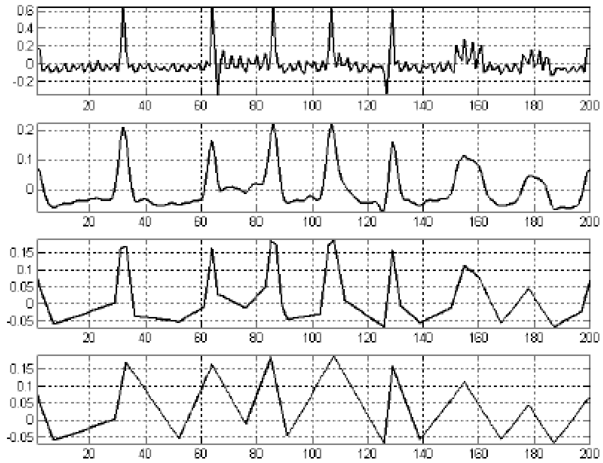
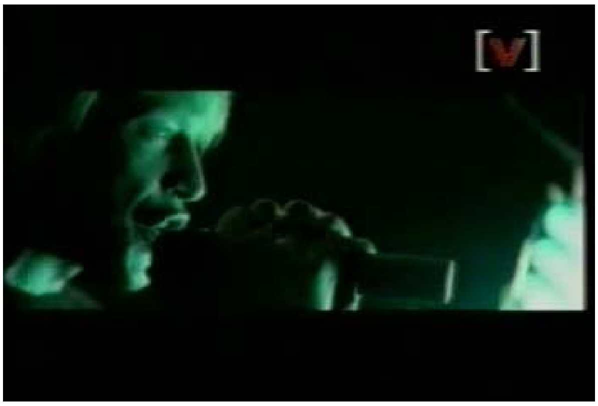


Fig 3. DCE algorithm results of a short commercial clip

The results after applying the **transform domain embedding** can be viewed in figure 4.



(a) Original key frame

(b) Key frame after embedding the message of size 4 kb.

(c) Key frame after embedding the message of size 64 kb.

Fig 4. Different embedding results on key frame

Cover Video File	Actual Length (MB)	Estimated Length (MB)
File 1	39	39
File 2	8	8.1
File 3	5	5.2
File 4	1981	1981

Table 1. Actual Length and Estimated Length of Hidden Bits

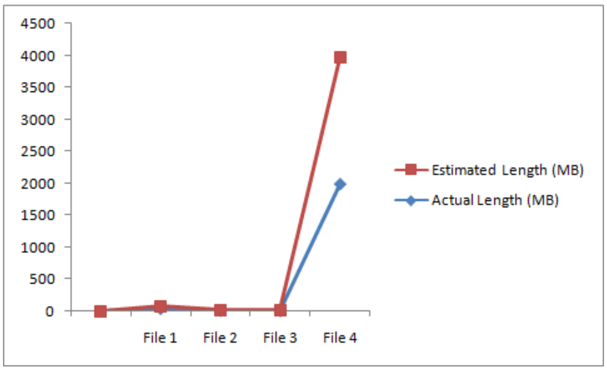


Fig 5. Graph representation between Actual Length and Estimated Length of Hidden Bits

In this above figure we can easily find out that the size of the data is not much change if we hid some data in these videos

Conclusion

In this paper, we presented a transform domain embedding algorithm for video steganography. It provides the user to give secret key for encryption. The length of message is more than the existing system. The quality of the video doesn't change variably. It cannot detect the lack in quality of video. The encryption key can be any combination of characters, symbols, or numbers. The key which is used for encoding is also used for decoding.

✓ This is a secret key where the both user have to agree upon a single common key. The improvement in robustness in presence of additive noise is obvious, as the proposed algorithm obtains significantly lower bit error rates than the standard algorithm. The resulting program is simple and easy-to-use and runs on Python framework. It offers three features: embedding any type of file in the video, extracting those same hidden data, and compressing the video to decrease its file size.

References

[1] A.J. Mozo, M.E. Obien, C.J. Rigor, D.F. Rayel, K. Chua, G. Tangonan "Video Steganography using Flash Video (FLV)", Electronics, Communications and Computer Engineering Department, Ateneo de Manila University, Proceedings of the IEEE, 5-7 May 2009, pp. 978-984.

[2] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding-a survey," Proceedings of the IEEE, Vol. 87, Issue 7, July 1999, pp. 1062-1078

[3]. J.C. Judge, "Steganography: Past, Present, Future," SANS white paper, 30 November, 2001, http://www.sans.org/rr/papers/index.php?id=552.

[4] Kavitha, R. and A Murugan, "Lossless Steganography on AVI File using Swapping Algorithm," SRM University. Conference on Computational Intelligence and Multimedia Applications, 2007. International Conference on Volume 4, 13-15 Dec. 2007 Page(s):83 – 88.

[5] F. Hartung and B. Girod, "Copyright Protection in Video Delivery Networks by Watermarking of Pre-Compressed Video," Proceedings of Multimedia Applications, Services and Techniques - ECMAST '97', Springer Lecture Notes in Computer Science, Vol. 1242, Milan, Italy, May 1997, pp. 423-436.

[6] K. Papapanagiotou, E. Kellinis, G. F. Marias, and P. Georgiadis, "Alternatives for Multimedia Messaging System Steganography," Proceedings of the IEEE International Conference on Computational Intelligence and Security (CIS 2005), Part II, LNAI 3802, Xian, China, December 2005, pp. 589-596 Authorized

[7]K. Wong, K. Tanaka, K. Takagi, and Y. Nakajima, "Complete video quality-preserving data hiding,"

IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, pp. 1499-1512, 2009.

[8]Saif Alzhair and Arber Borici, "An Innovative Lossless compression Method for Discrete-Color Images," IEEE Transactions on Image Processing vol. 21 Issue 1 January 2015.

[9]Shivani khosla and Paramjeet Kaur, " Secure Data Hiding Technique Using Video Steganography and Watermarking - A Review " IJCA (0975-8887)Volume 95-No 20,June 2014

[10]Usha B.A, N K Srinath and N K Cauvery" Analysis Of Video Steganalysis Techniques To Defend Against Statistical Attacks – A Survey", IJRET, ISSN 2319-1163.

[11]RaviKumar and P.R.K Murti"Data Security and Authentication Using Stegnaography,"IJCSIT Vol 2(4),2011, ISSN :0975,pp. 1453-1456.

[12]Sonali S.Ekhande Prof. S.P.Sonavane Dr. P .J .Kulkarni" Universal Steganalysis using Feature Selection Strategy for Higher Order Video Statistics"IJCA(0975-8887) Volume 1 No-19 2013

[13]Rajalakshmi, K., Dr.K. Mahesh, 2016 "A Review on Video Compression and Embedding Techniques"International Journal of Computer Applications (0975 – 8887) 141: 12.

[14]P. K. Shivani Khosla, "Secure Data Hiding Technique Using Video Steganography and Watermarking," International Journal of Computer Applications, vol. 95, pp. 7-12, 2014.

[15]Teena M. Thomas "Efficient video watermarking with SWT and empirical PCA based decoding" IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-8727 Volume 16 Issue 5,2014.

[16] K.Rajalakshmi, Dr.K.Mahesh, "Video Embedding with Compression Based on Patchwise Code Formation", Australian Journal of Basic and Applied Sciences, Vol.10, No.13 (August), 2016, ISSN :

[17] P. Roy and A. Nath, "New Steganography approach using encrypted secret message inside Audio and Video media," International Journal of Advance Research in Computer Science and Management Studies" vol. 2, pp. 46-59, 2014.

Author's Profile

[1]Department of Mathematics and Computer Science, Sule Lamido University, Kafin Hausa, Jigawa State, Nigeria, Email: rashid65_its@yahoo.com

[2]MCA, Punjab Technical University, India, Email: r_khan01@yahoo.com