An Enhanced Cryptography for ECG Steganography to Satisfy HIPAA Privacy and Security Regulation for Bio-Medical Datas

S. P. PREDEEP KUMAR[1] and E. BABU RAJ[2]

[1]Research Scholar, St. Peter's University, Chennai, Tamilnadu, India.
[2]Principal, N S College of Engineering, Kanyakumari Dist., Tamilnadu, India.
*Corresponding author E-mail: sppredeep@gmail.com

http://dx.doi.org/10.13005/bpj/1053

(Received: April 10, 2016; accepted: May 13, 2016)

ABSTRACT

The Health Insurance Portability and Accountability Act(HIPAA) privacy and security regulations are two crucial provisions in the protection of healthcare privacy. According to HIPAA patient information sent through the internet should be protected and secured. In this paper, to comply with the HIPAA regulations, chaos cryptographic technique is used to encrypt the confident data into unreadable form. After the data decryption process, the encrypted data concealment in ECG signals.A wavelet basedsteganography technique is used to merge the encrypted patient confident data with the host ECG signal. To examine the performance of the proposed technique, Percentage Residual Difference (PRD) a distortion measurement technique is used. The proposed technique brings formidable security protection for bio-medical data with low distortion (less than 1%) and ECG signal remains diagnosable after watermarking and as well as after watermarks are detach from the watermarked signal.

Keywords: ECG, Steganography, Encryption, Wavelet, Watermarking, Confidentiality.

INTRODUCTION

The number of elderly patients is increasing dramatically due to the recent medical advancements. Accordingly, to reduce the medical labor cost, the use of remote healthcare monitoring systems and Point-of-Care (PoC) technologies have become popular[1,2]. Monitoring patients at their home can drastically reduce the increasing traffic at hospitals and medical centres. Moreover, Point-of-Care solutions can provide more reliability in emergency services as patient medical information (ex. for diagnosis) can be sent immediately to doctors and response or appropriate action can be taken without delay. However, Remote health care systems are used in large geographical areas essentially for monitoring purposes, and, the Internet represents the main communication channel used to exchange information. Typically, patient biological signals and other physiological readings are collected using body sensors.

Next, the collected signals are sent to the patient PDA device for further processing or diagnoses. Finally, the signals and patient confidential information as well as diagnoses report or any urgent alerts are sent to the central hospital servers via the Internet. Doctors can check those biomedical signals and possibly make a decision in case of an emergency from anywhere using any device[3]. Using Internet as main communication channel introduces new security and privacy threats as well as data integration issues. According to the Health Insurance Portability and Accountability Act (HIPAA), information sent through the Internet should be protected and secured. HIPAA mandates that while transmitting information through the internet a patient's privacy and confidentiality be protected as follows[4]:

Patient privacy

It is of crucial importance that a patient can control who will use his/her confidential health information, such as name, address, telephone number, and Medicare number. As a result, the security protocol should provide further control on who can access patient's data and who cannot.

Security

The methods of computer software should guarantee the security of the information inside the communication channels as well as the information stored onserver. Accordingly, it is of crucial importance to implement a security protocol which will have powerful communication and storage security[5]. Several researchers have proposed various security protocols to secure patient confidential information. Techniques used can be categorized into two subcategories. Firstly, there are techniques that are based on encryption and cryptographic algorithms. These techniques are used to secure data during the communication and storage. As a result, the final data will be stored in encrypted format[4,6,2,7]. The disadvantage of using encryption based techniques is its large computational overhead. Therefore, encryption based methods are not suitable in resource-constrained mobile environment. Alternatively, many security techniques are based on hiding its sensitive information inside another insensitive host data without incurring any increase in the host data size and huge computational overhead. These techniques are called steganography techniques. Steganography is the art of hiding secret information inside another type of data called host data[8]. However, steganography techniques alone will not solve the authentication problem and cannot give the patients the required ability to control who can access their personal information as stated by HIPAA. In this paper, a new security technique is proposed to guarantee secure transmission of patient confidential information combined with patient physiological readings from body sensors. The proposed technique is a hybrid between the two preceding categories. Firstly, it is based on using steganography techniques to hide patient confidential information inside patient biomedical signal. Moreover, the proposed technique uses encryption based model to allow only the authorized persons to extract the hidden data. In this paper, the patient ECG signal is used as the host signal that will carry the patient secret information as well as other readings from other sensors such as temperature, glucose, position, and blood pressure. The ElectroCardioGram (ECG)

Table 1: prd Results For Different Normal Ecg Segments

Sample Number	PRD Value
1	0.1925
2	0.2452
3	0.2367
4	0.1967
5	0.2762
6	0.2388
7	0.1937
8	0.1833
9	0.2876
10	0.2760

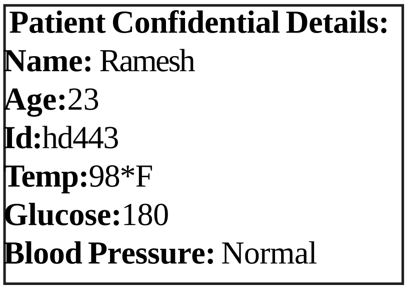
Table 2: iprd Results For Ventricular Tachycardia Ecg Samples

Sample Number	PRD Value
1	0.2567
2	0.2786
3	0.1765
4	0.2976
5	0.1287

Table 3: iiprd Results For Ventricular Fibrillation

Sample Number	PRD Value
1	0.2875
2	0.1634
3	0.2833
4	0.1774
5	0.2552

signal is used here due to the fact that most of the healthcare systems will collect ECG information. Moreover, the size of the ECG signal is large compared to the size of other information. Therefore, it will be suitable to be a host for other small size secret information. As a result, the proposed technique will follow HIPAA guidelines in providing open access for patients biomedical signal but prevents unauthorized access to patient confidential information. In this model body sensor nodes will be used to collect ECG signal, glucose reading, temperature, position and blood pressure, the sensors will send their readings to patient's PDA device via Bluetooth. Then , inside the patient's PDA device the steganography technique will be applied and patient secret information and physiological readings will be embedded inside the ECG host signal. Finally, the watermarked ECG signal is sent to the hospital server via the Internet. As a result, the real size of the transmitted data is the size of the ECG signal only without adding any overhead, because the other information are hidden inside the ECG signal without increasing its size. At hospital server the ECG signal and its hidden information will be stored. Any doctor can see the watermarked ECG signal and only authorized doctors and certain administrative personnel can extract the secret information and have access to the confidential patient information as well as other readings stored in the host ECG signal. The proposed steganography technique has been designed in such a way that guarantees minimum acceptable distortion in the ECG signal, Furthermore, it will provide the highest security that can be achieved. The use of this technique will slightly affect the quality of ECG signal. However, watermarked ECG signal can still be used for diagnoses purposes as it is proven in this paper. In this work the following research questions are answered:



Patient Confidential Details:
Name: Ramesh
Age:23
Id:hd443
Temp:98*F
Glucose:180
Blood Pressure: Normal

Fig. 1: Original data consisting of patient information

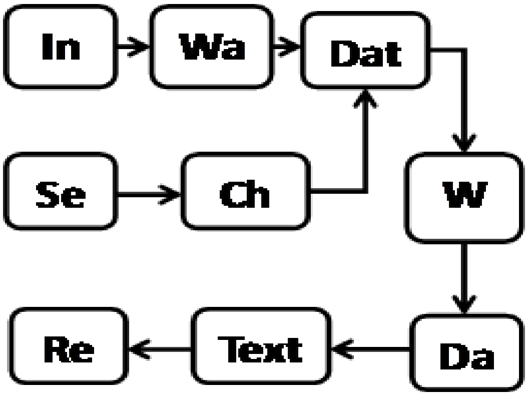


Fig. 2: Proposed Method System Architecture

- Can the proposed technique protect patient confidential data as explained in the HIPAA security and privacy guidelines?
- What will be the effect on the original ECG signal after applying the proposed steganography technique in terms of size and quality?

Rest of the paper is organized as follows. Section II briefly discusses the related works and what other researchers did in this area. In section III we discuss the basic system design, the embedding process (i.e patient sensitive data into ECG signal) and the extraction process. Next, in section IV security analysis is proposed. Then Section V explains diagnosability measurement. Section VI shows the results of PRD calculated before and after secret data extraction. Finally, section VII concludes the paper.

Table II and III shows 10 different cases taken and their corresponding average PRD values. It can be clearly seen how the values are approximately equal for different cases. The obtained results further prove that our proposed technique will cause minimum distortion for different cases of the scrambling matrix.

Related Work

There are many approaches to secure patient sensitive data[9,7,2,10]. However, one approach[11,12,13] proposed to secure data is based on using steganography techniques to hide secret information inside medical images. The challenging factors of these techniques are how much information can be stored, and to what extent the method is secure. Finally, what will be the resultant distortion on the original medical image or signal.

Kai-mei Zheng and Xu Qian[13] proposed a new reversible data hiding technique based on wavelet transform. Their method is based on applying B-spline wavelet transform on the original ECG signal to detect QRS complex. After detecting R waves, Haar lifting wavelet transform is applied again on the original ECG signal. Next, the non QRS high frequency wavelet coefficients are selected by comparing and applying index subscript mapping. Then, the selected coefficients are shifted one bit to the left and the watermark is embedded. Finally, the ECG signal is reconstructed by applying reverse haar lifting wavelet transform. Moreover, before they embed the watermark, Arnold transform is applied for watermark scrambling. This method has low capacity since it is shifting one bit. As a result only one bit can be stored for each ECG sample value. Furthermore, the security in this algorithm relies on the algorithm itself, it does not use a user defined key. Finally, this algorithm is based on normal ECG signal in which QRS complex can be detected. However, for abnormal signal in which QRS complex cannot be detected, the algorithm will not perform well.

H. Golpira and H. Danyali[12] proposed a reversible blind watermarking for medical images based on wavelet histogram shifting. In this work medical image such as MRI is used as host signal. A two dimensional wavelet transform is applied to the image. Then, the histogram of the high frequency subbands is determined. Next, two thresholds are selected, the first is in the beginning and the other is in the last portion of the histogram. For each threshold a zero point is created by shifting the left histogram part of the first threshold to the left, and shifting the right histogram part of the second threshold to the right. The locations of the thresholds and the zero points are used for inserting the binary watermark data. This algorithm performs well for MRI images but not for ECG host signals. Moreover, the capacity of this algorithms is low. Moreover, no encryption key is involved in its watermarking process.

Finally, S.Kauf and O.Farooq[11] proposed new digital watermarking of ECG data for secure wireless communication. In their work, each ECG sample is quantized using 10 bits, and is divided into segments. The segment size is equal to the chirp signal that they use. Therefore, for each ECG segment a modulated chirp signal is added. Patient ID is used in the modulation process of the chirp signal. Next, the modulated chirp signal is multiplied by a window dependent factor, and then added to the ECG signal. The resulting watermarked signal is 11 bits per sample. The final signal consists of 16 bits per sample, with 11 bits for watermarked ECG and 5 bits for the factor and patient ID.

METHODOLOGY

The sender side of the proposed steganography technique consists of four integrated stages. The proposed technique is designed to ensure secure information hiding with minimal distortion of the host signal. Moreover, this technique contains an authentication stage to prevent unauthorized users from extracting the hidden information.

Encryption

The aim of this stage is to encrypt the patient confidential information in such a way that prevents unauthorized persons - who does not have the shared key- from accessing patient confidential data. In this stage, a chaos cryptography technique is used. Chaos cryptography is selected because of its powerful encryption. Fig 1 shows an example of what information could be stored inside the ECG signal.

Wavelet Decomposition

Wavelet transform is a process that can decompose the given signal into coefficients representing frequency components of the signal at a given time. Wavelet transform can be defined as shown in equation C(S,P),

$$C(S,P) = \int_{-\infty}^{\infty} f(t)\,\psi(S,P)\,dt \qquad \text{...(1)}$$

Where represents wavelet function. S and P are positive integers representing transform parameters. C represents the coefficients which is a function of scale and position parameters. Wavelet transform is a powerful tool to combine time domain with frequency domain in one transform. In most applications discrete signals are used.

Therefore, Discrete Wavelet Transform (DWT) must be used instead of continuous wavelet transform. DWT decomposition can be performed by applying wavelet transform to the signal using band filters. The result of the band filtering operation will be two different signals, one will be related to the high frequency components and the other related to the low frequency components of the original signal. If this process is repeated multiple times, then it is called multi-level packet waveletdecomposition. Discrete Wavelet transform can be defined as in equation W(i,j),

$$W(i,j) = \sum_i \sum_j X(i)\,\psi ij(n)$$

...(2)

Where W(i,j) represents the DWT coefficients. i and j are the scale and shift transform parameters, and ij(n) is the wavelet basis time function with finite energy and fast decay. The wavelet function can be defined as in below equation,

$$\psi ij(n) = 2^{-\frac{i}{2}}\psi\left((2^{-i})n - j\right)$$

...(3)

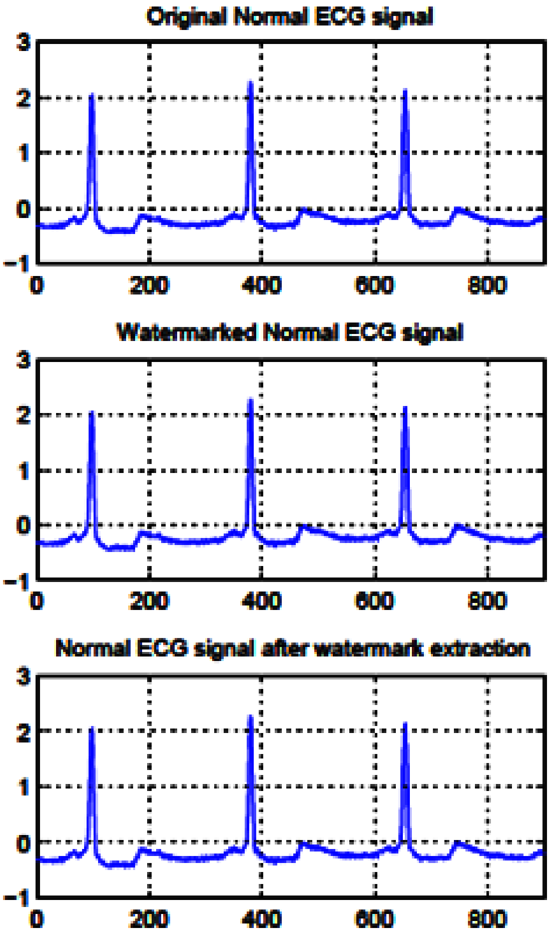


Fig. 3: ECG signals for before applying the steganography operation and after the steganography operation as well as after extracting the hidden data

In this paper, a 5-level wavelet packet decomposition has been applied to the host signal. Accordingly, 32 sub-bands resulted from this decomposition process. In each decomposition iteration the original signal is divided into two signals. Moreover, the frequency spectrum is distributed on these two signals. Therefore, one of the resulting signals will represent the high frequency component and the other one represents the low frequency component.

Most of the important features of the ECG signal are related to the low frequency signal. Therefore, this signal is called the approximation signal (A). On the other hand, the high frequency signal represents mostly the noise part of the ECG signal and is called detail signal (D). As a result, a small number of the 32 sub-bands will be highly correlated with the important ECG features while the other sub bands will be correlated with the noise components in the original ECG signal. Therefore, in our proposed technique different number of bits will be changed in each wavelet coefficient (usually called steganography level) based on its sub-band.

As a result, a different steganography level will be selected for each band in such a way that guarantees the minimal distortion of the important features for the host ECG signal. The process of steganography levels selection was performed by applying lot of experimentations.

It is clear that, hiding data in some sub-bands will highly affect the original signal, while hiding in other sub bands would result in small distortion effect. Accordingly, the selected steganography levels for bands from 1 to 17 are 5 bits and 6 bits for the other bands.

The Embedding Operation

At this stage the proposed technique will use a special security implementation to ensure high data security. In this technique a scrambling operation is performed using two parameters. First is the shared key known to both the sender and the receiver.

The embedding operation performs the data hiding process in the wavelet coefficients according to the sub-band sequence from the fetched row. The embedding process will start by reading the current wavelet coefficient in sub-band 32 and changing its LSB bits. Then, it will read the current wavelet coefficient in sub-band 22 and changing its LSB bits, and so on.

On the other hand, the steganography level is determined according to the level vector which contains the information about how many LSB bits will be changed for each sub-band. For example if the data is embedded in sub-band 32 then 6 bits will be changed per sample, while if it is embedded into wavelet coefficient in sub-band 1 then 5 LSB bits will be changed.

Inverse wavelet re-composition

In this final stage, the resultant watermarked 32 sub-bands are recomposed using inverse wavelet packet re-composition. The result of this operation is the new watermarked ECG signal. The inverse wavelet process will convert the signal to the time domain instead of combined time and frequency domain. Therefore, the newly reconstructed watermarked ECG signal will be very similar to the original unwatermarked ECG signal. The detailed is in embedding algorithm.

The algorithm starts by initializing the required variables. Next, the coefficient matrix will be shifted and scaled to ensure that all coefficients values are integers. Then, the algorithm will select a node out of 32 nodes in each row of the coefficient matrix. The selection process is based on the value read from the scrambling matrix and the key. The algorithm will be repeated until the end of the coefficient matrix is reached. Finally, the coefficient matrix will be shifted again and re-scaled to return its original range and inverse wavelet transform is applied to produce the watermarked ECG signal.

Watermark Extraction Process

To extract the secret bits from the watermarked ECG signal, the first step is to apply 5-level wavelet packet decompositionto generate the 32 sub-bands signals. Next, using the shared key and scrambling matrix the extraction operation starts extracting the secret bits in the correct order according to the sequence rows fetched from the scrambling matrix.

Finally, the extracted secret bits are decrypted using the same shared key. The watermark extraction process is almost similar to the watermarking embedding process. Except that instead of changing the bits of the selected node, it is required to read values of the bits in the selected nodes, and then resetting them to zero.

Security Analysis

The security of the proposed algorithm is mainly based on the idea of having several parameters shared between the transmitter and the receiver entities. Any change in any parameter will lead to extraction of wrong data. Accordingly,the receiver and transmitter should agree on the following information:

1. The scrambling matrix
2. The encryption key i.e. shared secret.
3. Steganography levels vector

As a result, even if the key is stolen the attacker will require knowing the scrambling matrix as well as the steganography levels vector. The scrambling matrix is stored inside the transmitter/ receiver pair and it will not be transmitted under any circumstance. For example, each patient could have his own device from his medical service provider and the matrix is burnt on this device. Therefore, for each pair of transmitter and receiver, it must be a unique scrambling matrix.

Diagnosability Measurement Of Thewatermarked Ecg Signal

To evaluate the proposed model, the PRD (Percentage Residual Difference) is used to measure the difference between the original ECG host signal and the resul ting watermarked ECG signal.

$$PRD = \sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N} x_i^2}}$$

Where x represents the original ECG signal, and y is the watermarked signal.

EXPERIMENTS AND RESULTS

In this paper three different types of ECG signals are used for experimentation. A test bed of 20 ECG samples is used for experimentation. The set of samples consist of 10 normal (NSR) ECG samples, 5 Ventricular fibrillation ECG samples and 5 Ventricular Tachycardia ECG samples. Each sample is 10 seconds long with 250 Hz sampling frequency.

Table II and III shows 10 different cases taken and their corresponding average PRD values. It can be clearly seen how the values are approximately equal for different cases. The obtained results further prove that our proposed technique will cause minimum distortion for different cases of the scrambling matrix.

CONCLUSION

In this paper a novel steganography algorithm is proposed to hide patient information as well as diagnostics information inside ECG signal. This technique will provide a secured communication and confidentiality in a Point-of-Care system. A 5-level wavelet decomposition is applied. A scrambling matrix is used to find the correct embedding sequence based on the user defined key. Steganography levels (i.e. number of bits to hide in the coefficients of each sub-band) are determined for each sub-band by experimental methods. In this paper we tested the diagnoses quality distortion. It is found that the resultant watermarked ECG can be used for diagnoses and the hidden data can be totally extracted.

REFERENCES

1. Y. Lin, I. Jan, P. Ko, Y. Chen, J. Wong, and G. Jan, “A wireless PDA-based physiological monitoring system for patient transport,” *IEEE Transactions on information technology in biomedicine*, vol. **8**, no. 4,pp. 439–447, (2004).

2. F. Hu, M. Jiang, M. Wagner, and D. Dong, "Privacy-preserving telecardiology sensor networks: toward a low-cost portable wireless hardware/ software codesign," *IEEE Transactions on Information Technology in Biomedicine,*, vol. **11**, no. 6, pp. 619–627, (2007).

3. A. Ibaida, I. Khalil, and F. Sufi, "Cardiac abnormalities detection from compressed ECG in wireless telemonitoring using principal components analysis (PCA)," in *5th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*. IEEE, 2010, pp. 207–212 (*2009*).

4. W. Lee and C. Lee, "A cryptographic key management solution for hipaa privacy/ security regulations," *IEEE Transactions on Information Technology in Biomedicine,*, vol. **12**, no. 1, pp. 34–41, (2008).

5. K. Malasri and L. Wang, "Addressing security in medical sensor networks," in *Proceedings of the 1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments*. ACM, p. 12 (2007).

6. I. Maglogiannis, L. Kazatzopoulos, K. Delakouridis, and S. Hadjiefthymiades, "Enabling location privacy and medical data encryption in patient telemonitoring systems," *IEEE Transactions on Information Technology in Biomedicine,*, vol. **13**, no. 6, pp. 946–954, (2009).

7 H. Wang, D. Peng, W. Wang, H. Sharif, H. Chen, and A. Khoynezhad, "Resource-aware secure ecg healthcare monitoring through body sensor networks," *Wireless Communications, IEEE*, vol. **17**, no. 1, pp. 12–19, (2010).

8 L. Marvel, C. Boncelet, and C. Retter, "Spread spectrum image steganography," *IEEE Transactions on Image Processing*, vol. **8**, no. 8, pp. 1075–1083, (1999).

9 A. De la Rosa Algarin, S. Demurjian, S. Berhe, and J. Pavlich-Mariscal, "A security framework for xml schemas and documents for healthcare," in *Bioinformatics and Biomedicine Workshops (BIBMW), 2012 IEEE International Conference on*, pp. 782–789 (2012).

10. M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, "Scalable and secure sharing of personal health records in cloud computing using attributebased encryption," *Parallel and Distributed Systems, IEEE Transactions on*, vol. **24**, no. 1, pp. 131–143, (2013).

11 S. Kaur, R. Singhal, O. Farooq, and B. Ahuja, "Digital Watermarking of ECG Data for Secure Wireless Commuication," in *2010 International Conference on Recent Trends in Information, Telecommunication and Computing*. IEEE, pp. 140–144 (2010).

12 H. Golpira and H. Danyali, "Reversible blind watermarking for medical images based on wavelet histogram shifting," in *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT),2009*. IEEE, pp. 31–36 (2010).

13 K. Zheng and X. Qian, "Reversible Data Hiding for Electrocardiogram signal Based on Wavelet Transforms," in *International Conference on Computational Intelligence and Security, 2008. CIS'08*, vol. **1**, (2008).

14 D. Stinson, *Cryptography: theory and practice*. CRC press, (2006).

15. A. Poularikas, *Transforms and Applications Handbook*. CRC, (2009).

16. A. Al-Fahoum, "Quality assessment of ECG compression techniques using a wavelet-based diagnostic measure," *IEEE Transactions on information Technology in Biomedicine*, vol. **10**, no. 1, (2006).