Compression and encryption for ECG biomedical signal in healthcare system

Mustafa Emad Hameed*[1], Masrullizam Mat Ibrahim[2], Nurulfajar Abd Manap[3]
[1,2,3]Centre for Telecommunication Research and Innovation (CeTRI), Faculty of Electronic and Computer Engineering (FKeKK), Universiti Teknikal Malaysia Melaka (UTeM), Malaysia
[1]Department of Computer Techniques Engineering, Bilad Al-rafidan University College, Diyala, Baqubah
*Corresponding author, e-mail: Mihhh221@gmail.com

Abstract

The ECG data needs large memory storage device due to continuous heart rate logs and vital parameter storage. Thus, efficient compression schemes are applied to it before sending it to the telemedicine center for monitoring and analysis. Proper compression mechanisms can not only improve the storage efficiency but also help in faster porting of data from one device to another due to its compact size. Also, the collected ECG signals are processed through various filtering techniques to remove unnecessary noise and then compressed. In our scheme, we propose use of buffer blocks, which is quite novel in this field. Usage of highly efficient methods for peak detection, noise removal, compression and encryption enable seamless and secure transmission of ECG signal from sensor to the monitor. This work further makes use of AES 256 CBC mode, which is barely used in embedded devices, proves to be very strong and efficient in ciphering of the information. The PRD outcome of proposed work comes as 0.41% and CR as 0.35%, which is quite better than existing schemes. Experimental results prove the efficiency of proposed schemes on five distinct signal records from MIT-BIH arrhythmia datasets.

Keywords*: CBC mode, compression, electrocardiogram, encryption, security*

1. Introduction

The Heart related diseases are becoming much common all over the world due to several factors that is actively affecting stress level and physical condition of the individuals. In medical field, electrocardiogram (ECG) signal processing is researched numerous times in the past due to its significance in monitoring heart status [1]. ECG signals carry electrical waves of heart functions, which mainly consists of five kinds of peaks i.e. P, Q, R, S and T. QRS complex in this context is very significant to be recognized. It illustrates left and right ventricular depolarization and has normal duration of 80ms-100ms. Normally, a person has a heart rate of 80-100 beats per minute. If the heart rate goes beyond 100 beats per minute, the condition is termed as tachycardia and if the beats are less than 60 per minute, it is termed as bradycardia [2]. Advancements in medical science and portable instruments have drastically reduced sensor equipment size. Portable ECG sensing devices obtain the signals and send it directly to the monitor for further analysis. In conventional systems, a sensor keeps on sending each n every bit of signal upon sensing to the monitor without any buffer [3].

A few complexities arise here due to the improper transmission rate or low bandwidth in communication channel. Often, continuous flow of signal may lead to an increase in sheer volume of data [4, 5]. Therefore, the volume of data is significantly affecting to the improper transmission rate or low bandwidth in communication channel. Often, continuous flow of signal may lead to packet drop and significant loss of vital information. However, the compression with eliminating redundancy in data is needed to optimize storage space and reduce the time required for data transmission causes stability [6-8]. Additionally, the ECG biomedical signals contain sensitive private health information as well as details that serve to individually distinguish patients, hence must be encrypted prior to transmission across public media so as to prevent unauthorized access by adversaries from cyber-attacks [9, 10].

Thus, In the research [11] proposed schemes for compression-then-encryption of the MIT-BIH Arrhythmia ECG signals. Their scheme involves complexity sorting (Beat detection, 2D ECG array formation, Period Normalization, Dc Equalization, Complexity Sorting, Codec

Received July 2, 2018; Revised May 29, 2019; Accepted July 18, 2019

Quantization and JPEG2000 codec) and coupled chaotic map mutation. In their scheme, they assume wireless transmission (using Rayleigh fading wireless channel) of ECG through OFDM (Orthogonal Frequency Division Multiplexing) and enhance it to correct impair samples using MMF (Moving Median Filtering). They show storage space minimization through 2D compression mechanism and combined it with chaotic based on mutation scheme to randomize ECG vector for maintaining shield to data confidentiality thus prevent from eavesdropping. As per Mahsa Raeiatibanadkooki [12] have propose a scheme that can compress the data without any loss of important information and also apply cryptographic scheme to preserve confidentiality from unauthorized access. In their work, they use mobile computing devise to eliminate usage of computers. They perform preprocessing such as removal of gaussian and baseline noise, detect peaks, do heart rate analysis and compress the ECG signal. At the compression they apply 3 level wavelet transformation (db04) and use threshold mechanisms. Next, Huffman coding technique is used to compress and encrypt the signal. They get the compression rate of 97.72% which is quite decent for any compression scheme. Further, the ECG signals are transmitted over TCP/IP to telemedicine clinic for specialist' assessment.

The study presented in [13] states the significance of compression with respect to ECG signals. According the researchers, ECG illustrates an individual's heart electrical movements. It helps in monitoring and diagnosis of heart related disease. Remote monitoring applications such as telemedicine necessitates storage of a big amount of data for assessment and diagnosis. Wireless transmission even consumes more energy while transmitting uncompressed data. Thus, compression of data is much required to reduce storage space, improve transmission rate and bandwidth usage. They compared diverse lossless compression schemes w.r.t ECG records and compared time efficiency and compression rate of those schemes. Through their study, they concluded that minimum variance Huffman coding is the best option to compress ECG signals. They considered MIT-BIH arrhythmia dataset for their study and MATLAB tool for simulations. As per their results, almost half of the storage memory can be saved with Minimum variance Huffman code having computational complexity of NLog2N. In their scheme they shown the better utilization of bandwidth with simple buffer design. Hence, the compression mechanism quality in reconstructed ECG needs to be measured by most popular function such as PRD (Percentage Root-mean-square Difference). Though, there are many chances that quality of reconstruction and the results may vary during evaluation of scheme due to different levels of data compressions. In their work they demonstrated that if efficiency of ECG compression scheme is tested alone in terms of quality through PRD method then variable outcomes may occur. Thus, they propose the use of multiple methods (PRD1 and CR) in order to achieve more accurate and reliable results and better conclusions during simulations. They performed simulations on MIT-BIH ECG (Arrythmia) dataset with different compression levels to analyze its influence on performance using PRD1 and CR (Compression Ratio) [14].

The study shows that, most of the previous works concentrate more on quality of signal based on peak detection but lacks security aspect. Though few works are found with security model along with filtering and compression schemes, but those schemes are not so efficient in terms of lossless compression or noise filtering. We propose a new system models based on lossless scheme such as Huffman coding for compression so that there is no information loss upon reconstruction. Also, to avoid the chances of data tampering we make use symmetric key of ciphering based on the AES-CBC algorithm with 256-bit key size. Remaining sections of this paper are organized as, section 1 contains Introduction section, 2 contains research method of proposed mechanisms, section 3 contains experimental results and section 4 summarizes the paper as conclusion.

2. Research Method

The block diagram of the proposed ECG compression and encryption method has been represented in the Figure 1. At first stage, the ECG biomedical signal has been loaded into proposed system. The second stage is the remove nosing from the signal using Discrete Wavelet Transform (DWT) combined with a Thresholding. The third stage is the block creation of ECG signal. Next to fourth stage the compression method is applied by use Huffman coding on each block of ECG signal. After completing the previous stage, the encryption method has been applying based on AES-CBC algorithm in the fifth stage and then transmission block ECG signal into monitoring. Finally, after receiving ECG signal block the decryption and decompress processes applying respectively based on inversion AES-CBC and inversion Huffman coding then, the block aggregation to recuperate the original ECG biomedical signal.

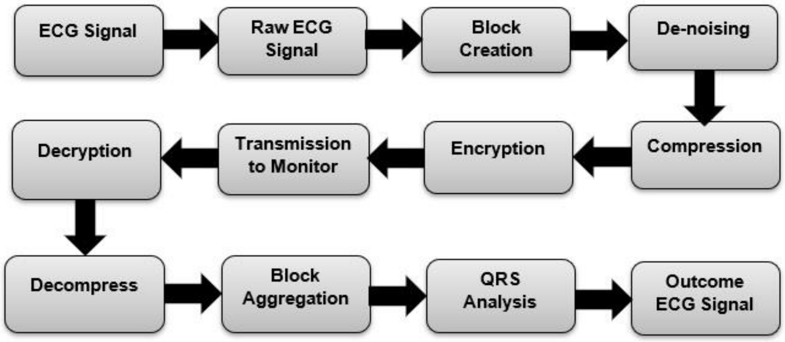


Figure 1. Proposed operations

2.1. DWT (Discrete Wavelet Transform)

Wavelets are arithmetic functions that work on signal data as per resolution or scale. DWT can be categorized as a type of wavelet that are discretely sampled the Figure 2 shown wavelet flow of operation. One of the key merits of DWT over fourier transforms is it analyses both frequency and time (location) in parallel. Soft or hard threshold methods define shrinkage rules [15].

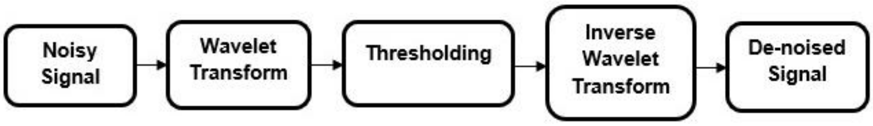


Figure 1. Wavelet flow of operations

Thresholding can be applied to the signal vector based on its characteristics. Majorly, there are two thresholding methods namely, soft and hard where hard is the easiest one. Hard thresholding:

$$\eta(x) = \begin{cases} x & |x| \geq T \\ 0 & |x| < T \end{cases} \tag{1}$$

soft thresholding:

$$\eta(x) = \begin{cases} x - T & x > T \\ 0 & |x| \leq T \\ x + T & x < -T \end{cases} \tag{2}$$

with wthresh, hard or soft threshold can be applied.

The daubechies (Db) wavelets are most popular discrete wavelet transforms in signal processing. It was developed in 1988 by Belgium based mathematician Ingrid Daubechies. Db has a range of wavelets, the first of which is the Db1 or Haar wavelet [16, 17].

2.2. Compression

Compression is the process of compacting size of the data through formulas or arithmetic operations [18]. Compression ratio is calculated to measure the performance of compression scheme and compactness of data. There are two types of compression schemes, i.e. lossy and lossless. Lossless schemes recover compressed file to its original status without any loss of data while decompressed. Particularly in the applications such as critical medical records, financial statement files and other vital files are always processed with lossless scheme as any loss of single bit also may affect aversively. For compression performance evaluation, various metrices functions are used such as PRD, CR and QS. The Compression Ratio (CR) is the measure of compression achieved in signal through encoding mechanisms. It doesn't provide information on compressed signal quality but measures efficiency of algorithm in reducing storage space. Thus, the Percentage Root-mean-square Difference (PRD) is a measure to evaluate error or difference between original. The quality score (QS) is used to evaluate the compression performance while considering the compromised reconstruction errors.

2.2.1. Huffman Code

Lossless compression techniques are very useful in applications where each bit of data is very significant for analysis such as medical field [19]. ECG signals are quite complex to analyze due to the property that a slight variation in signal value may misrecognize the peak type. In Huffman coding, input string characters are assigned with variable length codes (bit sequences), frequency of the individual character determines the allocated code length. Smallest code is assigned to character having the maximum frequency and largest code is assigned to the minimum frequency character. These variable length codes are known as prefix codes [20]. These prefix codes are unique for each character assignment. Thus, it is ensured that there should not be any ambiguity at the time of decoding operation on encoded bit stream. In Huffman, if there exists n unique characters to encode, then overall edge count (ec) would be:

$$ec = (2 * n - 2) \tag{3}$$

2.3. Cryptography

Cryptographic mechanism is a process in which information is converted to prevent it from being recognized by attackers [21]. The cryptography includes two processes first is the encryption process which entails converting the intelligible data into unintelligible data using a cryptography algorithm and encryption key. The second process is decryption, which involves converting the unintelligible data into intelligible data using the same algorithm and a decryption key [22]. The cryptography is divided into two broad categories; symmetric key and asymmetric key cryptography. The first category, symmetric key cryptography (otherwise called secret-key cryptography) uses the same key at the source and destination. The second category, asymmetric key cryptography uses different keys (called the public key) at the source and destination [23].

2.3.1. AES Algorithm

AES was one among the finalists in NIST competition and won the title of most secure cryptographic algorithm in October, 2000. It is also known as Rijndael and can have a variable key size of 128 bits, 192 bits or 256 bits with a fixed block size of 128 bits. It is a symmetric algorithm that uses single secret key for both encryption and decryption. There are four basic stages in each round of AES encryption or decryption. Permutation stage is ShiftRows and remaining three substitution stages are Substitute byte, MixColumns and AddRoundKey [24]. The encryption and decryption procedures of the Advanced Encryption Standard algorithm can show in Figure 3.

For 256 bits AES, Key Length (Nk) is 8 i.e. 8 words of 32-bits, Block Size (Nb) is 4 i.e. 4 words of 32-bits and no. of rounds (Nr) is 14. The ciphering function:

$$Enc(in[4 * Nb], out[4 * Nb]\,, w[Nb * (Nr + 1)]) \tag{4}$$

can be elaborated in following steps: Let st be the state and round be rd, AES works well with both hardware as well as software.There are five operational modes in AES i.e. ECB, CBC, CFB, OFB and CTR.

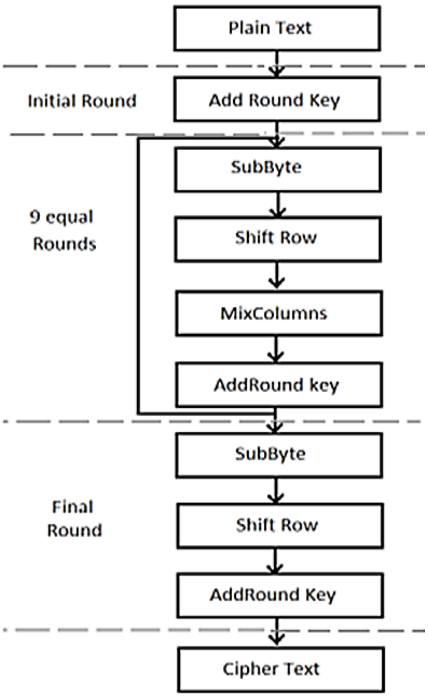


Figure 2. AES algorithm steps

2.3.2. Cipher Block Chaining Mode Operation (CBC)

To each ciphertext block produced earlier, an XOR is added to each plaintext block. Ciphering through CBC mode of AES can be illustrated as Figure 4 show the encryption and procedure of the CBC mode operation. The outcome of each succeeding ciphertext block relies on the preceding one. The initial plaintext block is added XOR to an unsystematic initialization vector (IV). An advancement to the ancestral block level ciphering version of AES such as ECB is Cipher Block Chaining (CBC). In this ciphering scheme, there is a dependency of every encrypted block on all plaintext blocks which are operated up to that stage. Due to this procedure, an additional level of computation is required while generating the cipher text [25].

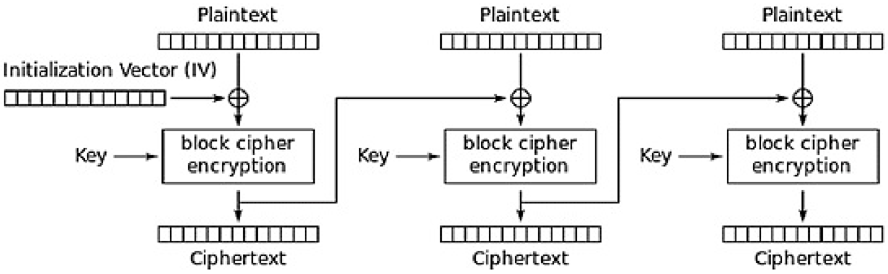


Figure 3. CBC mode operation

3. Results and Analysis

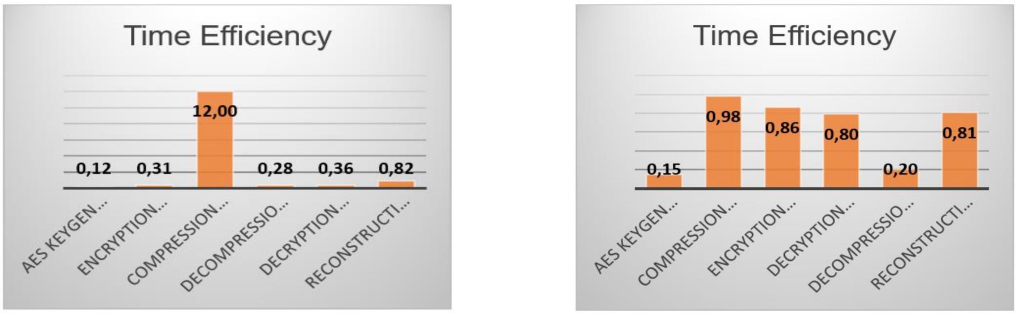
This portion reports and deliberates the results of the conducted investigations. In this work, we used DWT for clean signal with thresholding then compression signal used lossless method based on the Huffman coding algorithm and encryption the signal used symmetric key cryptography based on the AES-CBC algorithm block ciphering. The conducted experiments are performed on a computer with specifications (processor Intel(R) Core (TM) i5-4300U CPU@1.90GHz 2.49 GHz, RAM 4GB, under windows 10 professional 64-bit using MATLAB (R2018b). The execution of the proposed method was assessed using some parameters, for de-noise signal MSE, PSNR, SNR and CC the compression performance parameters are PRD and CR. In the end the timely execution and security level they are significant parameters for evaluation of proposed system performance. Therefore, the computational time of the system model can be defined as the time taken by each process in system. Performance parameter include the time taken by the algorithm for the compression and encryption of input ECG signal that is computational compression time and computational encryption time used for the processing on the ECG signal before transmutation to system monitoring.

Subsequently, after received the file calculated the estimated time taken by the algorithm for the decryption and decompression of file receiving that is computational decryption time and computational decompression time used for the processing on the ECG signal. As shown in Table 1 the time for different number of blocks from the whole signal until 60 blocks in the input of ECG signal file, which is the execution time for key generation, compression, encryption, decryption, decompression and total time of system proposed respectively. Therefore, for the four blocks of ECG signal the execution time are 3.2808, 3.8548, 3.5344, 0.616, in second respectively for each block. However, the increase for number of blocks on the ECG signal leads to computational overhead at the time of execution hence, the results show it that the proposed system within four blocks are consumes less time compared with other number of blocks.

Thus, the Figures 5 (a) and (b) shows the efficiency of the system in terms of the execution time for encryption mechanism then compression mechanism, the compression then encryption. As the experimental results show that the encryption process then the compression consumes more execution time in the process of implemented because of the entropy of ECG signal increased after encryption process. In the Figure 6 represents mean of CR and PRD retrieved on selected five datasets with 0.5db, 1dB and 2dB noise. As the results barely show any difference while changing dataset and noise ratio, thus average of it is demonstrated shown in Figure 6.

Table 1. Execution Time of Proposed System (in Second)

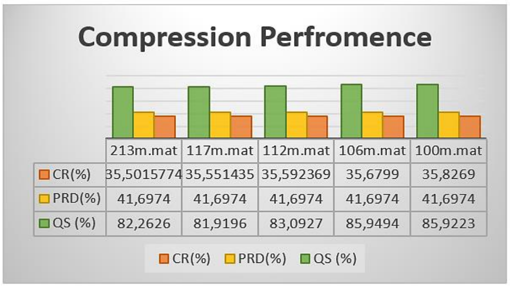
No. of block	AES Keygen Time	Compression Time	Encryption Time	Decryption Time	Decompression Time	Total Time of System Process
whole	0.1874	5.294	3.8862	3.8182	1.3216	14.325
4	0.1874	3.2808	3.8548	3.5344	0.616	11.4734
8	0.1874	3.8968	5.5792	2.9112	0.7944	13.369
16	0.1874	4.1648	6.6144	4.4832	0.8176	16.2674
24	0.1874	4.7176	9.1968	4.884	0.8408	19.8266
32	0.1874	4.8408	10.0832	5.36	0.8732	213446
40	0.1874	5.556	17.012	9.112	0.8812	32.7486
46	0.1874	6.2054	17.1386	9.6798	1.219	34.4302
60	0.1874	7.338	17.246	9.7558	1.4532	35.9804



(a) (b)

Figure 5. Time efficiency (a) encryption then compression (b) compression then encryption

The purpose of the ECG signal compression is to achieve high compression rates without changing the quality of the signal. The compression rate should be checked with the other parameters to evaluate the experimental results of the quality of the reconstructed signal. Table 2 comparison the performance of compression algorithm with other algorithms in previous studies. Figure 7 represents Denoising performance of HAAR wavelet on the noise ranges in between 0.5 dB to 6 dB. PSNR is relatively high in record no. 106 and 213 and lowest in record no. 117. Also, MSE is lowest in record no. 213 and highest in record no. 100. Thus, it can be clearly justified that why cross correlation is highest in record no. 213.

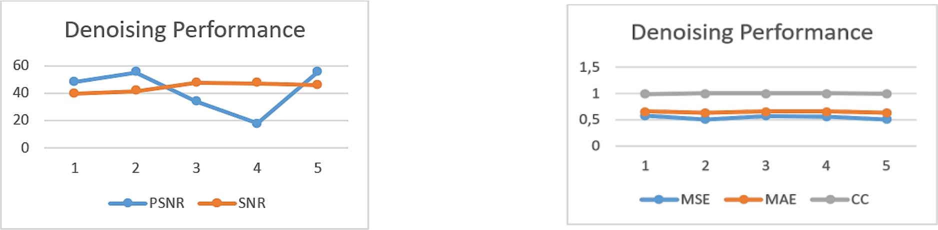


	213m.mat	117m.mat	112m.mat	106m.mat	100m.mat
🟧 CR(%)	35,5015774	35,551435	35,592369	35,6799	35,8269
🟨 PRD(%)	41,6974	41,6974	41,6974	41,6974	41,6974
🟩 QS (%)	82,2626	81,9196	83,0927	85,9494	85,9223

Figure 6. Compression performance

Table 2. Comparison between Proposed Method and Other Compression Alogorithms

		[26]	[27]	[28]	[12]	[29]	This Proposed Work
Comparison performance	PDR (%)	3.88	0.42	0.641	0.522	1.067 1.137 1.184	0.411
	CR (%)	0.40	0.401	16.91	0.9772	0.151	0.35015
	QS (%)	x	96.14	29.36	x	x	85.18



(a) (b)

Figure 7. Denoising Performance: (a) SNR and PSNR for MIT database, with different levels of noise, (b) MSE, MSA and CC for MIT database, with different levels of noise

4. Conclusion

This paper mainly aims to propose a new lightweight system model to process ECG signals efficiently and securely. To test the efficiency of proposed model, five distinguished datasets from MIT-BIH arrhythmia repository were processed through several mechanisms of Denoising, filtering, compression and encryption. The delay performance of compression algorithms is particularly important when time critical data transmission is required. As compressed signals take lesser time in computations compared to raw signal, thus Huffman lossless scheme is employed. Efficiency of those mechanisms are computed in terms of PRD, CR, PSNR, MSE, etc. The PRD outcome of proposed work comes as 0.41% and CR as 0.35%, which is quite better than existing schemes. The Experimental results prove the efficacy of algorithms used. Thus, the block level processing and encryption of signal using AES-CBC algorithm with 256-bit key size which could be prepare high level of security and quite novel in this work for further tested on real time embedded device.

References

[1] SP Awasarmol, S Ashtekar, A Chintawar. *Securely data hiding and transmission in an ECG signal using DWT*. 2017 Int. Conf. Energy, Commun. Data Anal. Soft Comput. ICECDS 2017. 2017: 2850–2854.

[2] JS Sahambi, SN Tandon, RKP Bhatt. Using wavelet transforms for ECG characterization. An on-line digital signal processing system. *IEEE Eng. Med. Biol. Mag.* 1997; 16(1): 77–83.

[3] MV Ramesh, GR Ragi, TK Abishek. *Low Power Intelligent Wearable Cardiac Sensor Using Discrete Wavelet Compression.* 2012 Int. Conf. Adv. Mob. Network, Commun. Its Appl. 2012: 107–110.

[4] H Kupwade Patil, R Seshadri. *Big Data Security and Privacy Issues in Healthcare 2014.* IEEE Int. Congr. Big Data. 2014: 762–765.

[5] J Andreu-Perez, CCY Poon, RD Merrifield, STC Wong, G-Z Yang. Big Data for Health. *IEEE J. Biomed. Heal. Informatics*. 2015; 19(4): 1193–1208 .

[6] SK Mukhopadhyay, S Mitra, M Mitra. An ECG signal compression technique using ASCII character encoding. *Measurement*. 2012; 45(6): 1651–1660.

[7] D Gurve, BS Saini, I Saini. An improved lossy and lossless combined ECG data compression using ASCII character encoding. *Int. J. Med. Eng. Inform.* 2016; 8(4): 758–764.

[8] AF Hussein, SJ Hashim, AFA Aziz, FZ Rokhani, WAW Adnan. A real time ECG data compression scheme for enhanced bluetooth low energy ECG system power consumption. *J. Ambient Intell. Humaniz. Comput.* 2017: 1-14.

[9] H Al-Hamadi, A Gawanmeh, J Baek, M Al-Qutayri. Lightweight Security Protocol for ECG Bio-Sensors. *Wirel. Pers. Commun.* 2017; 95(4): 5097–5120.

[10] A Vaniprabha, P Poongodi. Augmented lightweight security scheme with access control model for wireless medical sensor networks. *Cluster Comput.* 2018: 1-11.

[11] A Pandey, BS Saini, B Singh, N Sood. Complexity sorting and coupled chaotic map based on 2D ECG data compression-then-encryption and its OFDM transmission with impair sample correction. *Multimedia Tools and Applications.* 2018; 78(9): 11223-11261.

[12] M Raeiatibanadkooki, SR Quchani. Compression and Encryption of ECG Signal Using Wavelet and Chaotically Huffman Code in Telemedicine Application. *J Med Syst.* 2016; 40(3): 73.

[13] R Tornekar, S Gajre. *Comparative Study of Lossless ECG Signal Compression Techniques for Wireless Networks.* 2017 Comput. Cardiol. Conf. 2018; 44: 1–4.

[14] M Blanco-Velasco, F Cruz-Roldán, JI Godino-Llorente, J Blanco-Velasco, C Armiens-Aparicio, F López-Ferreras. On the use of PRD and CR parameters for ECG compression. *Med. Eng. Phys.* 2005; 27(9): 798–802.

[15] A Graps. An Introduction to Wavelets. *IEEE Comput. Soc.* 1995; 2: 1–18.

[16] AN Akansu, RA Haddad, H Caglar. Perfect Reconstruction Binomial Qmf-Wavelet Transform. *SPIE Vis. Commun. Image Process.* 1991; 1360: 609–618.

[17] AN Akansu, RA Haddad. Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets. 2001.

[18] V Kumar, SC Saxena, VK Giri. Direct data compression of ECG signal for telemedicine. *Int. J. Syst. Sci.* 2006; 37(1): 45–63.

[19] DA Huffmant. A Method for the Construction of Minimum-Redundancy Codes. *Proceeding of the I.R.E* 1952; 27: 1098–1101.

[20] D Richards. The Length of a Typical Huffman Codeword. *IEEE Trans. INFORMATION THEORY*. 1994; 40(4): 1246–1247.

[21] CF Lin, SH Shih, J De Zhu. Chaos based encryption system for encrypting electroencephalogram signals. *J. Med. Syst.* 2014; 38(5): 1–10.

[22] ME Hameed, MM Ibrahim, NA Manap. Review on Improvement of Advanced Encryption Standard (AES) Algorithm based on Time Execution, Differential Cryptanalysis and Level of Security. *J. Telecommun. Electron. Comput. Eng.* 2018; 10(1): 139–145.

[23] MS Reddy, YA Babu. Evaluation Of Microblaze and Implementation Of AES Algorithm using Spartan-3E. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.* 2013; 2(7): 3341–3347.

[24] Z Xinmiao, KK Parhi. Implementation approaches for the advanced encryption standard algorithm. *IEEE Circuits Syst. Mag.* 2002; 2(4): 24–46.

[25] M Vaidehi, BJ Rabi. *Design and analysis of AES-CBC mode for high security applications.* 2nd Int. Conf. Curr. Trends Eng. Technol. ICCTET 2014. 2014: 499–502.

[26] M Fira. *Applications of compressed sensing: Compression and encryption*. in *2015 E-Health and Bioengineering Conference (EHB)*. 2015: 1–4.

[27] TY Liu, KJ Lin, HC Wu. ECG data encryption then compression using singular value decomposition. *IEEE J. Biomed. Heal. Informatics.* 2018; 22(3): 707–713.

[28] H Kim, S Member, RF Yazicioglu, P Merken, C Van Hoof, H Yoo. ECG Signal Compression and Classification Algorithm With Quad Level Vector for ECG Holter System. *IEEE Transactions on Information Technology in Biomedicine* 2010; 14(1): 93–100.

[29] BS Kim, SK Yoo, MH Lee. Wavelet-Based Low-Delay ECG Compression Algorithm for Continuous ECG Transmission. *IEEE Transactions on Information Technology in Biomedicine.* 2006; 10(1): 77–83