Secure Processing and Delivery of Medical Images for Patient Information Protection

Ming Yang[1], Lei Chen[2], Shengli Yuan[3], and Wen-Chen Hu[4]
[1]School of Computing and Software Engineering, Southern Polytechnic State University, Marietta, GA, USA
[2]Department of Computer Science, Sam Houston State University, Huntsville, TX, USA
[3]Department of Computer and Mathematical Sciences, University of Houston Downtown, Houston, TX, USA
[4]Department of Computer Science, University of North Dakota, Grand Forks, ND, USA

Abstract - *In the delivery of medical imaging (such as X-ray, MRI) for remote diagnosis, the protection of the security and privacy of patient's information is extremely important. As conventional E-mail delivery is considered insecure, nowadays, people send medical images to a remote location using secure shared network storage space over IP protocol. While this is more reliable than traditional E-mail delivery, it introduces higher costs and dedicated devices. In this study, we propose a reliable and economical E-mail delivery approach which ensures the security and privacy of imaging contents and patient information. In the proposed methodology, patient information within the medical images (host image) is encrypted and embedded. Consequently, confidential data will not be visually available to unauthorized personnel. In order to further ensure the secure delivery of the medical images via E-mail over public network such as the Internet, the proposed system utilizes non-web-based Secure E-mail transmission using the Enigmail security extension installed on E-mail client software Mozilla Thunderbird. Thunderbird, Enigmail security extension, and Enigmail's essential component GNU Privacy Guard (GnuPG), are all open source and freely available online. With this system, any medical image that requires electronic transmission will have the patient's information protected, and will be readily available immediately upon the delivery at the destination. This system is an economical and reliable alternative to the IP-based delivery.*

Keywords*:* Encryption, Privacy, HIPAA, Information Hiding.

1. Introduction

The Health Insurance Portability and Accountability Act (HIPAA) [1] requires that medical providers and insurance companies implement procedures and policies to protect patient's medical information. Areas to be specifically addressed include ensuring that confidential data is secured during electronic transmission, and that access is limited only to authorized personnel. Today, as remote diagnosis is becoming increasingly popular, medical images, such as X-ray or MRI, often need to be delivered from one location to another. This imposes new challenges that need to be faced:

(1) The patient information is usually printed in the corner of the medical images for viewing. As a result, it is easily accessible to anyone and may be intercepted by a third party in the course of electronic transmission.

(2) For scenarios such as medical imaging research, the patient information should not be accessible either.

(3) Traditionally, medical images are delivered through printed films or in burned CDs. This is neither secure nor reliable.

(4) Nowadays, people start to send medical images to a remote location over IP protocol, or use shared network storage space. This is more reliable than traditional approach, but it introduces higher costs and needs dedicated devices.

(5) Traditional E-mail transmission of medical images is generally considered to be insecure.

In order to address these issues, we have proposed a reliable and inexpensive approach to ensure the electronic delivery of medical images while securing the confidentiality of imaging contents and patient information. We have developed an information hiding methodology that makes use of the RSA encryption algorithm and a Discrete Cosine Transform (DCT) based hiding technique. As a result, patient information will not be visually available to unauthorized personnel. To ensure the secure delivery of the medical images, our proposed system utilizes non-web-based Secure E-mail Transmission using Mozilla Thunderbird with its security extension Enigmail and the core security component GNU Privacy Guard (GnuPG), all of which are open source and freely available on the

Internet. Using this approach, any medical image that requires electronic transmission will have the patient's information protected, and will be readily available immediately upon the delivery at the destination. Secure E-mail delivery is also feasible through this approach.

2. System Overview

In the proposed system, patient information is not automatically visibly displayed in the corner of the medical image. Instead, this information is first encoded using ASCII character-encoding scheme and encrypted into a non-recognizable format using the RSA encryption algorithm.

Next the patient information is further secured by embedding it within a section of the image that is outside the Region-Of-Interest (ROI) [2]. This ensures that the encoded and encrypted information is embedded in a location that will not affect the image quality and further diagnosis. The area outside the ROI is located by using image segmentation techniques as discussed in Section III.A.

After the area outside of the ROI is located, the patient information (already encoded and encrypted) is embedded using a DCT domain methodology. This information hiding algorithm is robust enough that further attacks (including cropping, noise, lossy compression, etc.) will not remove the embedded information. This methodology effectively and securely protects patient information in situations of electronic transmission and medical imaging research.

The security of medical images and patient information is further reinforced using the free open source Enigmail [7] security extension, with its core component GnuPG [8], installed on E-mail client software Mozilla Thunderbird. All these three software components and applications will ensure the secure delivery of the medical images through E-mail transmissions.

The image can be viewed in one of two forms. If the viewer does not have the authority to access the patient's personal information, for example a medical or computer researcher (or network hacker for that matter), the image is viewed with no data displayed in connection to the image. On the other hand, if the viewer, such as the patient's doctor, has the authority to access the confidential information, it can then be extracted, decrypted, decoded, and displayed upon the image with the input of the correct encryption/decryption key. The above procedure can also be combined with a fragile watermark to validate data integrity. This approach is illustrated in Figure 1.

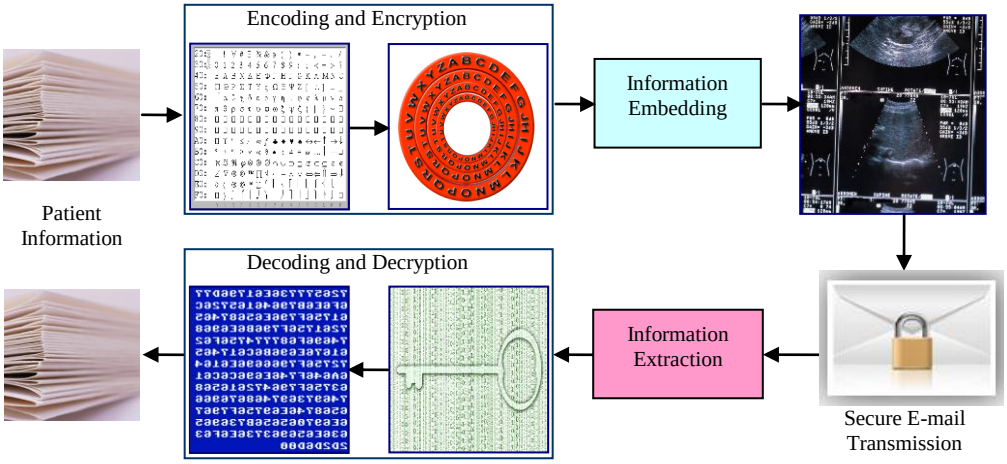


Figure 1. Flow Chart of Proposed Methodology

3. Implementation

In order to implement the proposed methodology, we first need to use image segmentation to identify the non-ROI region for information embedding, so that the embedded information will not affect the quality of the critical portion of the medical images. Next, patient information will be encoded, encrypted, and embedded for E-mail transmission. The patient information security system was implemented using MATLAB, a high-level programming language and interactive numerical computing environment. MATLAB has a wide variety of image processing capabilities and can process DICOM, BMP, JPG as well as other image formats.

3.1 Image Segmentation

Image segmentation is the process of dividing an image into sections, regions, or parts [2]. This process has numerous applications, such as automated inspection. Gonzalez gives the example that in the automated

inspection of electronic assemblies, image segmentation is used to find defects, such as a missing or broken path [2]. In respect to our project, we aim to identify the ROI, or the location where the actual picture is on the medical image. For example, if we have an X-ray of an elbow, our region of interest would be the elbow, NOT the black space surrounding it. Clearly, image segmentation is a vital part of this project. This process identifies the region of interest of an image, and draws a boundary within which the patient information should not be placed. This was achieved using MATLAB's built-in contour functions, and later implemented using a Java program. In the embedding procedure a simple image segmentation algorithm was employed to identify the ROI. Figure 2 is an x-ray of a skull that has been analyzed using image segmentation and has only the contour lines shown.

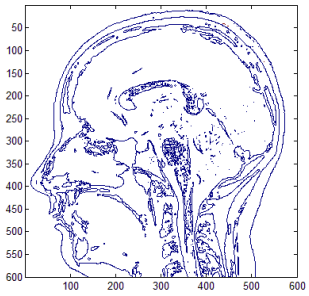


Figure 2. Image with Region of Interest (ROI) Boundaries

After image segmentation, the encoded/encrypted patient information is embedded in the portion of the background area that was determined outside of the ROI to preserve the quality of the host medical image [4].

3.2 Information Encryption

Patient data, delimited by commas and spaces, was read in from a test file for the first version of the system. In a later version the data was entered interactively by the user through a Graphical User Interface (GUI), written in Java and called by MATLAB.

The procedure to convert the text data to ASCII format in MATLAB resulted in seven-bit character strings instead of the expected eight bits. These strings then need to be broken apart and individually converted back into integer data types in order to perform the necessary mathematical operations for encryption and embedding.

The RSA encryption method was used to encrypt the patient's information. This particular method was chosen due to the simplicity of its algorithm. RSA is an asymmetric or public key algorithm, meaning it has both a public key and a private key [5]. The advantage of an asymmetric encryption lies in its higher level of security.

3.3 Information Embedding

Throughout the ages various methods have been devised to conceal information in transit. Tactics in previous times ranged from tattooing the message on a shaved head then waiting for the hair to re-grow before sending the message ([4]) to placing microfiche with the information under the postage stamp on a letter. With the creation of the Internet and other electronic data transmission mediums, steganography or the art of hiding information, has become even more important and commonplace.

Information can be hidden with success in text, image, audio/image, and protocol file formats. For image/video information hiding, there are two main groups of techniques: spatial domain algorithms and transform domain algorithms. Spatial domain algorithms generally involve manipulation of pixel intensity. Lossless image formats are most suited for spatial domain techniques [4]. The most well-known technique of information hiding in the image domain is Least Significant Bit (LSB) algorithm. Frequency domain algorithms try to modify the coefficients in the transform domain, which is more robust against transformation-based lossy compression [4].

3.4 Information Embedding and Extraction Algorithms

A high bitrate transform domain information hiding algorithm is designed to enable data embedding. In the proposed algorithm, a single bit is hidden within each 4x4 DCT coefficient block by means of vector quantization. Low-frequency coefficients are chosen for information hiding due to their relatively large amplitudes and the corresponding small step sizes in the quantization matrix [6].

The embedding algorithm is described in the following:

(1) DCT (4x4) transform of the original image;

(2) Scan the 4x4 DCT block along Zig-Zag scanning path;

(3) Convert the 8 low-frequency coefficients to an 1-D vector;

(4) V: the 1-D vector $V = (c_0, c_1, c_2, ..., c_6, c_7)$

T: the step size for vector quantization

|V|: the norm of vector V

[]: round-off operation

$l = |V| = \sqrt{\sum_{i=0}^{15} c_i^2}$: (norm of V)

$$l_T = \left[\frac{|V|}{T}\right] = \left[\frac{\sqrt{\sum_{i=0}^{15} c_i^2}}{T}\right]$$: (quantized norm of V)

(5) One single bit is embedded by modifying l_T: $l_T' = l_T \pm 0.25$ (+0.25 to embed 1, -0.25 to embed 0);

(6) $l' = l_T' * T$, $V' = \frac{l'}{l} * V$ (V' is the modified vector);

(7) Place the vector V' back to its original location in the 4x4 DCT block;

(8) Repeat the same operation for each 4x4 DCT block until all the information bits have been embedded.

The information retrieval algorithm is the following:

(1) DCT transform the stego-image (image with embedded information);

(2) For each 4x4 DCT block, scan the coefficients along Zig-Zag scanning path;

(3) Pick up the 8 low-frequency coefficients and convert them to a 1-D vector V";

(4) The norm of V" is: $l'' = |V''|$

(5) The quantized norm is: $l_T'' = \frac{l''}{T} = \frac{|V''|}{T}$

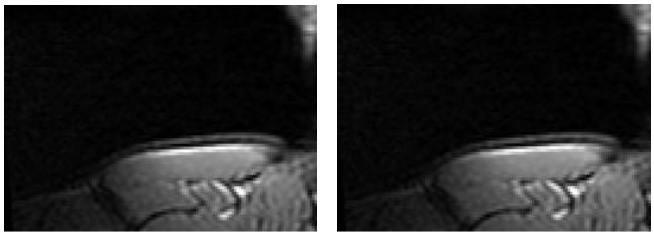
(6) $I = l_T'' - [l_T'']$

(7) If I>=0, then 1 is extracted as the information bit;

else (*I<0)*, then 0 is extracted as the information bit.

(8) Repeat the same operation to each 4x4 DCT block until all the information bits have been extracted.

With the embedding algorithm, the quality of the host image will not be visually degraded (Figure 3). Also, the hidden information can be extracted without the presence of



original image. This feature is extremely important in many applications. The proposed algorithm is also very robust to lossy compression, according to the experimental results.

Figure 3. Comparison of Original Image (left) and Stego-Image (right)

After the image segmentation, encoding, encryption, embedding procedures are completed, the medical image is ready for transmission. The transmission of the medical image with a secure E-mail approach will be discussed in Section IV.

With the supply of the correct decryption key, the extraction, decryption, and decoding of the data are simply the reverse of the embedding/encryption/ encoding procedures, with the addition of the display of the patient data below the image in the receiver's GUI. A copy of the program was placed on a remote computer and the full procedure was tested. The data file was encrypted and embedded into the image and transmitted via E-mail. The image was retrieved at the second computer, extracted, and decoded.

4. Secure E-mail Transmission

Not only is patient's personal information confidential and therefore prepared and embedded nicely using the proposed algorithm, the entire medical image also requires the guarantee of confidentiality and integrity on the path between sender and receiver. All of these security goals are achieved with the help from GNU Privacy Guard (GnuPG) [8], the core security component of Enigmail [7] extension installed on Thunderbird [11] E-mail client, as shown in Figure 4.

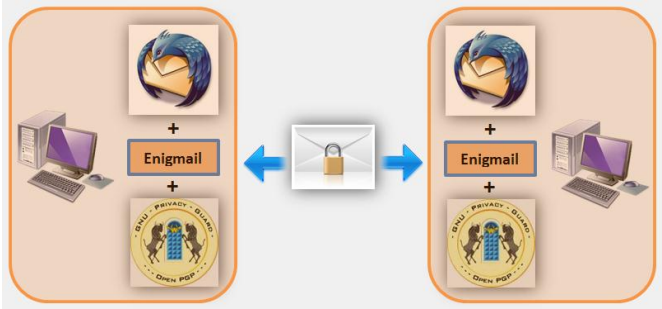


Figure 4. Securing Medical Images as E-mail Attachments using Thunderbird, Enigmail, and GnuPG

GnuPG, as described in OpenPGP standard [9] that it follows, makes use of both symmetric-key and public-key encryption to provide confidentiality. A unique random symmetric session key S created at the sender side, e.g. Dr. A, is used to encrypt the E-mail object content which consists of the message and medical images (with patient information embedded) as attachments. In order for the receiver, Dr. B, to be able to obtain key S in a secure manner and then decrypt the message and image attachments, S is encrypted at the sender using Dr. B's public key KB+, which can be posted on a public key server or sent directly to a sender via plaintext E-mails. To preserve the integrity of the E-mail message and the medical images, a message digest M is created using hash function SHA-1, then digitally signed by Dr. A with his private key KA- and attached to the E-mail object. As soon as the entire

E-mail arrives at Dr. B's computer, she can apply Dr. A's public key KA+ to what she has received. If the received message digest turns out to be the same as the SHA-1 hash value of the received E-mail object content, it proves that the E-mail message and all attachments have not been modified during transmission.

The stable version 1.0.4 version of GnuPG employed in our proposed system follows the Advanced Encryption Standard (AES) with 256-bit key. According to Bruce Schneier [10], current reported attacks can only break up to 11 rounds of the 14 rounds of AES-256. Therefore, the confidentiality of both E-mail messages and medical image attachments is ensured. Enigmail allows Thunderbird works seamlessly with GnuPG for protecting medical images in our system. The system uses Thunderbird for E-mails client software for several reasons: (1) it is free which adds no extra cost to the already expensive medical systems; (2) it supports most current operating systems; (3) it is open source that many plug-ins and extensions, especially those related to security, are freely available on the Internet.

5. Conclusion

In this study, we proposed a reliable and economical approach to ensure the electronic delivery of medical images while securing the security and privacy of image contents and patient information. The proposed methodology utilizes data encryption and high bitrate information hiding to ensure patient information security. It also makes use of secure E-mail transmission to ensure the secure delivery of medical images. It is a secure alternative of the existing medical imaging delivery approaches, such as server-to-server delivery through IP protocol. With this system, any medical image that requires electronic transmission will have the patient's information protected, and will be readily available immediately upon its delivery at the destination.

6. References

[1] "Health Insurance Portability and Accountability Act (HIPAA) and Its Impact on IT Security," Regulatory Compliance Series 3 of 6, Apani Networks White Paper Compliance Series. May 12, 2005. http://www.apani.com.

[2] R. C. Gonzalez, and R. E. Woods. "Digital Image Processing", Upper Saddle River: Prentice-Hall, 2002.

[3] D. Kundur, "Implications for high capacity data hiding in the presence of lossy compression", Proceeding of International Conference on Information Technology: Coding and Computing, March, 2000, pp. 16-21.

[4] T. Morkel, J.H.P. Eloff, and M.S. Olivier, "An Overview of Image Steganography," Proceedings of the Fifth Annual Information Security South Africa Conference. (ISSA2005), Sandton, South Africa, June/July 2005.

[5] M. Yang, S. Li, and N. Bourbakis, "Data-Image-Video Encryption", IEEE Potentials Magazine, Aug/Sept. 2004, pp.28-34.

[6] M. Yang, and N. Bourbakis, "A High Bitrate Multimedia Information Hiding Algorithm in DCT Domain", Proceeding of World Conference of Integrated Design and Process Technology (IDPT 2005), Beijing, China, June 13th-17th, 2005.

[7] "Enigmail Quickstart Guide", retrieved from http://enigmail.mozdev.org documentation/quickstart.php.html, March 23, 2011

[8] "GnuPG", retrieved from http://www.gnupg.org/documentation/index.en.html, March 23, 2011

[9] "OpenPGP Message Format", RFC 4880, retrieved from http://tools.ietf.org/html/rfc4880, March 23, 2011

[10] "Schneier on Security", Bruce Schneier, retrieved from http://www.schneier.com/blog/archives/2009/07/another_new_aes.html, March 23, 2011

[11] "Mozilla Thunderbird", retrieved from http://en.wikipedia.org/wiki/Mozilla_Thunderbird#Cross-platform_support, March 23, 2011