Chapter 3

Wavelets in ECG Security Application

Seedahmed S. Mahmoud and Jusak Jusak

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.74477

Abstract

Wavelet packet transform has been used in many applications of biomedical signal processing, for example, feature extraction, noise reduction, data compression, electrocardiogram (ECG) anonymisation and QRS detection. The wavelet analysis methods, in these applications, represent the temporal characteristics of a biological signal by its spectral components in the frequency domain. Furthermore, it has been shown in many works that the ECG signal can be used as a biometric method for robust human identification and authentication. In this case, it is necessary to anonymise the ECG data during the distribution and storage of the signal in a public repository. A neglectful system leads to an eavesdropper recording the ECG data and uses it as recognition data to gain access via an ECG biometric system. This chapter discusses and reviews recent researches on ECG anonymisation wavelets-based techniques. These techniques use discrete wavelet transform and wavelet packet transform. A comparative study between the wavelets-based methods will be presented.

Keywords: anonymisation, biometric, electrocardiogram, encryption, steganography, wavelet analysis

1. Introduction

In June 2006, Cisco released a virtual network index (VNI) forecast that projects global IP traffic over the next 5 years [1]. According to Cisco's paper, there has been quantitative evidence that proliferation of global IP traffic will exchange data to reach the order of zettabyte (ZB) by 2021. This massive amount of data will be driven mainly by the number of connected devices to IP networks, such as smart phones, tablets, sensors and machine-to-machine (M2M) applications that are estimated to be more than three times the global population. Hence, in this era, just about every physical object we see (e.g. health-care monitoring apparatus, machinery, appliances, autonomous cars and intelligent transportation, etc.) will be connected,

forming the Internet of Things (IoT) [2]. In order to handle the countless number and various types of devices as well as linking the existing radio-access technologies, a new architecture that will increase data rate, lower end-to-end latency and improve the coverage is urgently required. Therefore, to meet with this demand, a new standard on the fifth-generation (5G) networks is currently under consideration [3].

Health and medical care are considered as one of the most fascinating applications that can fully benefit from IoT deployment. The IoT that employs various sensor and smart medical devices may serve in, for example, tele-auscultation, remote health monitoring, remote diagnostics and possibly treatment as well as elderly care [4–6]. Such Internet of Medical Things (IMedT) is expected to reduce consultation and transportation cost and to shrink the gap for those who live in the isolated/remote areas where the presence of doctors is void. Nevertheless, transmitting medical data to health-care providers through the public networks require high data security as public networks are somehow vulnerable to spoof attack. In this chapter, two anonymisation techniques based on wavelet decomposition and wavelet packet (WP) transform for securing ECG signals will be discussed.

2. Motivations

An electrocardiogram (ECG) signal contains important health information of a patient. It is used to detect abnormal heart rhythms by measuring the electrical activity generated by the heart as it contracts. Recent studies show that an ECG signal can be used as a biometric method for robust human identification and authentication [7–9]. The ECG signal was found to be unique for each individual over a long period of time [10, 11]. An ECG biometric system consists of feature extraction and classifiers to identify and recognise a person. The selection of appropriate features is crucial for successful individual identification. In [12], ECG-based biometric features were grouped as fiducial based, non-fiducial based or hybrid.

An unsecure ECG signal can be subjected to *man in the middle* attack where fraudsters can use the spoofed recorded ECG data to gain access to a secured service [13–15]. A scenario where a man in the middle attack can be a real threat for health information transmission is presented in **Figure 1**. The figure illustrates possible attack points that include (1) wireless links between sensor nodes that collect health information data from wireless body area networks (WBAN) and gateways, (2) wire/wireless links between the gateway and the edge router, (3) wire/wireless links between the other side of the edge router and health-care provider router and (4) repository in the data centre/public server or health-care provider. In order to minimise such security threat to a system, a health-care provider needs to comply with certain widely accepted standards to protect medical records safely. For example, US government passed the Health Insurance Portability and Accountability Act (HIPAA) in 1996 for protecting medical privacy users [16], the European Union adopted the Directive on Data Protection in 1995 [17], the Health Information Privacy Code was passed by New Zealand government in 1994, which sets specific rules for agencies in the health sector to ensure protection of individual privacy [18] and the personally controlled electronic health record (PCEHR) eHealth system was launched by Australian government in 2012 [19].

http://dx.doi.org/10.5772/intechopen.74477

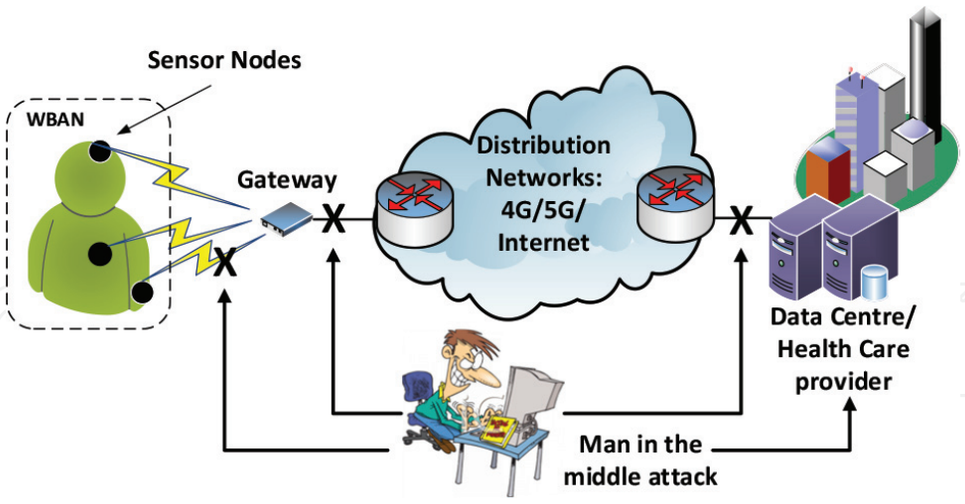

Figure 1. Possible attack points for unsecure ECG signals subjected to man in the middle attack.

There have been several proposed security techniques including image [20] and ECG steganography [21–24] to secure confidential patient information. In the steganography techniques, sensitive patient information is concealed inside public host data without incurring huge computational overhead or any increase in the size of the host data [21]. ECG data is used as the host signal to embed secret patient information and physiological readings. This may create watermarked ECG signals that is then transferred to a remote hospital server for further diagnosis. The effectiveness of ECG watermarking is dependent on the difference between the original host data and the watermarked data, that is, greater differences point to an ineffective steganography process. Unfortunately, all steganography methods bear some degree of information loss. This severe loss of information contributes to smeared/incorrect signal features and in some cases can lead to the failure of reconstructing the original ECG signal from the watermarked ECG signal [22]. However, even effective ECG watermarking can result in the delectability of ECG fiducial and non-fiducial features, which may allow for patient identification according to research in [7–9]. Therefore, a method combining the advantages of steganography with a technique that hides ECG fiducial and non-fiducial features is required. In this chapter, a review between two ECG anonymisation methods based on wavelet decomposition and wavelet packet transform (WPT) is presented.

3. Wavelet decomposition-based ECG anonymisation approach

Recent ECG anonymisation approaches based on wavelet decomposition were proposed in [13, 14]. During the wavelet decomposition process, filters of different cut-off frequencies were used to analyse the ECG signal at different scales (frequencies). It can be done by passing the ECG signal through a series of high-pass filters (i.e. the detail coefficients) for examining the high-frequency bands. The ECG signal was also passed through a series of low-pass filters (i.e. the approximation coefficients) to evaluate the low-frequency bands. Wavelet decomposition at level 3 was used during signal evaluation in the chapter [13]. Moreover, in the order to

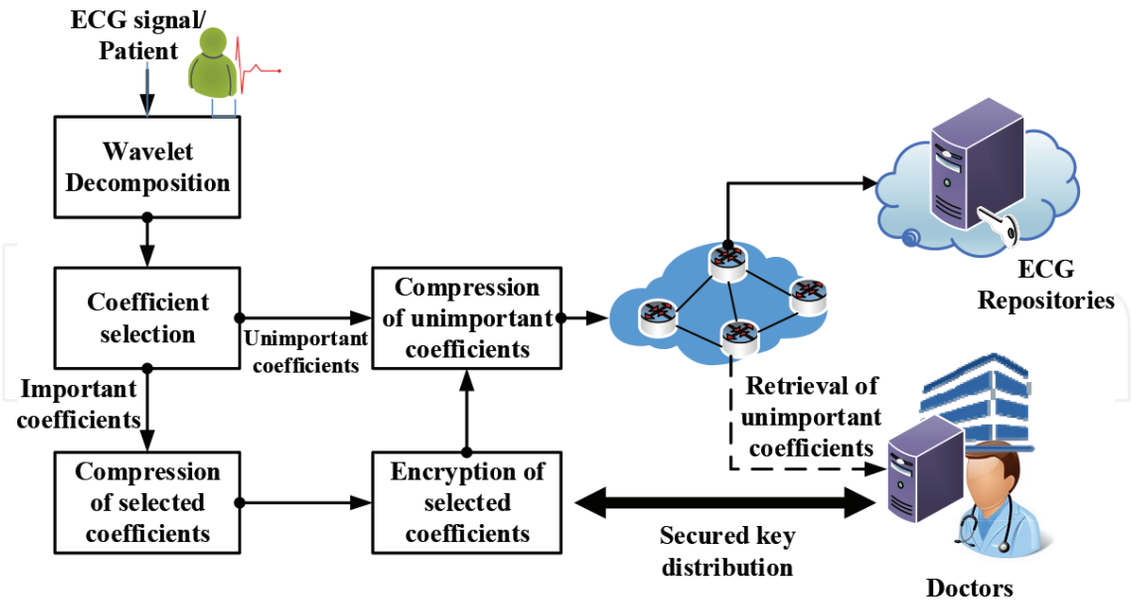

Figure 2. ECG anonymisation using wavelet decomposition.

construct a complete evaluation, two individual methods were studied during the experimentation [13]. Block diagram for the wavelet decomposition can be seen in **Figure 2**.

3.1. Method 1: discrete wavelet base anonymisation

In the first method, approximation (cA3) and detail (cD3) coefficients were removed after level 3 decomposition. Subsequently these nodes were encrypted using the well-known RSA symmetric cryptography. On the other hand, the remaining nodes, that is, cD1 and cD2, were compressed and transmitted to the ECG repository. **Figure 3** shows that without knowledge of nodes cA3 and cD3, the newly constructed signal in the repository completely hides P wave and T wave of the original ECG. It can be concluded that the first method hides most of the features required to reconcile the identity of a patient [7]. On the contrary, this method is not able to provide complete obfuscation of the cardiovascular conditions. This is mainly because

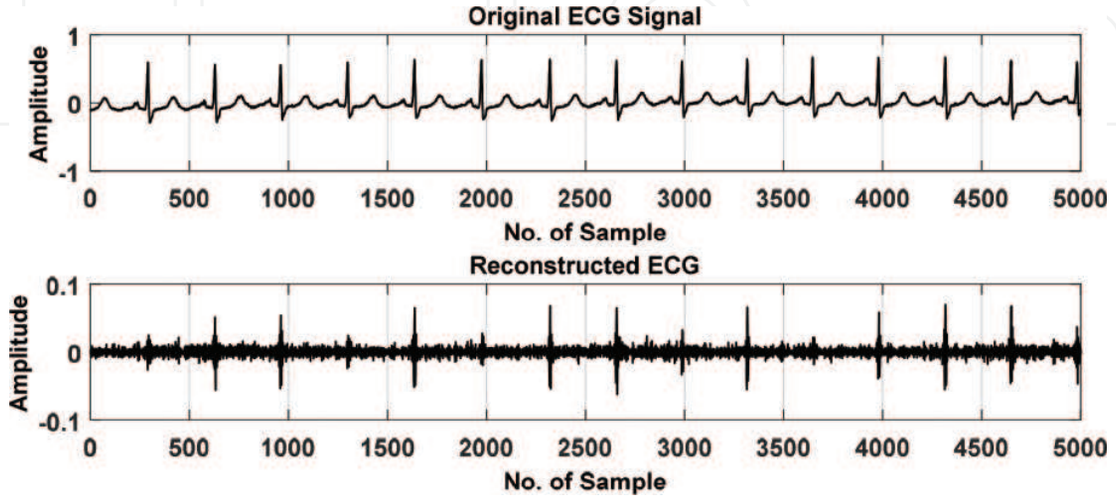

Figure 3. Normal ECG signal (top) and reconstructed annonymised ECG signal without nodes cA3 and cD3.

http://dx.doi.org/10.5772/intechopen.74477

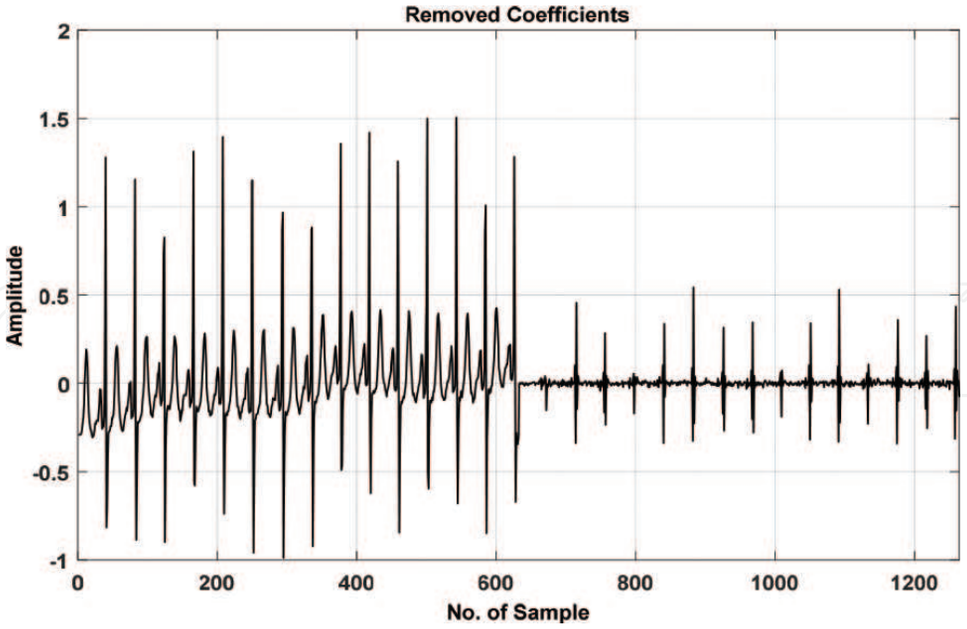

Figure 4. Removed (Selected for Encryption) Coefficients for Method 1.

the RR interval and certain types of arrhythmias are visible [25] as obvious in **Figure 3**. However, this method only required minimal selection of coefficient (approximately 25%) for encryption and key distribution. This is the main advantage of the first method. This method will perform well when faster distribution of key is priority and strong security is not deemed necessary. The removed coefficients are shown in **Figure 4**.

3.2. Method 2: discrete wavelet base anonymisation

In the second method, nodes cA3, cD3 and cD2 were selected for encryption, while the remaining coefficients cD1 were transmitted to the ECG repository. In contrast to the previous method, **Figure 5** shows that the reconstructed ECG from the coefficients that are extracted

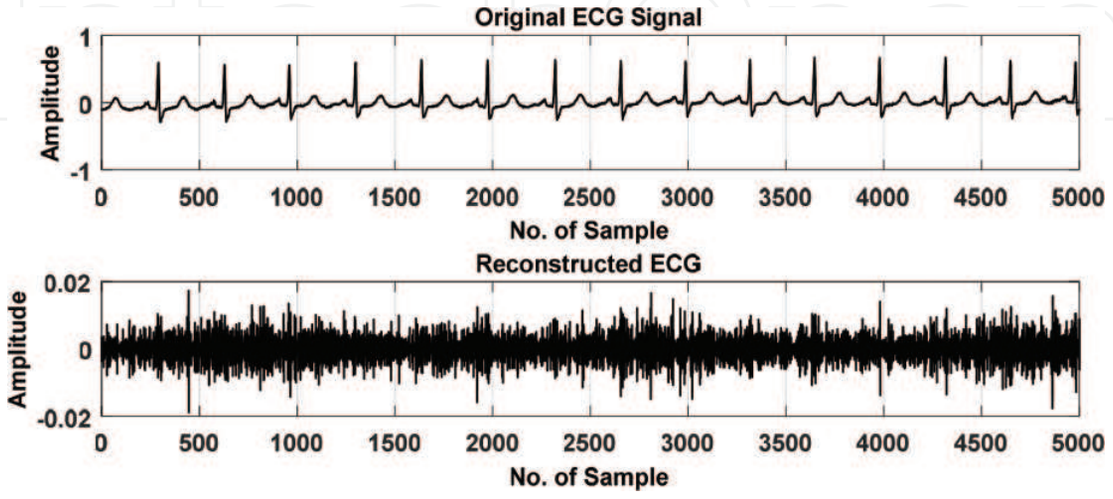

Figure 5. Normal ECG signal (top) and reconstructed annonymised ECG signal without nodes cA3, cD3 and cD2.

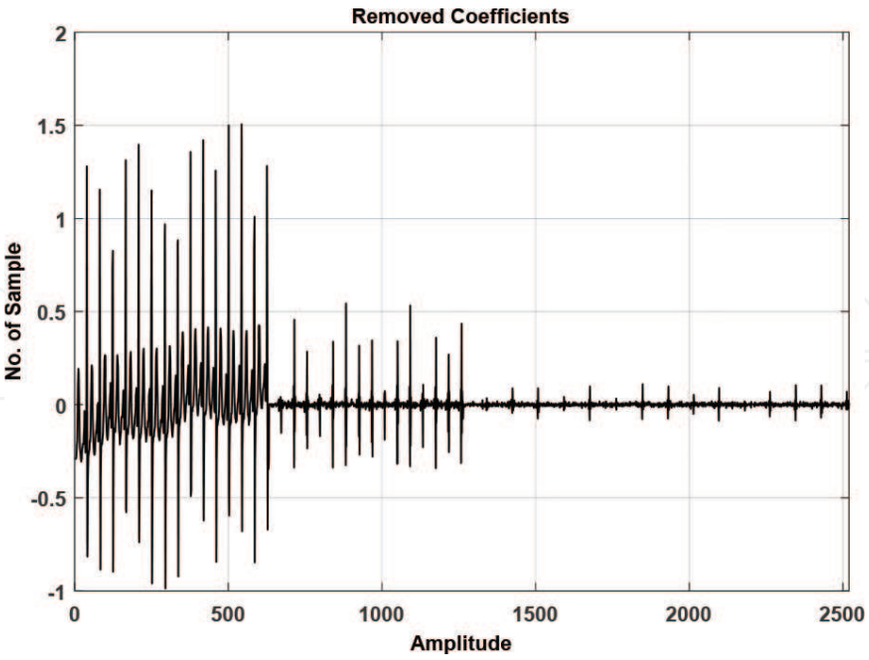

Figure 6. Removed (selected for encryption) coefficients for method 2.

from the repository is completely able to obfuscate features related to cardiovascular condition and person identification.

Therefore, this method provides higher ECG security by compromising larger key size (approximately 50%) as can be seen in **Figure 6**. **Figure 5** shows that the reconstructed ECG signal does not contain any ECG features.

Both methods described above suffer from long key size and lack of complete obfuscation to the ECG data. The long key size requires wider bandwidth during transmission process of the key to the ECG repository. On the other hand, lack of complete obfuscation results in trivial interpretation of the anonymised ECG signal. Therefore, due to these two main reasons, other methods based on the wavelet packet were proposed and developed.

4. Wavelet packet-based ECG anonymisation approach

4.1. Overview of wavelet packet transform

Wavelet packet transform has been used in many applications of biomedical signal processing, for example, feature extraction, noise reduction, data compression and QRS detection. Furthermore, wavelet packet transform has long been used for ECG signal analysis. A wavelet packet function [18] is defined as

$$\varphi_{j,k}^{n}(t) = 2^{\frac{1}{2}}\varphi^{n}\left(2^{j}t - m\right), \tag{1}$$

where j and m are the scale (frequency) and the translation (time) parameters, respectively, and $n = 0, 1, 3, \ldots$ is the oscillation parameter. The structure of wavelet packet (WP) decomposition

http://dx.doi.org/10.5772/intechopen.74477

is described as a binary tree structure E; each node is described as (j, n), where j is a node's scale level and n is a node's number on the corresponded level. The root node $(0, 0)$ of the WP tree corresponds to the entire frequency range, $\left[0, \frac{f_s}{2}\right]$, where f_s is the ECG sampling frequency of the ECG signal. Each internal node of the WP tree $(j, n) \in E$ is called a parent node that is divided into two child nodes: the first and the second nodes are associated with low-pass $h(m)$ and high-pass $g(m)$ filters. These nodes forms a quadrature mirror filter (QMF) pair [19].

The scaling function $\omega(t)$ and the mother wavelet $\varphi(t)$ for the wavelet packet when $n = 0, 1$ and $j = m = 0$ are given by

$$\varphi^0(t) = \omega(t), \varphi^1(t) = \psi(t). \tag{2}$$

The other wavelet packet functions for $n = 2, 3, \ldots$ and $j = 1$ are shown as follows:

$$\varphi^{2n}(t) = \sum_m h(m)\varphi^n_{j,m}(t), \tag{3}$$

$$\varphi^{2n+1}(t) = \sum_m g(m)\varphi^n_{j,m}(t). \tag{4}$$

By substituting Eq. (1) into Eq. (3) and (4), we can get

$$\varphi^{2n}(t) = \sqrt{2}\sum_m h(m)\varphi^n(2t - m), \tag{5}$$

$$\varphi^{2n+1}(t) = \sqrt{2}\sum_m g(m)\varphi^n(2t - m), \tag{6}$$

where the low-pass filter gives $h(m) = \frac{1}{\sqrt{2}}\langle\omega(t), \omega(2t - m)\rangle$, and the high-pass filter gives $g(m) = \frac{1}{\sqrt{2}}\langle\psi(t), \psi(2t - m)\rangle$. The operator $\langle ., . \rangle$ stands for the inner product. The wavelet packet coefficients of the ECG signal, $x(t)$, are expressed as follows:

$$Q^n_j(m) = \left\langle x, \psi^n_{j,m} \right\rangle = \int_{-\infty}^{\infty} x(t)\psi^n_{j,m}(t)\, dt \tag{7}$$

Each coefficient measures a specific sub-band frequency content, controlled by the scaling parameter, j, and the oscillation parameter, n. The ECG signal, $x(t)$, can be decomposed into a different time-frequency space with Eq. (6) and Eq. (7). By computing the full wavelet packet decomposition on the ECG signal, for the jth level of decomposition, we have 2^j sets of sub-band coefficients of length $\frac{N}{2^j}$, where N is the ECG signal length [20]. Each sub-band coefficient, node, has a frequency range in the interval $\left[\frac{n}{2^{j+1}}, \frac{n+1}{2^{j+1}}\right]$, $n = 0, 1, \ldots, 2^j - 1$. This is how wavelet packet decomposes the original ECG signal into two or more coefficients.

4.2. The generalised framework for the ECG anonymisation method

In this section, a generalised framework for the ECG anonymisation using wavelet packet transform (WPT) will be introduced. The proposed framework for ECG anonymisation can be seen in **Figure 7**, while its pseudo-code is listed in Algorithm 1. This framework comprises the following steps:

Step 1: Perform wavelet packet decomposition of the ECG signal, $x(t)$, at level j. The signal coefficients at this level are given by

$$C = \{c(j,n): n = 0, 1, \ldots, 2^j - 1\} \tag{8}$$

where $c(j,n)$ represents the coefficients of the nth node at level j.

Step 2: Exclude the first node, $c(j,0)$, from C in Eq. (8) to get

$$\overline{C} = \{c(j,n): n = 1, 2, \ldots, 2^j - 1\}. \tag{9}$$

The excluded node is set to

$$k = c(j,0) \tag{10}$$

where k is an unencrypted and uncompressed key that includes the low-frequency components of the ECG signal, $x(t)$.

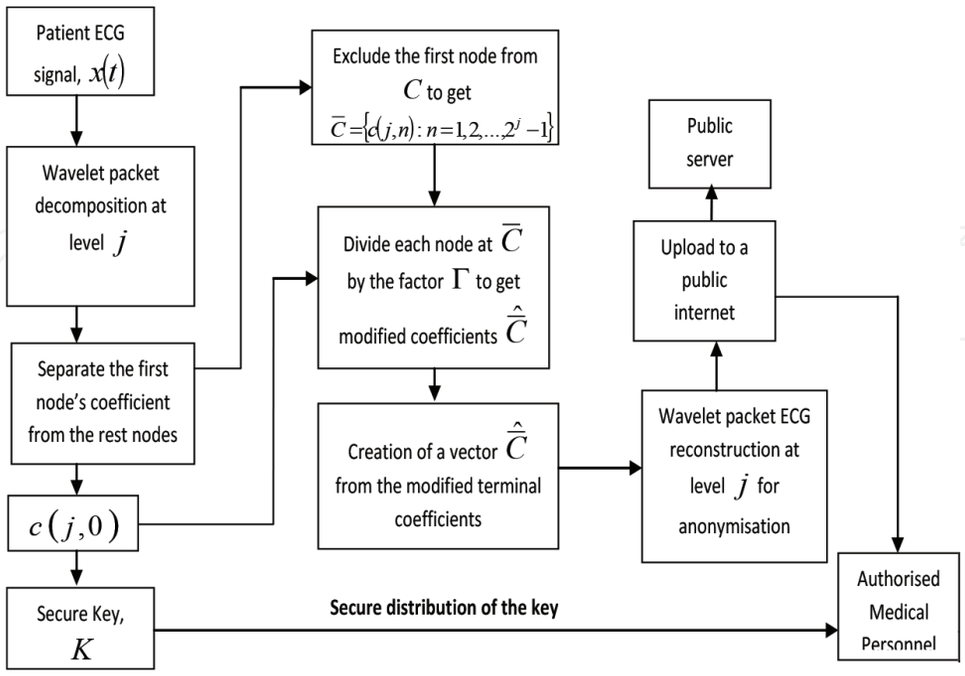

Figure 7. Wavelet packet-based ECG anonymisation process.

Step 3: Modify each node in $\overline{C}$, Eq. (9), using a reversible function/operation such as logarithm or division. In this chapter each node in $\overline{C}$ is divided by Γ. Γ is a reversible function driven from the key coefficients. Hence, the modified coefficients in $\overline{C}$ are given by

$$\widehat{\overline{C}} = \frac{\overline{C}}{\Gamma} = \left\{ \frac{c(j,n)}{\Gamma} : n = 1, 2, \ldots, 2^j - 1 \right\}, \tag{11}$$

where $\Gamma = k + \text{offset}$, $\text{offset} = |\min(k)| + \eta$, η is a constant and $|.|$ is the absolute operator. The offset term in Γ is used to prevent division by zero.

Step 4: Securely distribute the key, K, and the offset value to medical personnel. The key security will be achieved by compressing and encrypting the first node, k, and the offset as follows:

$$K = E(\Delta(k, \text{offset})) = E(\Delta(c(j,0), \text{offset})), \tag{12}$$

where $\Delta(.)$ is the compression operator and $E(.)$ is the encryption operator [9]. Compression and encryption are beyond the scope of this chapter.

Step 5: Perform wavelet packet reconstruction to the modified terminal nodes' coefficient, $\widehat{\overline{C}}$, to get the anonymised ECG, $y(t)$.

Step 6: Upload the anonymised ECG, $y(t)$, to the repository.

4.3. The ECG reconstruction method

The proposed reconstruction process for the anonymised ECG signal is shown in **Figure 8**, while the pseudocode is shown in Algorithm 2. The authorised personnel receives the secure key, K, and the anonymised ECG, $y(t)$, and performs the reconstruction process by the following steps:

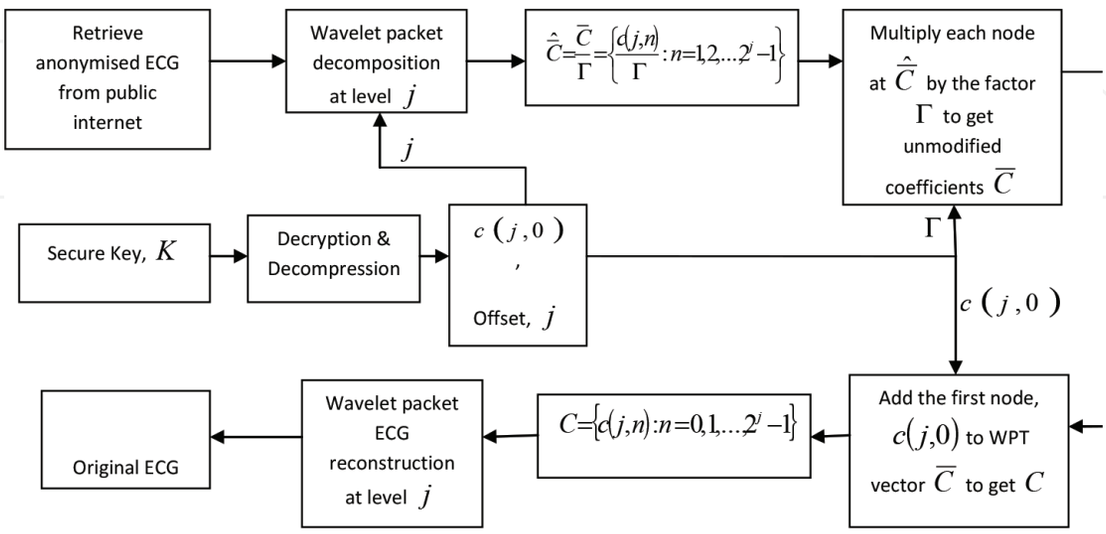

Figure 8. Wavelet packet-based reconstruction process for the anonymised ECG.

Step 1: Perform decryption and decompression to the key, K, to get Γ

$$\Gamma = \Lambda(D(K)), \tag{13}$$

where $\Lambda(.)$ and $D(.)$ are the decryption and decompression operators, respectively. Decryption and decompression are beyond the scope of this chapter.

Step 2: Perform wavelet packet decomposition of the ECG signal, $y(t)$, at level j to get $\widehat{\overline{C}} = \frac{\overline{C}}{\Gamma}$ as in Eq. (11).

Step 3: Multiply each node at $\widehat{\overline{C}}$ by the factor Γ to get

$$\overline{C} = f^{-1}\left(\widehat{\overline{C}}\right) = \widehat{\overline{C}} \times \Gamma = \left\{c(j,n) : n = 1, 2, \ldots, 2^j - 1\right\} \tag{14}$$

Algorithm 1: Wavelet packet-based ECG anonymisation process

1: Begin
2: $\boldsymbol{x(t)} \leftarrow$ ***ECG_signal***
3: $\boldsymbol{C} \leftarrow$ ***wpacket_decomposition***$\left(\boldsymbol{x(t)}, j\right)$
4: $\boldsymbol{k} \leftarrow \boldsymbol{c(j,0)}$ *// exclude the first node as a key*
5: $\overline{\boldsymbol{C}} \leftarrow \boldsymbol{c(j,n)}$
6: **offset** $\leftarrow |\mathbf{min}(\boldsymbol{k})| + \boldsymbol{\eta}$
7: $\boldsymbol{\Gamma} \leftarrow \boldsymbol{k}$ + **offset**
8: $\widehat{\overline{\boldsymbol{C}}} \leftarrow \frac{\overline{\boldsymbol{C}}}{\boldsymbol{\Gamma}}$
9: $\boldsymbol{K} \leftarrow \boldsymbol{E}\left(\boldsymbol{\Delta}\left(\boldsymbol{k}, \boldsymbol{offset}\right)\right)$ *//compression and encryption*
10: Send $\boldsymbol{K}$ to healthcare providers or doctors as a key
11: $\boldsymbol{y(t)} \leftarrow$ ***wpacket_reconstruction***$\left(\widehat{\overline{\boldsymbol{C}}}, j\right)$
12: Upload $\boldsymbol{y(t)}$ to public server
13: Save $\boldsymbol{y(t)}$ with unique ID for a particular individual
14: End

Algorithm 2: Wavelet packet-based reconstruction process

1: **Begin**
2: $\boldsymbol{k} \leftarrow \boldsymbol{\Lambda}(\boldsymbol{D}(\boldsymbol{K},$ **offset**$))$ *// decryption and decompression*
3: $\boldsymbol{y(t)} \leftarrow$ ***Anonymised_ECG_signal***
4: $\widehat{\overline{\boldsymbol{C}}} \leftarrow$ ***wpacket_decomposition***$\left(\boldsymbol{y(t)}, j\right)$
5: $\overline{\boldsymbol{C}} \leftarrow \widehat{\overline{\boldsymbol{C}}} \times \boldsymbol{\Gamma}$
6: $\boldsymbol{c(j,n)} \leftarrow \overline{\boldsymbol{C}}$
7: $\boldsymbol{k} \leftarrow \boldsymbol{\Gamma}$ − **offset**
8: $\boldsymbol{c(j,0)} \leftarrow \boldsymbol{k}$
9: $\boldsymbol{C} \leftarrow$ ***add_first_node***$\left(\boldsymbol{c(j,0)}, \boldsymbol{c(j,n)}\right)$
10: $\boldsymbol{x(t)} \leftarrow$ ***wpacket_reconstruction***$\left(\boldsymbol{C}, j\right)$
11: **End**

Step 4: Add the first node, $c(j,0) = \Gamma -$ offset, to the WPT vector $\overline{C}$ at Eq. (14) to get the WPT vector coefficients, C, of the original ECG signal, $x(t)$.

Step 5: Perform wavelet packet reconstruction of the coefficients vector, C, at level j to recover the original unanonymised ECG signal, $x(t)$.

5. Algorithm validation

Two types of electrocardiogram (ECG) signals were used to validate and investigate the performance and the effectiveness of the generalised ECG anonymisation framework. These signals are

1. normal ECG signal for a healthy subject, and
2. abnormal ECG signals for a patient with supraventricular arrhythmia and a patient with ventricular tachyarrhythmia.

The normal and abnormal ECG signals with different sampling frequencies were used in this chapter to study the robustness of the proposed anonymisation approach in concealing and smearing the ECG's fiducial and non-fiducial features. The normal and abnormal ECG data were obtained from the PTB ECG database [26] and the MIT-BIH arrhythmia database [27], respectively. These databases are publically available [26, 27].

In the evaluation process in the latter sub-section, bior5.5 wavelet was used. Besides this type of mother wavelet resembling the shape of an ECG signal, it is widely used for speech, video and biomedical signals providing that bior5.5 inherited linear phase. Nevertheless, it should be noted that for ECG anonymisation in this chapter, mother wavelet will not impact the anonymisation result since the ECG signal will be constructed back to its original at the receiver side.

The security of the proposed scheme depends on the following parameters that are required at the receiver side:

1. the encrypted security key which should be shared secretly,
2. the reversible function that should be used to reconstruct the original ECG information from the anonymised ECG, and
3. the type of transformation and the level of decomposition (wavelet packet transform at level 2 is used in this study).

An attacker with stolen key (i.e. able to decrypt the secure key) using brute force or any other method will require the knowledge of the reversible function and the level of decomposition. This information will be stored inside a patient/medical personnel PC and will not be transmitted under any circumstance. In this case, brute force attack is infeasible for the attack.

In the following sections, performance analysis using cross-correlation of normal and anonymised ECG signals, power spectral density of anonymised ECG signal and percentage residual difference (PRD) methods will be examined.

5.1. Performance evaluation over normal electrocardiogram

An electrocardiogram (ECG) signal has a well-defined P, QRS and T signature that is represented with each heartbeat. The P-wave arises from the depolarisation of the atrium. The QRS complex arises from depolarisation of the ventricles and T-wave arises from repolarisation of the ventricle muscles. The duration, shape and amplitude of these waves are considered as major features in time-domain analysis. Sometimes the time morphologies of these waves are similar.

The normal ECG was obtained from the PTB database (patient247, signal s0479). The sampling frequency, f_s, for this signal was 1 kHz. A total of 10 s of this signal was transformed by wavelet packet decomposition at level 2, $j = 2$. Decomposition level, j, depends on the ECG sampling frequency. Higher sampling frequency requires a low value of j to conceal all features in the anonymised signal. Node $c(2,0)$ of size $\frac{N}{4}$ (N = 10,000 samples) was removed from the wavelet packet coefficients of the normal ECG signal. This node was used to generate the key, K, which was distributed securely to medical personnel. The anonymised ECG is reconstructed

from the rest nodes, three nodes, using the anonymisation algorithm in Section II (B) and transmitted confidently over the public internet, since the anonymised ECG does not impose any threat to privacy.

Figure 9 (a) and **(b)** shows the time-domain representation of the 10-s normal ECG signal (patient247, signal s0479) and its anonymisation version, respectively. The frequency range for the anonymised ECG after node $c(2,0)$ removal is 125 and 500 Hz. From the time-domain representation of the ECG signal and its anonymisation in **Figure 9 (a)** and **(b)**, the proposed anonymisation algorithm conceals all fiducial features from the reconstructed ECG signal (**Figure 9(b)**). **Figure 10 (a)** and **(b)** shows the frequency representation of the 10-s normal ECG signal (**Figure 9 (a)**) and its anonymisation version, respectively. The non-fiducial features were also concealed as shown in the frequency-domain representation of the anonymised version of the normal ECG signal.

Figure 11 shows the time-domain representation of the coefficients $c(2,0)$ which was used to create the secure key, K. The frequency range for $c(2,0)$ in this data is between 0 and 125 Hz. This node preserves all fiducial features in the original ECG signal. **Figure 12 (a)** and **(b)** shows the reconstructed ECG signal at the medical personnel side and its cross-correlation with the original ECG signal at the patient side, respectively. From **Figure 12 (b)**, both signals are highly correlated, which guarantees a lossless reconstruction.

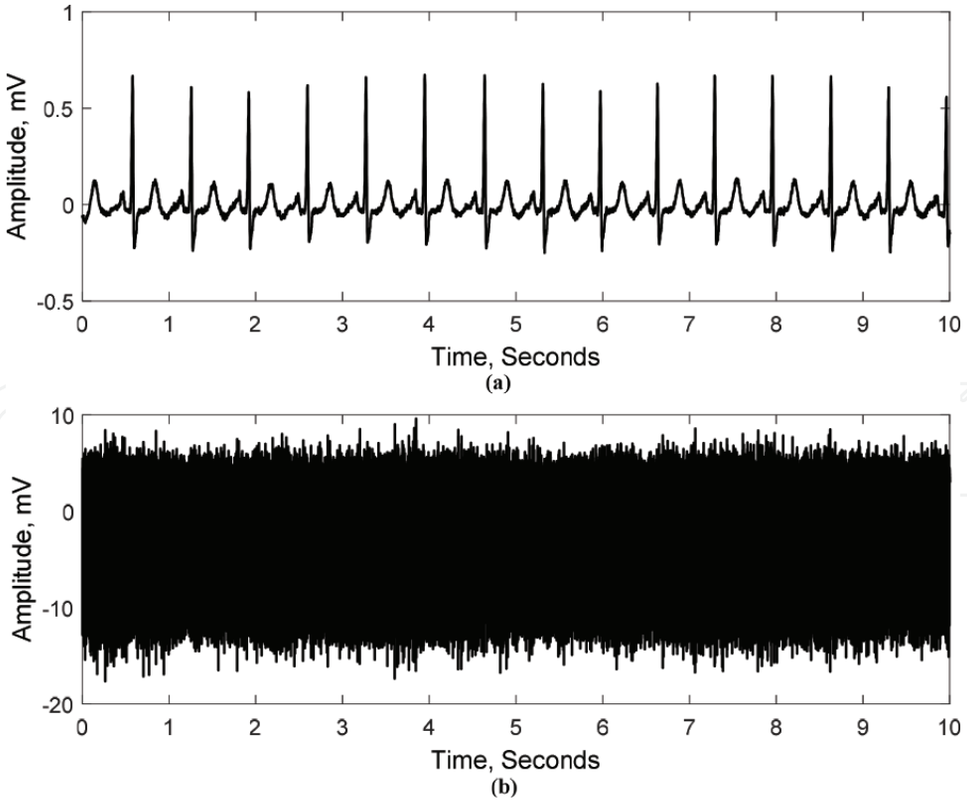

Figure 9. Time-domain representation of 10-s normal ECG signal, (a) unanonymised ECG signal and (b) anonymised ECG signal. The sampling frequency was f_s = 1 kHz.

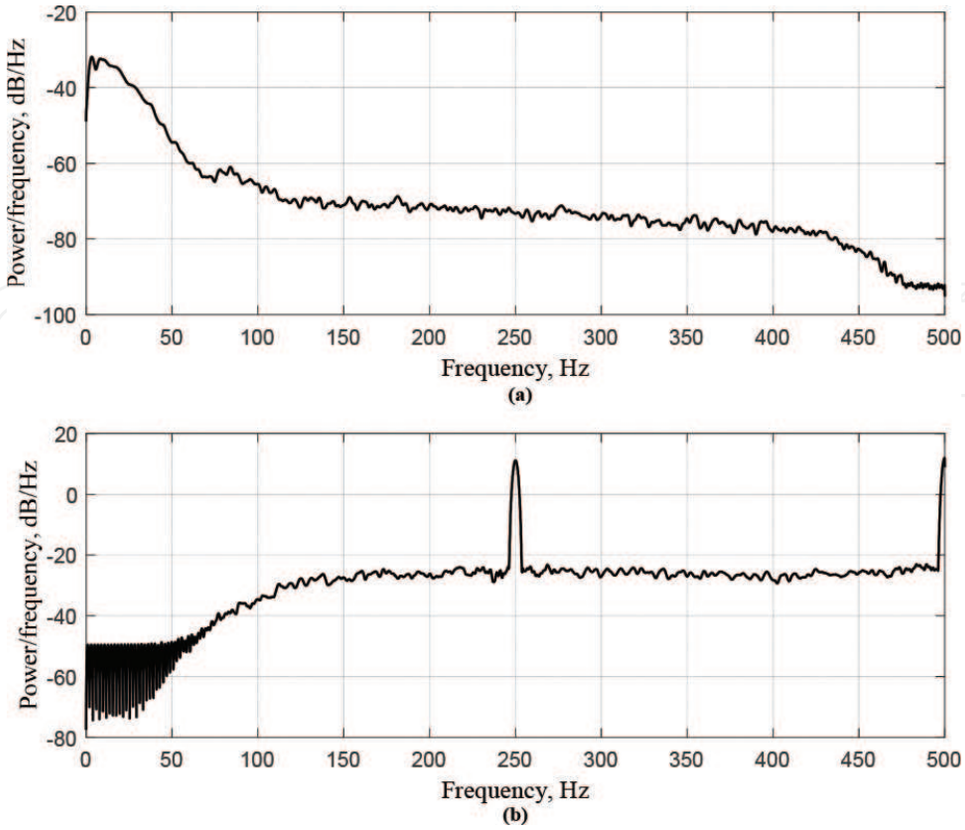

Figure 10. Power spectral density of 10-s normal ECG signal, (a) unanonymised ECG signal and (b) anonymised ECG signal. The sampling frequency was f_s = 1 kHz, the power spectral method was Welch periodogram.

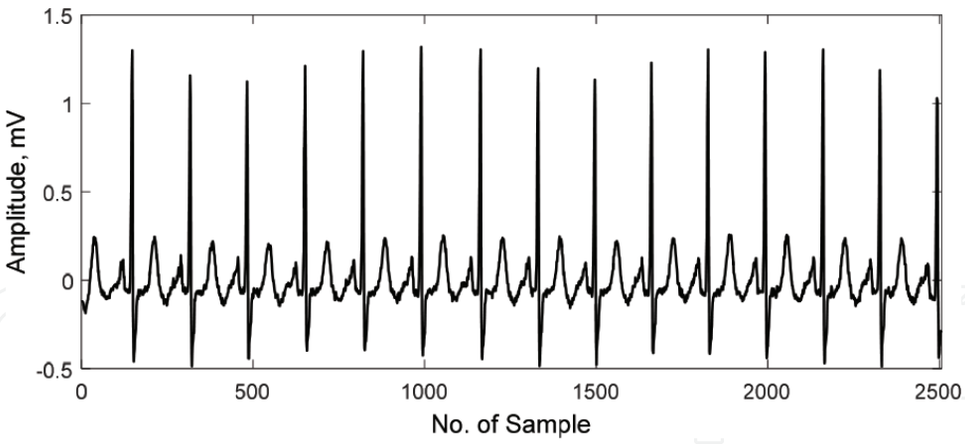

Figure 11. Time-domain representation of the first node $c(2,0)$ coefficients for the 10-s normal ECG signal in **Figure 9(a)**. This node was used to create the secure key.

5.2. Performance evaluation over abnormal electrocardiogram

An arrhythmia is an abnormality in the heart's rhythm or heartbeat pattern. The heartbeat can be too slow, too fast, have extra beats or otherwise beat irregularly [28]. The types of abnormal ECG signals investigated in this study were supraventricular arrhythmia and ventricular tachyarrhythmia. Supraventricular arrhythmia occurs in the upper areas of the heart and

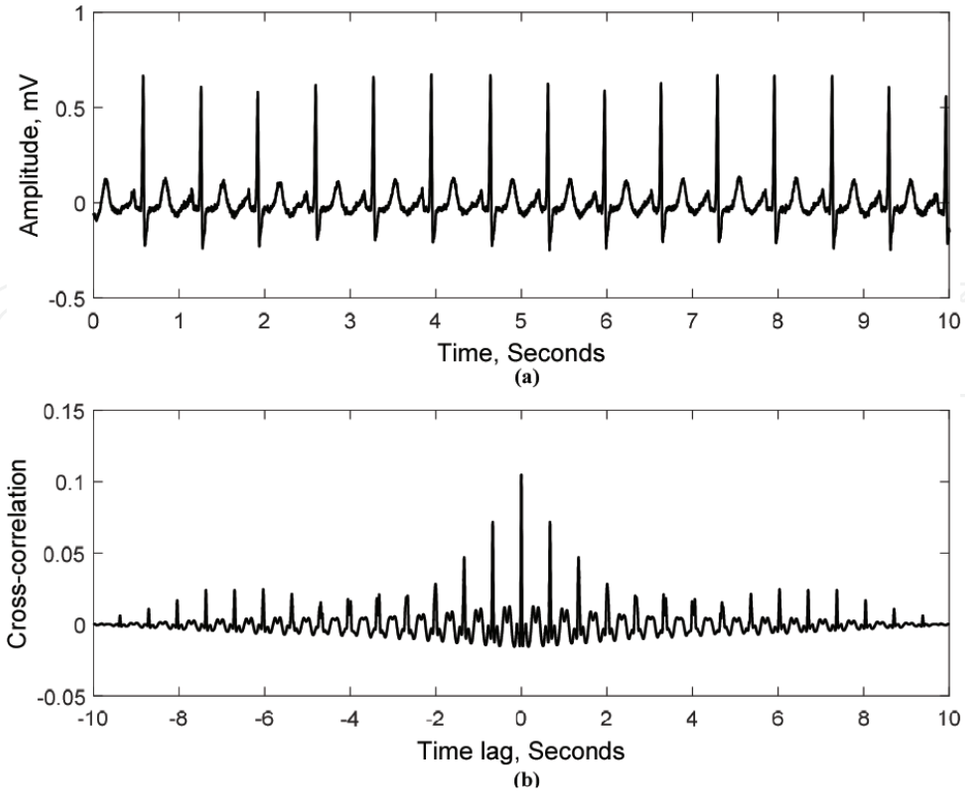

Figure 12. Ten seconds reconstructed ECG signal, (a) time domain representation of the reconstructed ECG signal, (b) cross correlation between the normal ECG signal in **Figure 9(a)** and its reconstructed version.

is less serious than ventricular arrhythmia. It has irregular shapes of QRS complexes [28]. These arrhythmia data—supraventricular arrhythmia and ventricular tachyarrhythmia—were obtained from the MIT-BIH arrhythmia database [26].

5.2.1. Supraventricular arrhythmia

The sampling frequency, f_s, for this signal was 128 Hz. A total of 10 s of this signal was transformed by wavelet packet decomposition at level 2, $j = 2$.

Node $c(2,0)$ of size $\frac{N}{4}$ (N = 1280 samples) was removed from the wavelet packet coefficients of the supraventricular arrhythmia signal. This node was used to generate the key, K, which was distributed securely to medical personnel. The frequency range for $c(2,0)$ in this data is between 0 and 16 Hz. The other nodes at level 2 with the frequency range between 16 and 64 were used to construct the anonymised signal.

Figure 13 (a) and **(b)** shows the time-domain representation of the 10-s ECG signal of a patient with supraventricular arrhythmia and its anonymisation version, respectively. The frequency-domain representation for both signals is shown in **Figure 14 (a)** and **(b)**. The fiducial and non-fiducial features were concealed in the time-domain and frequency-domain representation of the anonymised supraventricular arrhythmia signal.

Figure 15 shows the time-domain representation of the coefficients $c(2,0)$, which was used to create the secure key, K. This node preserves all fiducial features in the original supraventricular arrhythmia signal.

http://dx.doi.org/10.5772/intechopen.74477

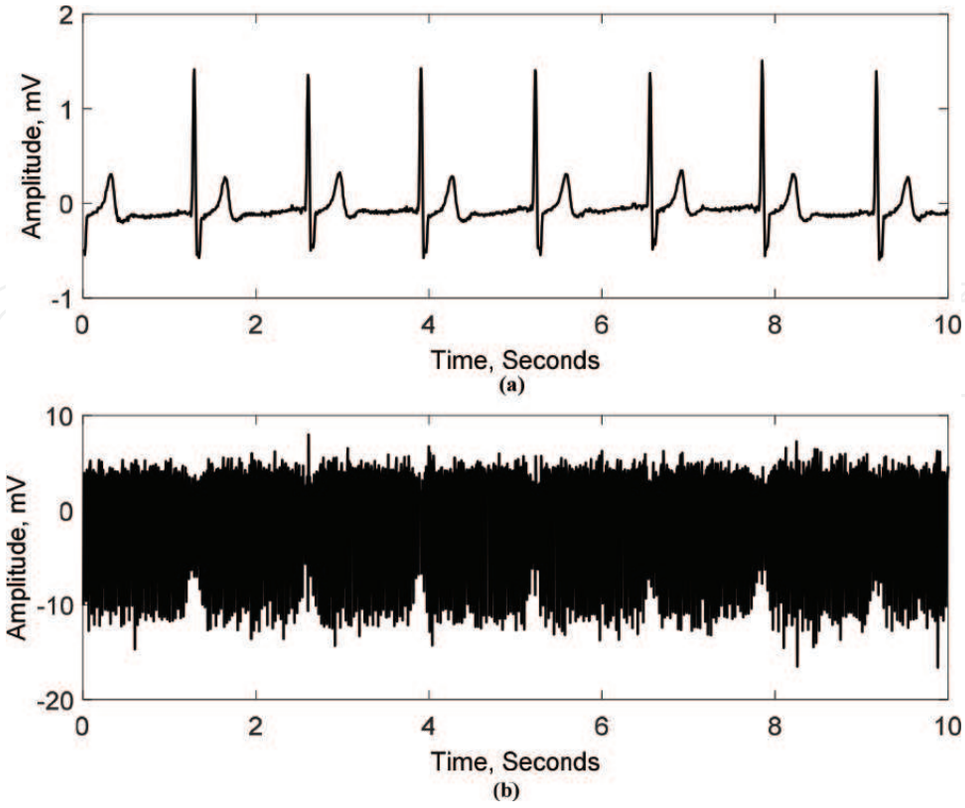

Figure 13. Time-domain representation of 10-s ECG signal of a patient with supraventricular arrhythmia, (a) unanonymised ECG and (b) anonymised CG. The sampling frequency was f_s = 128 Hz.

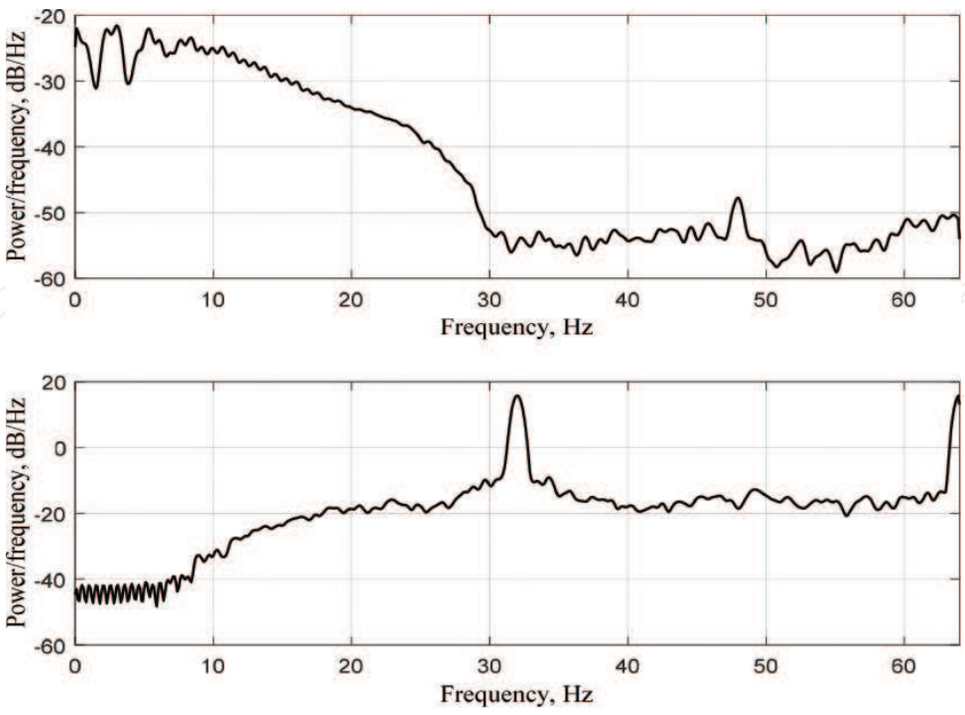

Figure 14. Power spectral density of ten seconds normal ECG signal of a patient with supraventricular arrhythmia, (a) unanonymised ECG and (b) anonymised ECG. The sampling frequency was f_s = 128 Hz, and the power spectral method was Welch periodogram.

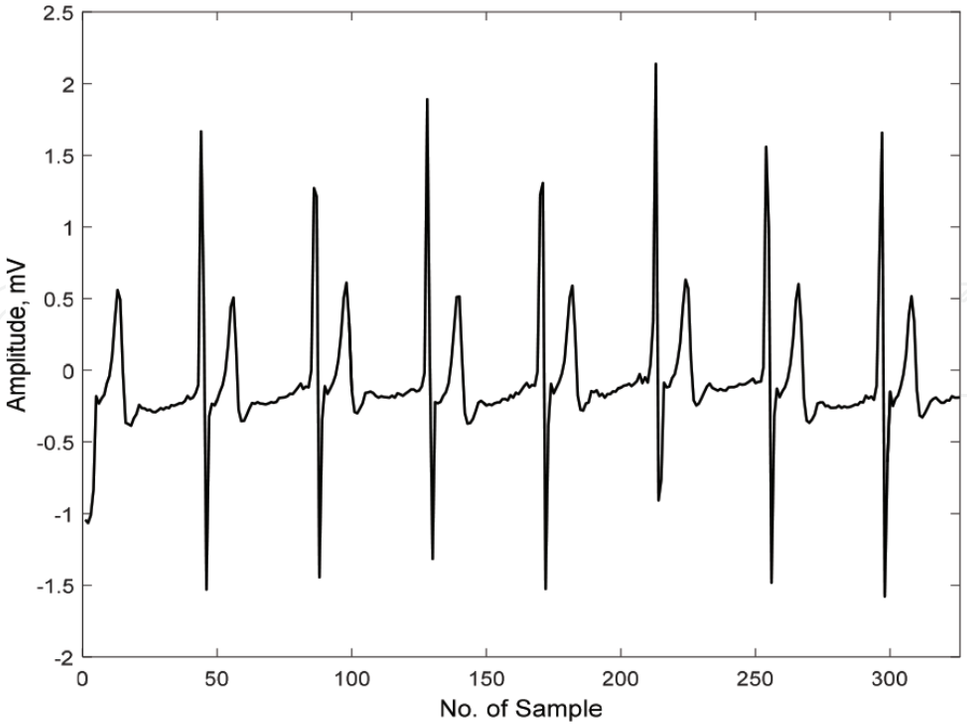

Figure 15. Time-domain representation of the first node $c(2,0)$ coefficients for the 10-s abnormal ECG signal in **Figure 7 (a)**. This node was used to create the secure key.

5.2.2. Ventricular tachyarrhythmia

The sampling frequency, f_s, for this signal was 250 Hz. A total of 10 s of this signal was transformed by wavelet packet decomposition at level 2. Node $c(2,0)$ of size $\frac{N}{4}$ (N = 2500 samples) was removed from the wavelet packet coefficients of the ventricular tachyarrhythmia signal. This node was used to generate the key *K*, which was distributed securely to medical personnel. The other nodes were used to reconstruct the anonymised ventricular tachyarrhythmia signal.

Figure 16 (a) and **(b)** shows the time-domain representation of the 10-s ECG signal of a patient with ventricular tachyarrhythmia and its anonymisation version, respectively. The frequency-domain representation for both signals is shown in **Figure 17 (a)** and **(b)**. The fiducial and non-fiducial features were concealed in the time-domain and frequency-domain representation of the anonymised supraventricular arrhythmia signal.

Figure 18 shows the time-domain representation of the coefficients $c(2,0)$ which was used to create the secure key, *K*. This node preserves all fiducial features in the original supraventricular arrhythmia signal.

5.3. Performance evaluation with the PRD metric

The percentage residual difference (PRD) is used to measure the difference between the original ECG signal and the anonymised ECG signal using the following equation.

http://dx.doi.org/10.5772/intechopen.74477

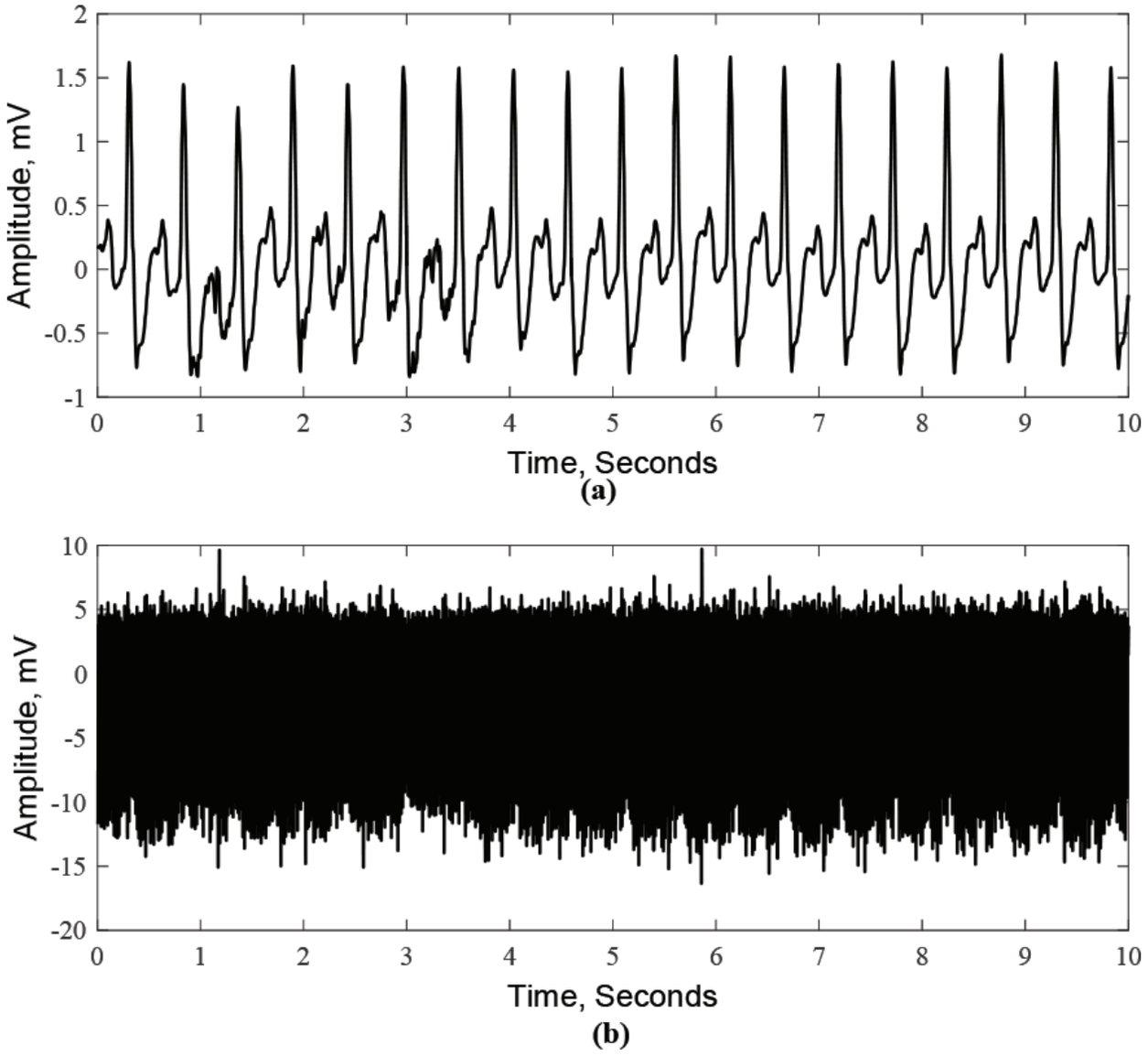

Figure 16. Time-domain representation of 10-s ECG signal of a patient with ventricular tachyarrhythmia, (a) unanonymised ECG and (b) anonymised ECG. The sampling frequency was f_s = 250 Hz.

$$\mathrm{PRD} = \sqrt{\frac{\sum_{i=1}^{N}(x(i)-y(i))^2}{\sum_{i=1}^{N}x(i)^2}} \tag{15}$$

where $x(i)$ is the original ECG signal, $y(i)$ is the anonymised ECG signal and $i = 1 \ldots N$, where N is the total number of the sample.

Performance of the proposed anonymisation algorithm using PRD metric is shown in **Table 1**. It can be seen from the table that the minimum and the maximum PRD measured were 14.8 and 70.6%, respectively. The PRD value depends on the ECG frequency bandwidth and its sampling frequency.

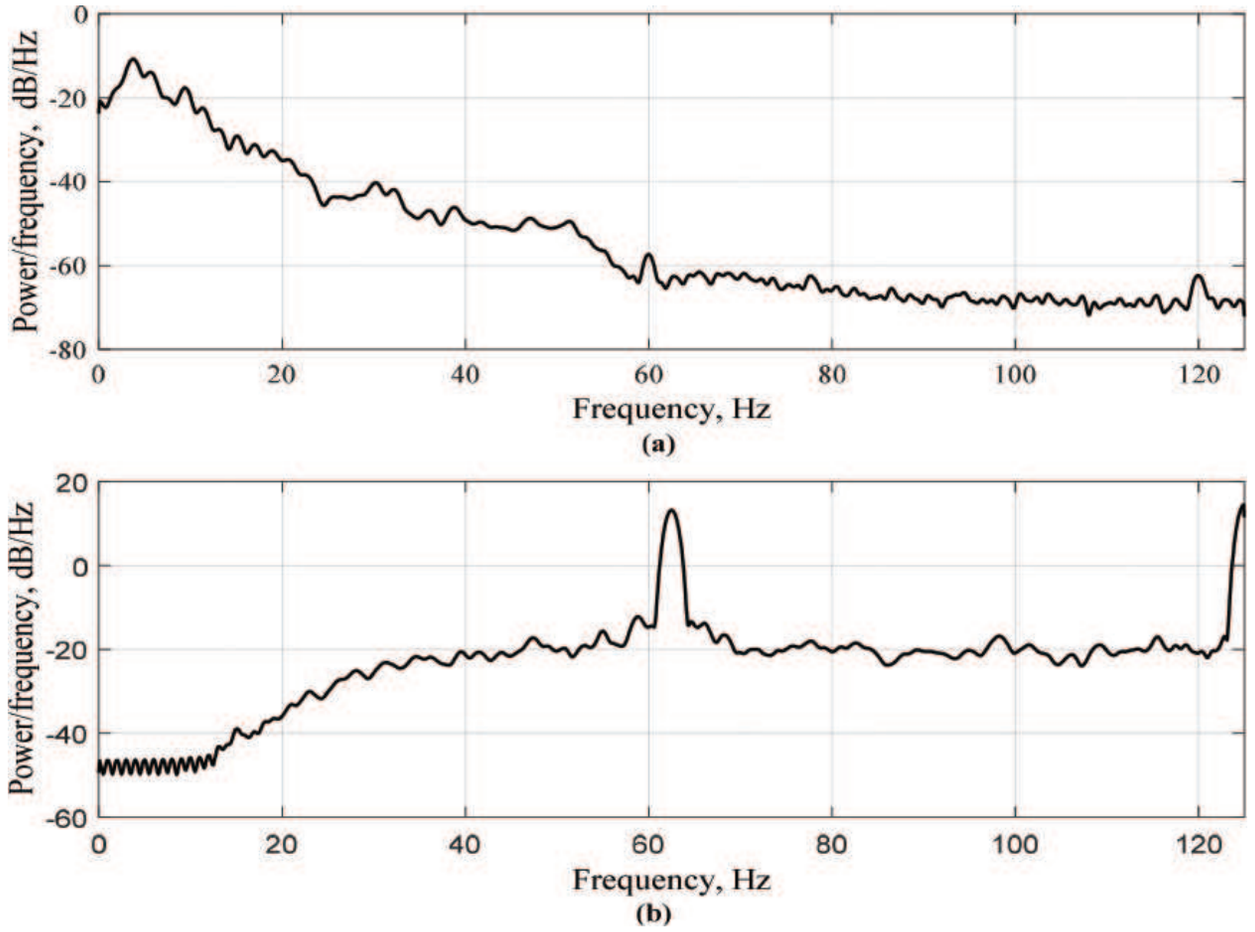

Figure 17. Power spectral density of 10-s ECG signal of a patient with supraventricular arrhythmia, (a) unanonymised ECG and (b) anonymised ECG. The sampling frequency was f_s = 250 Hz, the power spectral method was Welch periodogram.

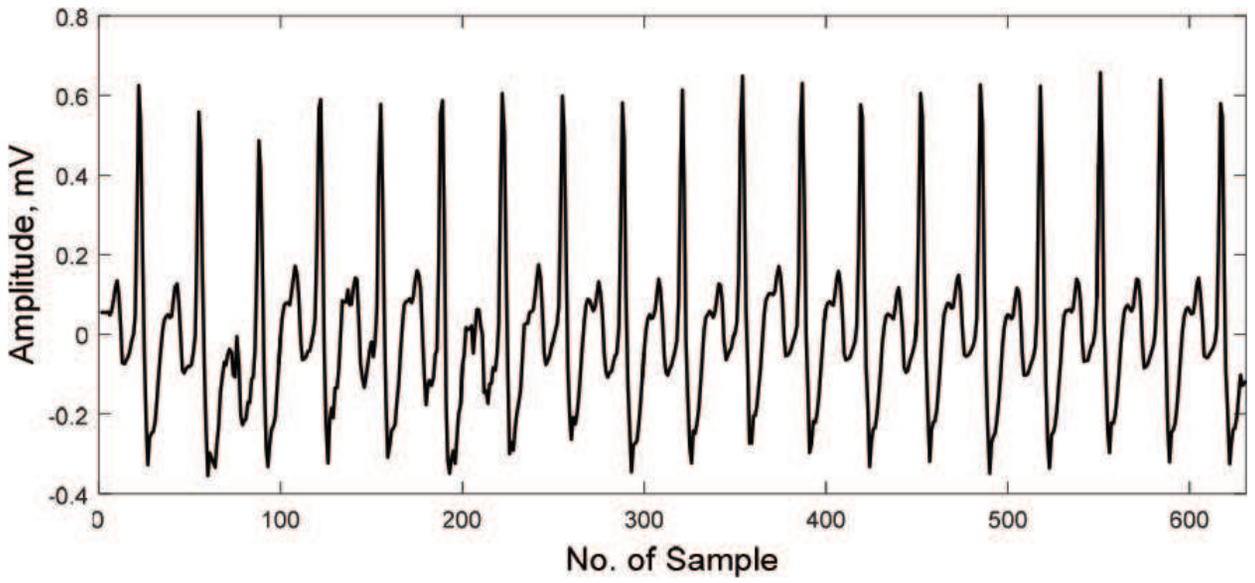

Figure 18. Time-domain representation of the first node $c(2,0)$ coefficients for the 10-s abnormal ECG signal in **Figure 10 (a)**. This node was used to create the secure key.

ECG type	Sampling frequency f_s, Hz	PRD %
Normal ECG	1000	70.6
Supraventricular arrhythmia	128	29.6
Ventricular tachyarrhythmia	250	14.8

Table 1. PRD performance results of normal and abnormal ECG signal for the proposed algorithm.

Comparing with ECG steganography methods, ECG steganography has a low PRD value between original and watermarked ECG signal. For example, in [14], the maximum PRD measured was 0.6%. Low PRD is essential in ECG steganography to guarantee correct diagnosis of the ECG watermarked signal. However, the lower value of PRD makes the ECG vulnerable to attack [1–3, 9].

6. Conclusions

A generalised wavelet packet-based ECG anonymisation framework has been presented in this chapter. This proposed anonymisation technique was used to conceal fiducial and non-fiducial features from normal and abnormal ECG signal for secure transmission over the public internet. Normal and abnormal ECG signals with different sampling frequencies have been investigated by the proposed method. Signal transformations other than wavelet packet transform can be used in this framework. Such transformations should have inverse property.

The performance analysis revealed that the proposed method is able to conceal both fiducial and non-fiducial features in normal and abnormal ECG signals under examination. Moreover, the analysis showed that the reconstructed ECG is highly correlated with the original ECG signal. It achieved a lossless reconstruction of the ECG data and proved the robustness of the proposed method. The security measures taken to secure the key and other information such as the level of decomposition and the knowledge of the reversible function make attacks using methods such as brute force is infeasible.

Author details

Seedahmed S. Mahmoud[1]* and Jusak Jusak[2]

*Address all correspondence to: seedahmed.sharif@gmail.com

1 Department of Electrical and Electronic Technology, Applied Engineering College, Lincoln College International, Buraidah, Riyadh, Kingdom of Saudi Arabia

2 Department of Computer Engineering, Institut Bisnis dan Informatika Stikom Surabaya, Surabaya, East Java, Indonesia

References

[1] Cisco. The Zettabyte Era: Trends and Analysis. White Paper at Cisco.Com, June 2016

[2] Atzori L, Iera A, Morabito G. Internet of things: A survey. Computer Networks. October 2010;**54**(15):2787-2805

[3] Andrews JG, Buzzi S, Choi W, Hanly SV, Lozano A, Soong ACK, Zhang JC. What will 5G be? IEEE Journal on Selected Areas in Communications. June 2014;**32**(6):1065-1082

[4] Islam SMR, Kwak D, Kabir MDH, Hossain M, Kwak KS. Internet of things for health care: A comprehensive survey. IEEE Access. 2015;**3**

[5] Jusak J, Puspasari I. Wireless tele-auscultation for phonocardiograph signal recording through the zigbee networks. IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob), Bandung, Indonesia, August 27–29, 2015

[6] Jusak J, Pratikno H, Putra VH. Internet of medical things for cardiac monitoring: Paving the way to 5G mobile networks. IEEE International Conference on Communication, Networks and Satellite, Surabaya, Indonesia, Dec. 2016

[7] Biel L, Petersson O, Philipson L, Wide P. ECG analysis: A new approach in human identification. IEEE Transaction on Instrumentation and Measurement. 2001;**50**(3):808-812

[8] Irvine JM, Wiederhold BK, Gavshon LW, Israel SA, McGhee SB, Meyer R, Wiederhold MD. Heart rate variability: a new biometric for human identification. International Conference on Artificial Intelligence, Las Vegas, Nevada; 2001. pp. 1106-1111

[9] Israel SA, Scruggs WT, Worek WJ, Irvine JM. Fusing face and ECG for personal identification. In: Proc. 32nd IEEE Appl. Imagery Pattern Recog. Workshop; 2003. pp. 226-231

[10] Chan ADC, Hamdy MM, Badre A, Badee V. Wavelet distance measure for person identification using electrocardiograms. IEEE Transactions on Instrumentation and Measurement. 2008;**57**(2):248-253

[11] Wubbeler G, Stavridis M, Kreiseler D, Bousseljot R-D, Elster C. Verification of humans using the electrocardiogram. Pattern Recognition Letters. 2007;**28**:1172-1175

[12] Odinaka I, Lai P, Kaplan AD, O'Sullivan JA, Sirevaag EJ, Rohrbaugh JW. ECG biometric recognition: A comparative analysis. IEEE Transactions on Information Forensics and Security. 2012;**7**(6):1812-1824

[13] Sufi F, Mahmoud SS, Khalil I. A Wavelet Based Secured ECG Distribution Technique for Patient Centric Approach. In: The Proceedings of 5th International Workshop on Wearable and Implantable Body Sensor Networks, Hong Kong, China; 2008

[14] Fahim KS, Mahmoud SS, Khalil I. A novel wavelet packet-based anti-spoofing technique to secure ECG data. International Journal of Biometrics. 2008;**1**(2):191-208

[15] Reinsmidth E, Schwab D, Yang L. Securing a connected mobile system for healthcare. IEEE 17th International Symposium on High Assurance Systems Engineering (HASE), Orlando, FL, USA. Jan. 2016. pp. 19-22

[16] Department of Health & Human Services USA. Security 101 for covered entities. HIPAA Security Series. 2007;**2**:1-11

[17] European Parliament and of the Council. Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official Journal of the European Communities. 1995;**1**:281/31-281/50

[18] Privacy Commisioner. Health Information Privacy Code 1994. Ed. 2008. Auckland, New Zealand: KB Printed Ltd; 2008

[19] Pearce C, Bainbridge M. A personally controlled electronic health record for Australia. Journal of the American Medical Informatics Association. 2014;**21**(4):707-713

[20] Özkaynak F, Yavuz S. Analysis and improvement of a novel image fusion encryption algorithm based on DNA sequence operation and hyper-chaotic system. Nonlinear Dynamics. 2014;**78**(2):1311-1320

[21] Ibaida A, Khalil I. Wavelet-based ECG steganography for protecting patient confidential information in point-of-care systems. IEEE Transactions on Biomedical Engineering. 2013; **60**(12):3322-3331

[22] Tseng K, He X, Kung W, Chen S, Liao M, Huang H. Wavelet-based watermarking and compression for ECG signals with verification evaluation. Sensors. 2014;**14**(2): 3721-3736

[23] Liji CA, Indiradevi KP, Babu A. Integer-to-integer wavelet transform based ECG steganography for securing patient confidential information. Procedia Technology. 2016;**24**: 1039-1047

[24] Chen C-K et al. Personalized information encryption using ECG signals with chaotic functions. Information Sciences. 2012;**193**:125-140

[25] Bartolo A, Clymer BD, Bugess RC, Turnbull JP, Golish JA, Perry MC. An arrhythmia detector and heart rate estimator for overnight polysomnography studies. IEEE Transactions on Biomedical Engineering. 2001;**48**(5):513-521

[26] The PTB Diagnostic ECG Database, http://www.physionet.org/ physiobank/database/ ptbdb (viewed Dec. 2015)

[27] MIT-BIH Arrhythmia Database, http://physionet.org/physiobank/database/svdb/ (viewed Dec. 2015)

[28] Hebbar AK, Hueston WJ. Management of Common Arrhythmias: Part I. Supraventricular arrhythmias. Journal of American Family Physician. 2002;**65**(12):2479-2486

IntechOpen