LCS based Secure Authentication Scheme for Health Care Information Exchange

B. Anne Vinitha Vardhini, T. S. Sivapriya & S. Poonkuntran

Velammal College of Engineering and Technology, Madurai-09, Tamilnadu, India.
E-mail : annbernardz@gmail.com, technosivapriya@gmail.com,s_poonkuntran@yahoo.co.in

Abstract **– Information Hiding has captured the imagination of many researchers. Digital watermarking and steganography techniques are used to address digital rights management, protect information, and conceal secrets. Information hiding techniques provide an interesting challenge for digital forensic investigations. Information can easily traverse through firewalls undetected. Research into steganalysis techniques aids in the discovery of such hidden information as well as leads research toward improved methods for hiding information. Medical data is the one that has to be given high level of security, because of the significant information it holds. Thereby, various researches have been done and it is in progress to provide security mechanism to the data. In this paper, Longest Common Subsequence (LCS) has been used as a tool along with integer transform to provide a highly secure way of transmitting the medical data. This approach mainly ensures combined binding of the data such that the data remains protected and any unauthorised access to the data can also be easily detected. This scheme achieved PSNR value around 53 dB attaining a fair level of imperceptibility. When subjected to jittering attack the modified range of bits is around 33%.**

Keywords – Information Security; Text Steganography; Information Hiding; watermarking; Longest Common Subsequence; Digital fundus image; Integer transform

I. INTRODUCTION

The worldwide use of computer technology has its roots in the healthcare sector too. Many hospitals have embraced the use of information technology to record, create, store, retrieve and process the various medical information. The complexity of modern healthcare has made IT inevitable in playing a significant role in improving the health care quality. With the emergence of electronic medical records and its uses, the field became active and seeking the practical application of computing and communications technology to healthcare, health education and biomedical research [1-25].

According to the Institute of Medicine (IOM) Committee on Quality Health Care in America, IT must play a central role in the redesign of the health care system if a substantial improvement in quality is to be achieved over the coming decade [14]. It focuses on researches to improving the clinical decision making, clinical information management, communication, costs, and access to care.

Such IT revolutions in healthcare are also creating an environment where these information can easily be accessed, replicated and distributed without any loss in quality. Hence, there is a need of technologies for secret communications and to ensure the security of content that is being exchanged.

The information has been recognized as a technique to provide secret communications and to ensure the authentication while the health care information is being exchanged.

While considering the transmission of medical information over the internet confidentiality and data integrity must be taken into consideration so as to ensure protection of the information from unauthorised access [17]. The main root of medical identity theft is the falsification of the medical charts which usually happens by accessing the information based on legitimate needs, dumpster driving and hacking of computerised records. False entries in victim's medical record may remain in victim's medical files for years and may not be corrected or even discovered. False entries in victim's medical record may cause the victims to receive wrong medical treatment. Victims may find their health insurance exhausted, and become uninsurable for both life and health insurance coverage. Therefore security is

an important issue in the transmission of the medical data and reports via internet [6].

This paper presents a novel authentication scheme for health care information exchange. The reversibility, imperceptibility and fragility have been emphasized on the scheme, since it is proposed for healthcare information exchange.

II. DIGITAL FUNDUS IMAGES

Medical data usually takes the form of narrative/textual, recorded signals, audio and video formats. Compared with text information, image information has its own unique properties such as strong correlation, high redundancy and so on [19]. Digital fundus images are those which are used by ophthalmologists for diagnosis, monitoring and screening of the eye through the photograph of the interior surface of the eye such as macula, optic disc, posterior pole and retina. An example of a fundus image is shown in fig. 1.

Fig.1 : A Sample Digital Fundus Image.

In ophthalmology, the fundus is the interior surface of the eye, including the retina, optic disc and macula. The fundus can be viewed with an ophthalmoscope. The fundus images are taken using fundus camera. A fundus camera is a specialized low power microscope with an attached camera designed to photograph the interior surface of the eye. Fundus photograph is usually taken using a green filter to acquire images of retinal blood vessels [9]. Green light is absorbed by Blood and appeared darker colour in the fundus photograph than the background and the Retinal nerve fibre layer. Hence, the green channel of the fundus images possesses the valuable Information for diagnosis than other channels [21]. There has been a significant rise in the overall number of ophthalmic examinations which relies upon improving the ocular health care of the people but this has also forced the ophthalmologists to be pressurised. To address this issue, modern health care systems have been developed with Computer Aided Design (CAD) and networking for analysing the fundus images resulting in drastic improvement in the screening accuracy. These modern health care systems require highly secure communication techniques for exchanging information from one place to other [22].

III. RELATED WORKS

There are plenty of algorithms available for implementing steganography and a detailed survey has been conducted on the various works that has been done so far in the area of medical image security. For proposing this scheme which satisfies the various properties of watermarking such as reversibility, fragility, imperceptibility and so on, the studies mentioned below have their own unique contribution in understanding the existing available techniques. One among those is the LSB (Least Significant Bit) algorithms where the least significant bit (LSB) is the position of the bit in a binary integer giving the units value and determining whether the number is even or odd during computing. It is referred to as the right-most bit, due to the convention in the positional notation of writing less significant digit further to the right [13]. In VoIP steganography, proposed by there has been a modified the threshold secret sharing scheme which applies a two phase approach on the steganography mechanism to provide reliability and fault tolerance and to increase steganalysis complexity [4].

Yeung and Mintzer [10] made use of a look-up table (LUT) to embed a watermark. But the accuracy is not expected and external memory is needed for LUT. Wu designed a LUT for DCT coefficients to embed a binary watermark. The chaotic model based semi fragile watermarking technique uses Integer transform which is used for obtaining reversible image watermarking [8]. Also block linking method is widely used for the same.

Generalized collage steganography uses the capacity of cover images in the database for analysis in advance in order to select a proper cover according to the secret message [6]. The process involved in encoding and decoding uses a blend of media cryptography and asymmetric cryptographic algorithms called as visual cryptographic steganography method. Nighat mir proposed a Web based secure communication which makes use of Xml file as a carrier with cover data which is further encrypted using AES where text information is converted into a string message[11]. To ensure secure private communication between the users and to hide large amount of data with minimal change of the structure of the cover file is done through LCS based steganography on Indian languages.

IV. PROPOSED SCHEME

The key concept used in the proposed scheme includes Longest Common Subsequence (LCS) and Integer Transform. The LCS plays a vital role in finding

a key string input value that can be used for embedding purpose. Integer transform is the mathematical tool used to embed the bits and inverse integer transform is used on the extraction arena.

A. LCS(Longest Common Subsequence)

The Longest Common Subsequence (LCS) is a mathematical function that is used to calculate the longest subsequence common to all sequences in a set of sequences but it is different from a substring [1]. Let us define two sequences as $X = (x_1, x_2...x_m)$ and $Y = (y_1, y_2...y_n)$. The prefixes of X are $X_{1, 2,...m}$ and the prefixes of Y are $Y_{1, 2,...n}$. Then $LCS\ (X_i, Y_j)$ represent the set of longest common subsequence of prefixes X_i and Y_j [1]. To find the longest sub sequences common to X_i and Y_j, compare the elements x_i and y_j. If they are equal, then the sequence $LCS\ (X_{i-1}, Y_{j-1})$ is extended by that element, x_i. If they are not equal, then the longer of the two sequences, $LCS\ (X_i, Y_{j-1})$, and $LCS(X_{i-1}, Y_j)$, is retained. If they are both the same length, but not identical, then both are retained. Note that the subscripts are reduced by 1 in these formulas [1]. That can result in a subscript of 0. Since the sequence elements are defined to start at 1, it was necessary to add the requirement that the LCS is empty when a subscript is zero. The reason why LCS has been used for solving problem is because of the various motivating applications it has such as molecular biology, file comparison, and screen redisplay and so on. Since LCS uses dynamic programming technique it has reduced space complexity [25]. LCS has been used as a function for embedding because of the uniqueness of input generation in spite of the variety of medical and secret images used. The algorithm that is used for LCS calculation is as follows:

```
Algorithm LCS(X, Y):
m = length[X]
n = length[Y]
for k= 1 to m
        c[k,O] = 0
        for 1 = 1 to n
                c[O,l] = 0
        for k = 1 to m
                for 1 = 1 to n
loop
        if X[k] = Y[l]
        LCS[l] = I[k]
        c[k,l] = c[k-1 ,1-1]+ 1
        b[k,l] = '1'
        else if c[k-l,1]>=c[k,1-1]
        c[k,l] = c[k-1,1]
        b[k,l] = 'u'
        else
        c[k,l] = c[k,1-1]
        b[k,l] = '1'
        end if
end loop
        return b, c, LCS [ ]
end.
```

B. Integer transform

This paper uses integer transforms to hide the watermark in the image for the following reasons [9].

1. The transform is reversible. This is a main requirement given to medical images.
2. The difference between the pixel pair is expanded to hide the watermark information. In the extraction, the original pixel pair is exactly retrieved, when the watermark is removed.
3. It can be used in multilayer. Hence, the high capacity is achieved. Note that capacity and imperceptibly are in trade-off. Therefore, capacity should be identified optimally before embedding.

For a 8 bit gray scale pixel pair *(x, y)*, $0 \le x, y \le 255$, the integer transform is given by the pair *(m, d)*.Where *m* refers integer average and *d* refers difference.

$$m = \left\lfloor \frac{x+y}{2} \right\rfloor \quad (1)$$

$$d = x - y \quad (2)$$

The inverse transform is given by

$$x = m + \left\lfloor \frac{d+1}{2} \right\rfloor \quad (3)$$

$$y = m - \left\lfloor \frac{d}{2} \right\rfloor \quad (4)$$

Where $\lfloor\ \rfloor$ refers floor operation which rounds the value to nearest integer. In the integer transform, the difference *(d)* is modified based on the watermark bit

(bit) to hide the *bit* into the pixel pair. The modification of difference *(d')* is given by

$$d' = 2 * d + bit \quad (5)$$

The modification process requires overflow and underflow conditions to ensure that the difference is expandable or not. The expandable difference is satisfies the following condition.

$$\begin{aligned} |d'| &\le 2 * (255 - m) \qquad if\ 128 \le m \le 255 \\ |d'| &\le 2 * m + 1 \qquad if\ 0 \le m \le 127 \end{aligned} \quad (6)$$

Only expandable difference can be used for embedding. If all the expandable differences are used, the capacity will reach its limit. Let N and Ne denote the number of differences and the number of expandable differences, respectively. The hiding capacity of an image is defined as:

$$c = \frac{Ne}{N} \quad (7)$$

For multilayer embedding, the same pixel pairs are selected for further data embedding. Here some of the differences may not be expandable for longer time [9].

C. Proposed scheme

The proposed scheme mainly ensures combined binding that satisfies the goals of watermarking techniques like availability, authenticity, confidentiality, copy control, copyright protection, integrity and utility [24]. As like always the implementation can be done in two major steps of embedding and extraction. In the embedding section, Original image (Digital fundus image) and the secret image (logo) are taken as the inputs. The secret image is likely to be taken as one-fourth of the original image. First step is to find the LCS between the two inputs string1 and string2. The LCS output is used as a tool for embedding process on the original image using the integer transform function. The LCS value will be available to both the sender and the receiver. The extraction section considers the embedded image as the input. Here the inverse integer transform is used to extract the original and the secret image from the embedded image.

For the proposed scheme mentioned in this paper we have the following steps:

Embedding Process

Embedding involves usage of red and blue planes where in which the green plane pixels remain unused. But the green plane pixels are used in finding the string inputs for calculating LCS.

Step 1: Finding String 1 input

The original image is used to find the string1 input. Here the green plane of the original image is considered to construct string1. The green plane pixel values are divided into blocks of 4 rows. In every block, each of the pixel values are converted into the 8 bit binary and the LSB [2] is taken for each block in a zigzag pattern in order to get a single row output from every block. Thereby an output wherein which the number of rows of string1 which will be equal to the number of rows of the secret image (logo) is obtained.

Step 2: Finding String 2 input

The secret image is used for finding string2. The secret image is converted into a binary image [16]. This binary image is taken as string2.

Step 3: Finding LCS

The number of rows of string1 and string2 will be of equal size. LCS is taken between each row of string1 and string2.

Step 4: Embedding using integer transform

The LCS and the original image are the inputs for the integer transform function. Here the original image is divided into three planes. Only the Red and blue planes are used in integer transform.

Step 5: Creation of location map

Location map is created for the embedded image to identify the pixel values that has undergone modification during embedding process [20].

Extraction Process

In the extraction process, the watermarked medical image is processed in the same way as original image processed for embedding. Both original image and original watermarks are not used for the extraction process, in that the original image and the location map created is used for extraction. The extraction process works as follows. The extraction process is done by using the inverse integer transform function. The input for the inverse integer transform is embedded image. The embedded image is divided into three planes. The extraction process is done by using the following formula.

By using the inverse integer transform formula, the extracted image is obtained. This image will be equivalent to the original image, and also the extracted watermark will be similar to the calculated LCS.

V. RESULTS AND ANALYSIS

For the quantitative analysis of the proposed scheme, the test bed has been created using the digital fundus images taken from STARE and DRIVE databases [9] . The test images used is the experiment was taken in the size of 585 x 565 x 3 and in the Tagged Image File format (TIFF). The secret images were taken in the one fourth of the size of the original fundus image i.e. 146 x 140. All the simulation has been conducted on MATLAB R2010a. The test bed contains 10 fundus images and 10 secret images.

A. Reversibility

The watermarking scheme should be reversible. It means that the original image should be retrieved without any loss after removing the watermark. It is very important in medical images, since these images are exchanged for diagnosis purpose where loss in the quality of images is not compromised [1] [9-25].

It is found that the proposed scheme is reversible for any size of watermark. The Peak Signal to Noise Ratio (PSNR) has been used as metric for measuring the reversibility. The PSNR between original image (I) and extracted image (I_{ext}) is calculated as

$$PSNR\ (I, I_{ext}) = 10\ Log_{10}\ [(2^p\text{-}1)^2 / MSE] \qquad (8)$$

B. Imperceptibility

To measure the imperceptibility of the proposed scheme the above mentioned PSNR has been used between original images (I) and extracted image (I_{emb}). The experiments were conducted on a 10 test fundus images and 10 secret images in 10x10 combinations, totally 100 samples of values recorded. The proposed scheme gives around 53dB at an average of imperceptibility for the secret data size of above 20000 bits. The sample values are tabulated in tab. 1.

Table 1 The Imperceptibility of the Proposed Scheme.

Test Images	Average PSNR value
Test image 1	48.2964
Test image 2	51.61695
Test image 3	60.1056
Test image 4	46.74045
Test image 5	56.017
Test image 6	53.8823
Test image 7	53.0651
Test image 8	50.5658
Test image 9	59.9386
Test image 10	52.92295

C. Fragility

To measure the fragility, jittering attack has been taken in to the test bed and the experiment conducted on all the 10 test fundus images in various modification rates starting from 5% to 95%. It has been done for above mentioned 100 combinations, totally 1900 sample of values recorded. The proposed scheme provides 33% of fragility at an average. The recorded average values of various rates of modification are listed in tab. 2.

VI. CONCLUSIONS

This paper presents a novel authentication scheme for health care information exchange. It uses Longest Common Subsequence (LCS) for obtaining the secret data to be embedded inside the original data. The LCS is calculated between the original image and original secret image. The calculated LCS will then be embedded inside the original image using integer transform. The proposed scheme has been simulated on digital fundus images taken from STARE and DRIVE databases. The experimental results showed that the proposed scheme is reversible and provides around 53dB of imperceptibility. The proposed scheme is fragile against jittering attacks; thereby it modifies around 33% of secret data to ensure the authentication.

Table 2 The fragility rates of the proposed scheme for various rates of modifications in the embedded image.

Percentage of modification	Fragility
5	4.8054
10	9.6068
15	13.8725
20	17.886
25	21.8514
30	25.4875
35	28.8415
40	32.0019
45	34.9031
50	37.4926
55	39.7836
60	42.1494
65	43.8801
70	45.506
75	46.9286
80	47.9591
85	48.6889
90	49.4902
95	49.8155

VII. REFERENCES

[1] S.Changder, D. Ghosh,"LCS based text Steganography through Indian languages", in proceeding 2nd Information Hiding Workshop, vol. 1525, pp. 306-318, 1998

[2] R Chandramouli, N. Memon, "Analysis of LSB Based Image Steganography Techniques", IEEE pp. 1019-1022,2001.

[3] N.F. Johnson, S. Jajodia, "Staganalysis: The Investigation of Hiding Information", IEEE, pp. 113-116, 1998.

[4] Mohammad Hamdaqa, Ladan Tahvildari , "ReLACK: A Reliable VoIP Steganography Approach", Fifth International Conference on Secure Software Integration and Reliability Improvement,pp.277, 2011

[5] G. Simmons, "The prisoners problem and the subliminal channel," CRYPTO, pp.5I-{j7, 1983.

[6] Mei-Ching Chen, Sos S. Agaian, and C. L. Philip Chen, "Generalized collage steganography on images" Proc. 34th Ann. Hawaii Int'l Conf. System Sciences, IEEE CS Press, Los Alamitos,Calif., 2008.

[7] Piyush Marwaha, Paresh Marwaha,"Visual cryptographic steganography on images" IEEE Second International conference on Computing, Communication and Networking Technologies, pp. 1034-1041, 2010.

[8] Claerhout B, De Moor GJE. Privacy protection for clinical and genomic data: The use of privacy-enhancing techniques in medicine. Journal of Medical Informatics. 2005;74:257–265.

[9] S. Poonkuntran & R. S. Rajesh"Chaotic model based semi fragile watermarking using integer transforms for digital fundus image authentication" Journal: Multimedia Tools and Applications, Vol. 51, no. 3, 2012.

[10] Wu M, Liu B (1998) Watermarking for image authentication. In: Proceedings of the IEEE International Conference on Image Processing, Chicago, Illinoise,US, pp 437–441

[11] P. Wayner, "Mimic functions", Cryptologia XVI, pp. 193-214, July 1992.

[12] M. T. Chapman, "Hiding the hidden: A software system for concealing ciphertext as innocuous text", Master's thesis, University of Wisconsin-Milwaukee, May 1997.

[13] Peng Meng, Liusheng Huang, Zhili Chen, Wei Yang, Dong Li," Linguistic Steganography Detection Based on Perplexity", International Conference on MultiMedia and Information Technology, pp 217-220, 2008.

[14] T.Moeriand,"Steganography and Steganalysis", May 15, 2003, www.liacs.nl /home /tmoerianiprivtech.pdf.

[15] Nighat Mir , Sayed Afaq Hussain, "Secure web-based communication", Procedia Computer Science 3 (2011) 556–562.

[16] A.M. Alattar and O.M. Alattar, "Watermarking electronic text documents containing justified paragraphs and irregular line spacing ", Proceedings of SPIE - Volume 5306, Security, Steganography, and Watermarking of Multimedia Contents VI, June 2004, pp. 685-695.

[17] Neil F Johnson, "http :// www. jjtc. com / Steganography / information hiding and digital watermarking".

[18] D. Artz, "Digital Steganography: Hiding Data within Data", IEEE Internet Compufing, pp. 75-80, May-Jun 200 I.

[19] Ram kumar karki" http:// www.healthnet. org.np/reports/bpklcos/mrecord.html"

[20] K. Rabah, "Steganography-The Art of Hiding Data", Information Technology Journal, vol. 3, Issue 3, pp. 245-269, 2004.

[21] Poonkuntran S, Aju D, Anitha C (2007) A Smart system for retinal blood vessel identification and width computation. Proceedings of International conference on Trends in Intelligent Electronic Systems

[22] Fujita H et al (2008) The emerging of three CAD systems induced by Japanese health care needs. Comp Meth Prog Biomed. Doi:10.1016/j.cmpb.2008.04.003

[23] S. Changder, N.C. Debnath , "An Approach to Bengali Text Steganography", Proceedings of the International Conference on Software Engineering and Data Engineering (SEDE-08), ISBN: 978-1-880843-67-3, pp. 74-78, July, 2008, Los Angeles, California, USA.57

[24] S. Changder, N.C. Debnath , "A new approach for steganography in Bengali text", Journal of Computational Methods in Science and Engineering(JCMSE), lOS Press, ISSNI472-7978 , Pageslll-122,2009.

[25] S. Changder, Narayan C. Debnath, "New Techniques and Algorithms for Text Steganography through Hindi Text." Proceedings of the International Conference on Software Engineering and Data Engineering (SEDE-09), ISBN: 978-1-880843-71-0, pp 200-204 June2009, Las Vegas, USA.

◈◈◈