Walsh-Hadamard Based 3D Steganography for Protecting Sensitive Information in Point-of-Care

Alsharif Abuadbba, *Student Member, IEEE,* and I Khalil, *Senior Member, IEEE*

***Abstract*—Remote points-of-care has recently had a lot of attention for their advantages such as saving lives and cost reduction. The transmitted streams usually contain (1) normal biomedical signals (e.g. ECG) and (2) highly private information (e.g. patient identity). Despite the obvious advantages, the primary concerns are privacy and authenticity of the transferred data. Therefore, this paper introduces a novel steganographic mechanism that ensures (1) strong privacy preservation of private information by random concealing inside the transferred signals employing a key, and (2) evidence of originality for the biomedical signals. To maximize hiding, Fast Walsh-Hadamard Transform is utilized to transform the signals into a group of coefficients. To ensure the lowest distortion, only less-significant values of coefficients are employed. To strengthen security, the key is utilized in a 3-Dimensional random coefficients' reform to produce a 3D order employed in the concealing process. The resultant distortion has been thoroughly measured in all stages. After extensive experiments on three types of signals, it has been proven that the algorithm has little impact on the genuine signals ($<$ 1 %). The security evaluation also confirms that unlawful retrieval of the hidden information within rational time is mightily improbable.**

***Index Terms*—Steganography, Walsh-Hadamard, Privacy Preservation, Authenticity, Watermarking, Biomedical Signal.**

I. Introduction

THE traditional healthcare systems alone where the patients should be physically in hospitals to be monitored is presently considered as inappropriate to the current Century needs for several causes such as significant increase in the number of elderly people who suffer from cardiac diseases, hospitals and expertise shortage especially in rural areas [1]. Therefore, a new model called "Point-of-Care" (PoC) has recently emerged and can be used to remotely monitor patients (e.g. in homes) by collecting continuous samples every short time (i.e. minute) with smart sensors (e.g. temperature, blood pressure glucose levels and biomedical signals) and send them to health authorities using various techniques [2], [3]. The major advantages are saving people's life, reducing the medical labor cost and burden on hospital infrastructure.

However, despite the obvious benefits, remote PoC causes various security threats such as their existence in distant areas, transfer of the extremely private data of patients through public networks [4]. Consequently, many countries such as US (e.g. Health Insurance Portability and Accountability Act HIPAA) [5], and Australia (e.g. The Regulation of Health Information Privacy) [6] have laid hard regulations on hospitals imposing that patients' private information must be protected from unauthorized access especially when the health authorities perform off-site operations (i.e. cloud environment). Actually, there are two concerns: the primary concerns from patients' viewpoint are their sensitive information privacy (i.e. confidentiality issue), and the validity of the transferred samples and resultant professional decisions (i.e. Does this diagnosis is precise and belongs to me?); the major worry from health authorities' point of view is guaranteeing an effective and robust mechanism that assists in protecting the patients sensitive information.

Most of early models that treat these issues are using classical cryptography [7], [8], [9], [10], [11]. In spite of their appropriate functionalities, they are restricted in this field due to:

- Resource-limited PoC mobile devices having restricted memory and power usually make it tough to implant classic cryptography techniques because of the enormous introduced overhead from the required mathematical operations to guarantee strong security.
- Manipulating the original samples into a ciphertext, tightens the health authorities' job particularly when off-site processes are required to boost the system's mission and scalability.

To overcome some of these matters, a homomorphic (i.e. non-classical cryptography) encryption mechanisms have been utilized [12], [13], [14], [15]. The preference of homomorphism over the classical cryptography is that "ciphertext" (i.e. the encrypted form) can be used without revealing its content and thus satisfying both the patients worry by ensuring a robust end-to-end protection and the concerns of health authorities by permitting a direct use of cloud without any exposure. Nevertheless, due to its expensive computations, homomorphic cryptography is not yet feasible for similar applications.

Therefore, researchers are obliged to look for other solutions that tackle the primary worries of both the patients and health authorities together to: (1) ensure robust end-to-end privacy preservation for the patient's private information that will be transmitted such as personal data (e.g. name, Date-of-Birth DoB, geometric location), biometric data (e.g. iris and fingerprint) and diagnosis data (e.g. blood pressure and glucose level) as well as the authentication of normal collected biomedical signals (e.g. Electrocardiograms ECG, Electroencephalography EEG and Photoplethysmogram PPG), (2) be feasible with current PoC capacity such as memory, power consumption and bandwidth, and (3) avoid hindering the performance at health authorities (i.e. conducting quick and secure offshore operations).

Alsharif Abuadbba and Ibrahim Khalil are with the School of Computer Science and IT, RMIT University, Melbourne, VIC, Australia, 3001. E-mail: alsharif.abuadbba@gmail.com and ibrahim.khalil@rmit.edu.au

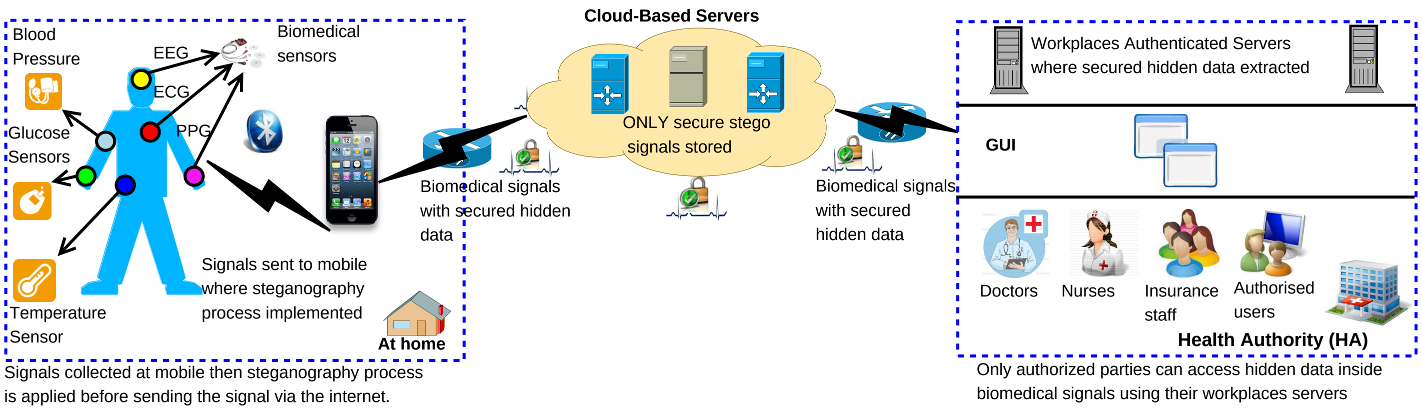

Blood Pressure
EEG
Biomedical sensors
ECG
PPG
Glucose Sensors
Temperature Sensor
Signals sent to mobile where steganography process implemented
At home
Signals collected at mobile then steganography process is applied before sending the signal via the internet.
Biomedical signals with secured hidden data
Cloud-Based Servers
ONLY secure stego signals stored
Biomedical signals with secured hidden data
Workplaces Authenticated Servers where secured hidden data extracted
GUI
Doctors
Nurses
Insurance staff
Authorised users
Health Authority (HA)
Only authorized parties can access hidden data inside biomedical signals using their workplaces servers

Fig. 1. The essential layout of the introduced technique where patient's confidential information is randomly concealed inside the periodically collected biomedical signals and only legitimate parties can extract this information.

Steganography is also a technique that is used to protect the private information where a secret piece (i.e. watermark) is embedded inside a wider transmitted message and which will be extracted by legitimate parties [16]. The advantage of steganography is that (1) it can be performed with much lower memory and power, and (2) it ensures the originality without manipulating the data which makes the technique as a possible candidate that can be used to overcome the limitations of PoC infrastructure, but the steganography alone can not rule who can retrieve the secured data.

A. Contribution

This article introduces a novel steganographic model that (1) provides privacy preservation for the patient's sensitive information and (2) guarantees the authenticity of normal collected biomedical signals (e.g. ECG, EEG, PPG). To overcome the steganography confidentiality problem, a security key is employed to randomly hide the private information on a bit level inside the biomedical signals. To get over the capacity issue, a simple signal processing mechanism called Fast Walsh-Hadamard (FWHT) is utilized to convert the biomedical signals from the spatial to frequency domain. The result will be a group of values called coefficients [17]. These coefficients can be categorised into: (1) low series samples regarded very significant to rebuild the signal, and (2) high series coefficients which have little impact on reconstructing the signal. To ensures the lowest deformation, only least-significant values are employed to embed the patient's private information. Our major contributions are three folds: (a) to the best of our knowledge, it is the first steganographic based privacy and authenticity preserving model that targets various biomedical signals (ECG, EEG and PPG), (b) it employs a light mathematical transformation (FWHT) in biomedical field, and (c) it increases hiding process strength by performing it following a dynamic random $3D$ generated matrix using the key.

In our paradigm (See Fig 1), remote sensors are used to gather various periodical samples from the patients. These readings usually contain two types of data: (1) normal biomedical signals (e.g. ECG, EEG and PPG) and (2) highly confidential information such as personal (e.g. name, DoB and geometric location), biometric (e.g. fingerprint or iris) and diagnosis (e.g. temperature, blood pressure and glucose level). The steganographic algorithm then is performed within the used PoC and private patient data is randomly concealed within the genuine biomedical signals. Lastly, the stego signals (i.e. contains the embedded sensitive) are transmitted to the distant health authorities via the Internet. Thus, the actual transferred data volume is just the genuine signals volume without any overhead, because the private bits are concealed within the signals. The stego signals are kept and managed by health authorities. However, only authorized parties are able to extract the private information utilizing an appropriate key, whereas others (e.g. cloud owners) can just access the stego signals. The second merit is that, from the empirical results, the diagnosis can be accurately performed on the stego signals without removing the hidden bits whenever utilizing the signals.

The rest of the work is organised as follows. Section II summarizes the related models. Section III shows our technique in main steps. Evaluation of various features of the introduced model is then presented in Section IV. Section V discusses our implemented experiments and the obtained results. Lastly, the work is concluded in Section VI.

II. Related Work

Any proposed solution to the remote PoC should carefully consider two main characteristics: security (i.e. solid privacy of the transmitted private data and the authenticity of the collected signals) and efficiency (i.e. permitting direct operations to be applied to the stego signals (e.g. at cloud) without sensitive information disclosure). Although steganography has been widely studied and used in the multimedia domain (e.g. Image [22], audio [23], video [24] and sensor streams [25], [26]), but it is rarely studied in the biomedical streams. This is because in the multimedia domain the imperceptibility (i.e. how people can sight) to human senses is the top priority, whereas in the biomedical signal's context the sensitivity is the most crucial factor in the diagnosis. This renders employing

TABLE I
RELATED WORK SUMMARY.

Model	Features	Comments
Zheng and Qian (2008) [18].	• Apply B-Spline wavelet into ECG samples. • Select only non-QRS top-frequency values. • Shift coefficients 1-bit for hiding. • Apply Arnold transform for scramble the watermark.	• Low hiding capacity. • Tested on ECG only. • No security key used.
Golpira and Danyali (2009) [19].	• Apply 2D wavelet on the host data. • Apply histogram on high frequency coefficients. • Select 2 threshold/ shift for hiding.	• Good with MRI images. • Bad with ECG. • No security key used.
Kaur et al (2010) [20].	• Quantize ECG using 10 bits. • Split the signal into two parts. • Use Window-Dependent Factor to modulate the signal. • Patient ID is used in modulation process. .	• Low hiding capacity. • Tested only with ECG. • Simple/easy-to-break key (i.e. Patient ID).
Ibaida and Khalil (2013) [21].	• Apply 2D five-level wavelet on the ECG signal. • Generate a 2D hiding order. • XOR's the key with sensitive info. • Follow 2D order to hide the secret bits. .	• Good hiding capacity. • Tested only with ECG. • 2D hiding order is static.

the biomedical signals as a cover medium in steganography much more difficult and complicated.

Table I summarizes most of the related work of steganography in the biomedical signals context. There are two primary limitations in common among most existing solutions: (1) although the traditional one-dimensional signal processing techniques (e.g. Haar Wavelet) are relatively simple (i.e. linear complexity), the currently used variations to produce the entire multidimensional sub-bands wavelet tree (i.e. N-dimensions) are more complicated (i.e. quadratic complexity) and can be computationally more expensive where it mainly uses multiplications [27], and (2) they were designed and experimented on a single type of biomedical signals (i.e. ECG).

III. METHODOLOGY

In the proposed steganographic algorithm, a suitable balance between the two main concerns (i.e. security and efficiency) has been carefully considered as follows: (1) it is highly improbable for illegitimate users to extract the concealed private data, and (2) there is no deformation on the original transmitted biomedical signals, to be directly usable without extracting the stego (i.e. embedded bits). The procedures at the distant PoC are explained in the subsequent sections.

A. Walsh-Hadamard Transform

WHT is a recognized linear mathematical operations which targets transforming a signal from its time to its frequency domain as a group of coefficients [28]. The importance of that is the obtained values can be categorised into: (1) low series values representing the crucial part to rebuild the signal, and (2) wider series of values that have little effect on signal reconstruction. The interesting observation is that the ability of almost reconstructing the genuine signal using just a small percentage of coefficients.

For clarity, Fig 2 shows an instance of three types of biomedical signals (i.e. ECG, EEG and PPG). (a) The plot of $>$ 2000 samples. (b) The obtained coefficients after performing WHT. This plot highlights the crucial parts of coefficients (i.e. low series) between 0 and <400, while the rest are less significant. Based on this observation, we wiped all values between >400 and 2048 to monitor their impact on the recomposed signal. (c) The rebuilt biomedical signals using just <400 values. This example obviously emphasizes the capacity and the flexibility that can be obtained by exploiting these values. This motivates us to investigate the possibility of employing such a technique to embed more patient's private information while strengthening the security into 3D level.

There is also a rapid variation model of WHT that achieves a linearithmic computational complexity $n \log n$ and only uses additions and subtractions [29], whereas the complexity of traditional form of WHT is quadratic $O(n^2)$ [30]. Therefore, our algorithm uses FWHT as it is shown in Eq 1.

$$b_n = \frac{1}{M} \sum_{1=0}^{M-1} a_i FWHT(n,i), n = 1, 2, ..., M-1 \qquad (1)$$

where b_n is the obtained FWHT values, a_i is the genuine readings, and $FWHT(n,i)$ is the performed transform.

The main idea is that a Walsh preconfigured matrix that should be compatible with the signal samples is applied. This matrix contains only $+1/-1$ [31]. The uniqueness of this transformation is that the values of the matrix can be tuned based on the targeted application. For examples, signal processing applications can employ Sequence order, control applications can use Hadamard arrangement, and mathematics applications mainly use Dyadic order. Four readings mathematical representations with compatible FWHT matrix are depicted in Eq 2.

$$\begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} \cdot \begin{pmatrix} +w_{11} & +w_{12} & +w_{13} & +w_{14} \\ +w_{21} & +w_{22} & -w_{23} & -w_{24} \\ +w_{31} & -w_{32} & -w_{33} & +w_{34} \\ +w_{41} & -w_{42} & +w_{43} & -w_{44} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \qquad (2)$$

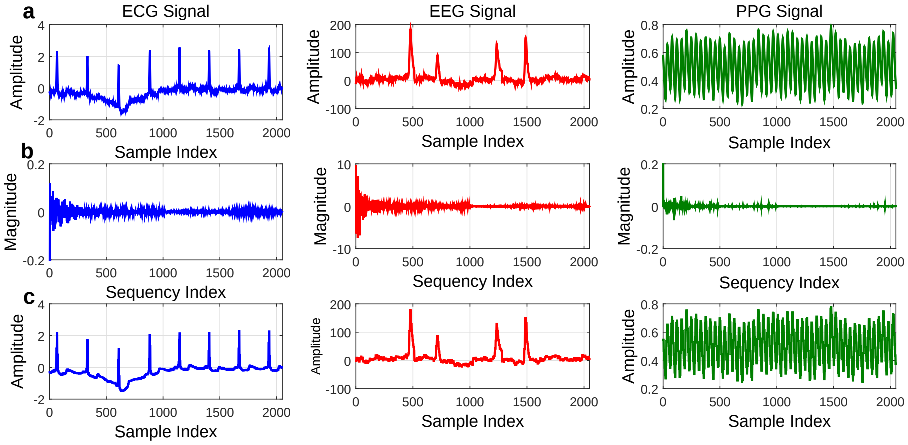


Fig. 2. Three types of biomedical signals (i.e. ECG, EEG and PPG) (a) Original samples, (b) FWHT coefficients and (c) Reconstructed signal using $< \%20$ - just low sequence values.

Where S_s are the original biomedical signal, w_s are the values of FWHT and c_s are the obtained output.

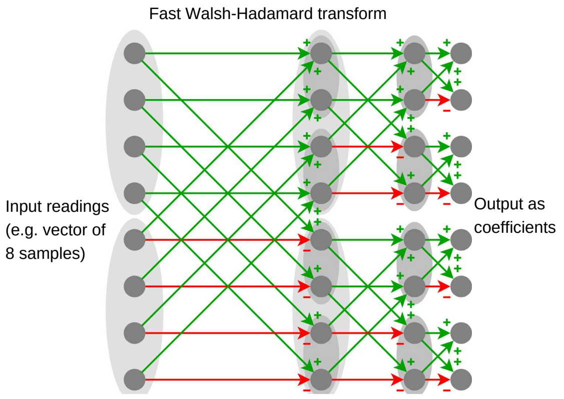


Fig. 3. Graphical representation of the fast Walsh Hadamard transform implemented with an eight samples vector.

FWHT is chosen in this paper for various purposes: (1) the genuine biomedical signal can be precisely recomposed relying on few low series values, opening the door for employing the majority of remaining coefficients in both strengthening the security into 3D and embedding a larger number of secret bits, and (2) due to the simplicity of FWHT mathematical model (i.e. only additions and subtractions) [17], it requires less computations, memory and power compared to many already used signal processing in biomedical field such as Fourier and Wavelet. This is because they are using multiplication operations.

Thus, in this work FWHT is implemented to various continuously gathered biomedical signals (e.g. ECG, EEG and PPG) and the obtained values are reformatted to a $3D$ template. The beginning serious values are not touched due to their importance in reconstructing the origin. In contrary, the reset of coefficients are used to accommodate various bits per value. To ensure the possible number of bits (i.e. stego level) can be changed per coefficient without noticeable effect, many empirical tests have been done to select an accurate stego level as depicted in Fig 4. From that, it is acceptable to embed up to five bits in every value of all high serious coefficients.

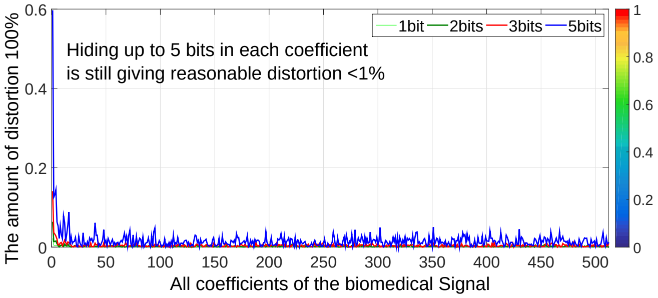


Fig. 4. Noise impact of embedding different levels 1bit-to-5bits and the obtained distortion.

B. Hiding Operation

In this step, confidential patient data is embedded within the obtained values, after implementing FWHT to the collected biomedical signals. However, to strengthen the protection while prohibiting unauthorised extraction, a patient key is generated for every distributed PoC and should be shared only with the end legitimate recipient of the information. This key is used to enforce three security layers.

1) Encryption: The key is employed to cipher the private information such as personal (e.g. ID, name, address, DoB and geometric location), diagnosis (e.g. temperature, glucose level and blood pressure) and biometric (e.g. iris and fingerprint) prior to the embedding operation utilizing a light encryption (e.g. AES), that is strong and fits remote PoC technical capabilities (See Fig 5). This is defined in Eq 3.

$$\widetilde{P_c} \Leftarrow f_e(K, P_c) \tag{3}$$

where f_e is an AES algorithm, K is the key, P_c is the original patient confidential information and $\widetilde{P_c}$ is the encrypted form.

2) Coefficients Shuffling: The key is employed along with a rotation factor to dynamically reform the obtained FWHT values from a vector into $3D$ $X \times Y \times Z$ matrix (See Fig 6).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2016.2631885, IEEE Transactions on Biomedical Engineering

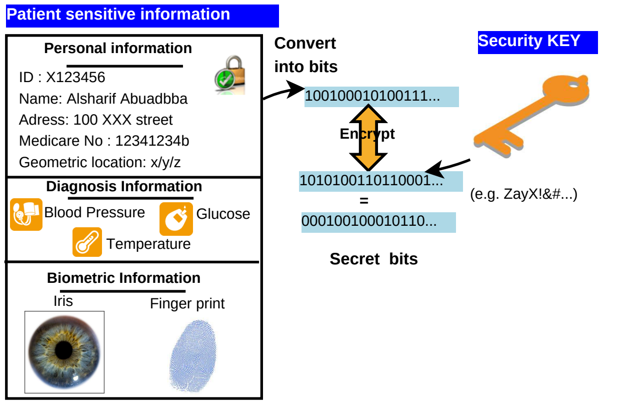


Fig. 5. A paradigm reflecting encrypting a patient private information prior to embedding.

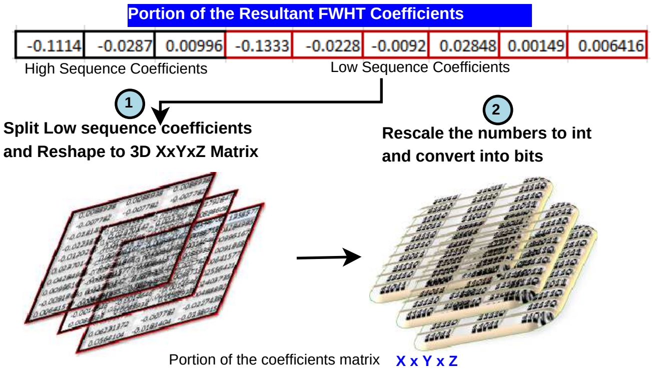

Portion of the Resultant FWHT Coefficients
-0.1114
-0.0287
0.00996
-0.1333
-0.0228
-0.0092
0.02848
0.00149
0.006416
High Sequence Coefficients
Low Sequence Coefficients
1
Split Low sequence coefficients and Reshape to 3D XxYxZ Matrix
2
Rescale the numbers to int and convert into bits
Portion of the coefficients matrix
X x Y x Z

Fig. 6. Block diagram presents how the FWHT coefficients are split, reshaped, rescaled and converted into bits.

3) Key-Driven Random Order: The key is utilized to produce a unique sequence in a $3D$ template which has to be followed precisely to embed the sensitive patient's data. This can be shown in Eq 4.

$$\widetilde{X} \times \widetilde{Y} \times \widetilde{Z} \Leftarrow f_x(K) \tag{4}$$

Where $\widetilde{X}\times\widetilde{Y}\times\widetilde{Z}$ is the generated $3D$ sequence of coefficients. f_x represents the random coefficients order algorithm. K is the chosen key to seed this algorithm.

Fig 7 demonstrates the steps followed to produce the chosen key-driven random FWHT values order. It begins by rotating the key using an introduced rotation value δ. The main goal is to generate different random order every single time. Next, the direct rotated key form is converted into the American Standard Code followed by allocating a sequential order. After that, the codes are ordered in a scaling up manner along with assigning a second position sequence. The key codes are also reshuffled to its initial status relying on the first allocated sequence resulting in $\widetilde{X}$. This step is to reduce the chances of producing two identical sequences in short period from various codes. On the other hand, $\widetilde{Y}$ is produced following identical operation but with differences: (i) an inverse key form, and (ii) scaling down the order. Before last, three tokens are randomly pulled from the key using the rotation factor after a half splitting process. They are used to generate $\widetilde{Z}$ following the previous steps. Lastly, a 3D template is built by interleaving the three produced sequences representing rows $\widetilde{X}$, columns $\widetilde{Y}$ and depth $\widetilde{Z}$. Also, the rotation factor is increased to be ready for next usage.

In our model, the codes' length obtained from the key are ≥ 128 bits.

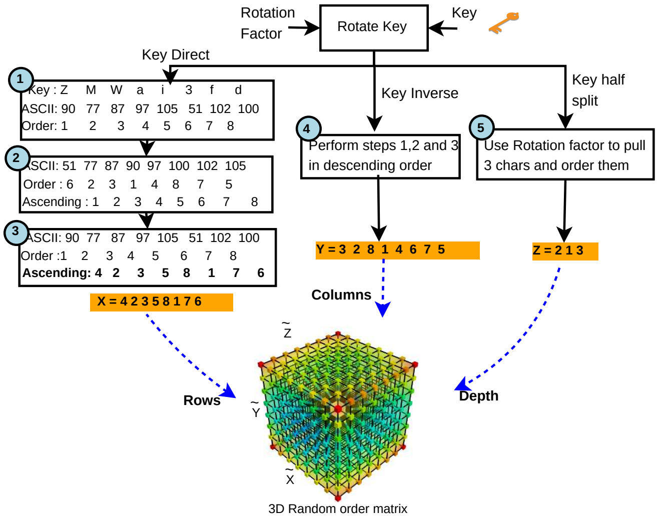


Fig. 7. A paradigm of generating the $3D$ key-driven dynamic order template.

The above steps ensuring that without the right key with its rotation factor, it would be highly infeasible for unlawful parties to accurately retrieve the private bits. The idea of rotating the key dynamically is to generate an extensive number of 3D templates even with the same patient key.

The overall operations of embedding are depicted in Fig 5, 6, 7 and summarised in Fig 8. The obtained coefficients (i.e. from implementing FWHT into the genuine readings) are reformatted to $X \times Y \times Z$ $3D$ matrix. The key codes are then employed to encrypt the patient's sensitive information. Next, the key-driven $3D$ template is produced which will be followed to randomly embed the secret bits.

C. Inverse Walsh-Hadamard Re-transform

After the embedding operations, the obtained values are named stego coefficients. To recompose the original time domain of the signal, the stego form should be (i) re-shuffled from $3D$ to a vector format, and (ii) reversed by applying FWHT re-composition. The outcome is a recomposed stego signal (i.e. hiding private patient data) almost identical to the genuine form. The beauty is that even the stego biomedical signal may be utilized as the genuine form; However, the patient's identity is only possible to be retrieved by legitimate recipients holding both the secret key and rotation factor. The FWHT re-composition is shown by Eq 5

$$a_i = \frac{1}{M}\sum_{1=0}^{M-1} y_n IFWHT(n,i), n = 1,2,...,M-1 \tag{5}$$

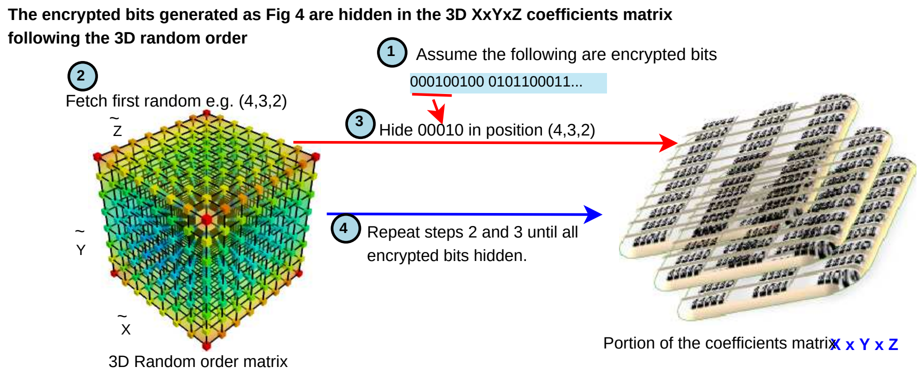

The encrypted bits generated as Fig 4 are hidden in the 3D XxYxZ coefficients matrix following the 3D random order
1 Assume the following are encrypted bits
000100100 0101100011...
2 Fetch first random e.g. (4,3,2)
3 Hide 00010 in position (4,3,2)
4 Repeat steps 2 and 3 until all encrypted bits hidden.
~Z
~Y
~X
3D Random order matrix
Portion of the coefficients matrix X x Y x Z

Fig. 8. A holistic view of the embedding operations combining Figures 5, 6 and 7.

where a_i is the genuine readings, b_n is the obtained coefficient from the re-composition, and $IFWHT(n, i)$ is the inverse transformation.

D. Retrieval of Confidential Information

To precisely retrieve and decrypt the hidden confidential information of the patient, the receiver must have the security key and set the rotation factor. The operations are largely similar to the concealing steps, but the bits will be pulled rather than placed. Fig 9 shows the overall steps. It begins by applying FWHT to the biomedical signal. The key is then used to reshape the FWHT coefficients into a $3D$ matrix and generate the selected coefficients' order. Next, the secret bits' retrieval is started, following to formed $3D$ template. Finally, the retrieved bits are integrated to be decoded for identity verification.

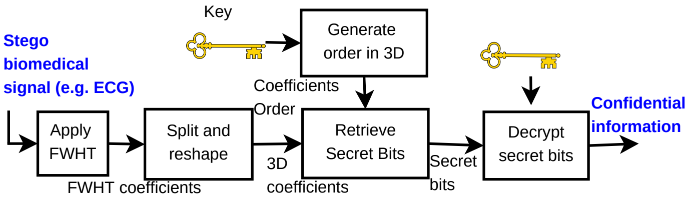


Fig. 9. The overall operations of retrieval process.

IV. Evaluation

This part concentrates on evaluating the introduced model from various angles such as the key solidity, embedding operations, the possible volume of embedded bits and the deformation monitoring.

A. Key Solidity

The strength of the introduced steganographic technique relies on infeasibility of retrieving the embedded bits without obtaining (i) the patient key, (ii) the rotation factor, and (iii) the entire biomedical signal in addition to FWHT decomposition parameters.

However, the most essential factor is the patient key because it is involved in generating three layers of security (See Figs. 5, 6 and 7): (1) encrypt the confidential information, (2) reshape FWHT coefficients into $3D\ X \times Y \times Z$ matrix, and (3) produce a $3D\ \widetilde{X} \times \widetilde{Y} \times \widetilde{Z}$ random order template followed to embed the bits. Hereby, the privacy of this key has to be highly ensured by: (a) the sender (i.e. remote PoC) where the key should be burned and utilized whenever collected biomedical signals are sent, and (b) the legitimate recipient (e.g. hospitals) which can properly retrieve, read and check the validity of the hidden confidential bits. The rest (e.g. cloud providers) should be only able to obtain stego biomedical signals. In this paper, the key is produced and is preserved by PoC's and receiver's ends. The key strength of our algorithm can be measured by entropy's bits number H (See Eq. 6) where 2^H represents total possible combinations which would need to be tried by unlawful parties in a brute-force attack.

$$H = log_2 P_s^T \tag{6}$$

Where T is the total symbols length and P_s is the probabilities of the symbols. Table II presents an overview of some used keys strength along with their length, set and the achieved combinations. This emphasizes that 256 key along with $UTF-16$ set is a very secure option.

TABLE II
An Overview on Key Solidity Samples

Key Range	Token Set	Combinations
64	US-ASCII	1.27e+101
128	US-ASCII	1.63e+202
256	US-ASCII	4.23e+253
64	UTF-8	5.43e+120
128	UTF-8	2.96e+241
256	UTF-8	5.34e+277
64	UTF-16	7.67e+274
128	UTF-16	3.45e+301
256	UTF-16	9.21e+341

B. Illegitimate Retrieval

To protect the hidden sensitive information from exhaustive key search extraction, the reshaped $3D\ X \times Y \times Z$ template

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBME.2016.2631885, IEEE Transactions on Biomedical Engineering

obtained from FWHT transformation should be in acceptable volume (i.e. > Key range) (See Eq 7).

$$T_c = \sum_{i=t}^{r} R! \times \sum_{j=t}^{c} C! \times \sum_{k=t}^{d} D! \times N^L \tag{7}$$

where T_c is the total possible combinations, R, C and D are the reshaped $3D$ values template and t representing the lowest possible selection from every dimension.

For example, assume a collected biomedical signal of 10-second length (e.g. 8200 readings), and reformatted to $3D$ coefficients template of volume $32 \times 16 \times 16$ next to implementing FWHT. The presumed t is $16 \times 8 \times 8$, the token set is $UTF - 16$ and its key range is 128 (See Eq 8).

$$T_c = \sum_{i=1}^{32} 32! \times \sum_{j=16}^{32} 32! \times \sum_{k=16}^{32} 32! \times 65536^{128} \tag{8}$$

This proves that accurately revealing the hidden patient identity is highly impossible.

C. Hidden information Size

The maximum volume can be embedded essentially based on: (i) the maximum samples of transferred signal and (2) the stego bits level that can possibly be embedded in every FWHT value (See Eq 9).

$$b = \sum_{i=1}^{n} ((R \times C \times D) - h_c) \times S \tag{9}$$

Where b represents highest embedded bits, n is the biomedical signal's samples, R, C and D are the dimensions of the reshaped $3D$ template after implementing FWHT transformation, h_c represents the less-significant sequence coefficients and S represents the stego scale per value.

For clarity, suppose that FWHT transformation is applied to normal collected biomedical signal and the size of reshaped $3D$ matrix is $512 \times 32 \times 32$ (i.e. values of R, C and D). Also, assume the high sequence coefficients is $\leq$200 (i.e. value of h_c) and the value of S is five bits per coefficient. Consequently, around (328 Kilobyte) of patient's private information are possibly embedded within these samples.

D. Stego Effectiveness

To accurately measure the impact of our model on the transferred biomedical signals (e.g. ECG, EEG and PPG) and to confirm the setgo validity, the variance forms of the signal before and after the stego have been thoroughly screened employing a percent of root-mean-square difference (PRD). This matrix is widely-used for its ability to detect any variation between two signals as given in Eq. 10 [21].

$$PRD = \sqrt{\frac{\sum_{i=1}^{L}(\upsilon(i) - \widetilde{\upsilon}(i))^2}{\sum_{i=1}^{L}(\upsilon^2(i))}} \times 100\% \tag{10}$$

where $\upsilon(i)$ and $\widetilde{\upsilon(i)}$ are the genuine and recomposed signals, and L represents samples length.

Identically, the distortion obtained from extraction operations also monitored. This is achieved by calculating PRD variance between genuine and extracted form of the signal. All outcomes are introduced in Section V.

V. Implementations

A. Datasets

To test the effectiveness of our algorithm with different types of biomedical signals, three datasets have been used in our experiments. (1) Electrocardiograms (ECG) and Photoplethysmogram (PPG) signals dataset collected and published in ”PhysioNet” repository, which is funded by the National Health Institute (U.S) [32]. It contains extensive periodical readings over many years for different biomedical signals (e.g. ECG and PPG). (2) Electroencephalography (EEG) signals dataset gathered by Henri Begleiter at the Neurodynamics Laboratory at the University State of Brooklyn New York Health Center and published by University of California [33]. It offers detailed continuous EEG monitoring for a large number of people.

B. Experiments and Results

In this work, all the above sorts of samples were thoroughly utilized to prove the accuracy of performing the introduced model on a wide range of biomedical signals. The tests were applied by embedding and retrieving patient's private data according to our model steps explained in Sections III-B and III-D. The private data represents a group of sensitive information that has to be preserved such as personal data (e.g. patient ID, name, DoB, geometric location), biometric data (e.g. iris and fingerprint) and diagnosis data (e.g. temperature, blood pressure and glucose level) which all turned into bits for embedding within transferred biomedical signals.

The categorisation of our tests can be as follows. (1) Hiding, which is done by remote PoCs to embed the patients' private data in their gathered signals as explained in Section III-B. (2) Sensitive information retrieval that is at the receivers' end (e.g. hospitals) as explained in III-D. This resulting in that any intercepting or brutal altering of the stego signals by unlawful parties will (i) not be able to reveal any private bits, and (ii) be easily verified.

To obtain unbiased outcomes, our introduced model was tested with a wide range of key lengths along with many biomedical signals of sizes like 512, 1024, 2048, 4096 and more. To achieve the highest deformation, all less-significant FWHT values (i.e. around 80% of the total coefficients) have been used. For this work brevity, the important contribution outcomes summarized as follows. (1) Fig. 10 shows an example of three original biomedical signals (ECG, EEG and PPG) used to hide patient's confidential data along with the stego and extracted signals. (2) Tables III, IV and V show the exact PRD results from all aforementioned biomedical signals between the genuine vs stego shapes as well as the normal vs retrieved shapes.

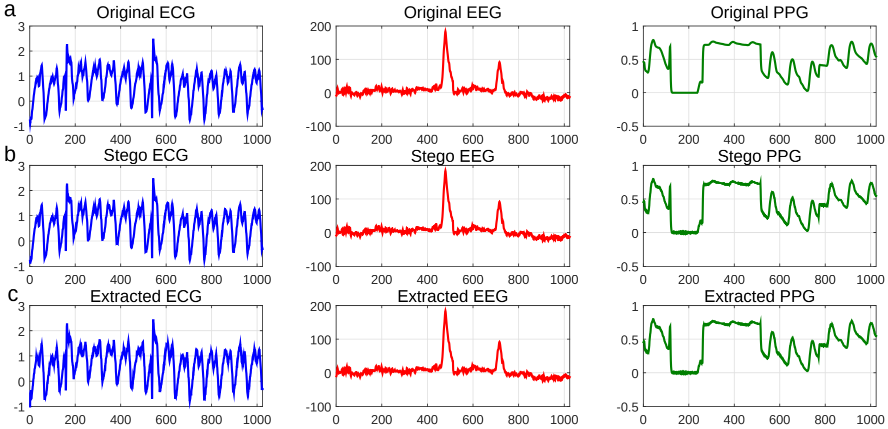


Fig. 10. Three examples of biomedical signals: (a) original samples (b) stego shape hiding the patient sensitive bits and (c) retrieved signal after extraction.

C. Discussion

In all conditions, in spite the variant sample ranges of the biomedical signals and different characteristics and values, all PRDs are $< 1\%$. This emphasizes that our introduced model has a little constant impact on the genuine transmitted biomedical signals. In contrary, it offers a robust solution for protecting the privacy of transmitted patient's confidential information as well as the authenticity of periodically collected signals. The advantages of this solution are as previously stated. (1) There are strong privacy preservation and authenticity where the embedded private identity can only be retrieved and verified by legitimate recipients (e.g. hospitals), but the rest can only access and use the stego signal form. (2) There is no overhead in the real transmitted signals volume. (3) There is no noticeable manipulation to the original signal's form that helps the legitimate receivers to directly exploit off-site services (i.e. cloud) without revealing patient's confidential data. In other words, all mathematical and diagnosis operations can be directly applied to the transmitted stego form of signals even at intermediate hopes and cloud while maintaining the privacy and the authenticity.

TABLE III
PRD RESULTS FOR EEG DATA SET READINGS

	512 samples EEG		1024 samples EEG	
Segment No	PRD % Stego	PRD % Extracted	PRD % Stego	PRD % Extracted
1	0.0257	0.1056	0.0245	0.1757
2	0.0477	0.1952	0.0305	0.1997
3	0.0933	0.4179	0.0402	0.2896
4	0.0215	0.0983	0.0858	0.5975
5	0.0299	0.1631	0.0332	0.1810
6	0.0625	0.2924	0.0460	0.2975
7	0.0714	0.3506	0.0704	0.5178
8	0.0567	0.2846	0.0976	0.5440
9	0.1114	0.4276	0.0675	0.4721
10	0.0513	0.1655	0.0906	0.5684
11	0.0540	0.2310	0.0562	0.4036
12	0.0540	0.2310	0.0632	0.3681
13	0.0697	0.3213	0.0632	0.3680
14	0.0567	0.2520	0.0849	0.5276

TABLE IV
PRD RESULTS FOR ECG DATA SET READINGS

	512 samples ECG		1024 samples ECG	
Segment No	PRD % Stego	PRD % Extracted	PRD % Stego	PRD % Extracted
1	0.4978	0.8840	0.0513	0.1710
2	0.9661	0.1852	0.0857	0.1271
3	0.9643	0.0572	0.5310	0.6833
4	0.2302	0.3426	0.9463	0.8580
5	0.0364	0.6458	0.5453	0.5873
6	0.8963	0.8120	0.2375	0.5073
7	0.7411	0.5518	0.7632	0.5023
8	0.6836	0.8223	0.9488	0.7442
9	0.5345	0.6071	0.9405	0.7350
10	0.7744	0.1155	0.9331	0.2011
11	0.7986	0.4063	0.1894	0.5958
12	0.5334	0.3519	0.0318	0.0342
13	0.9345	0.6071	0.9117	0.7723
14	0.6072	0.7953	0.2267	0.3055

TABLE V
PRD RESULTS FOR PPG DATA SET READINGS

	512 samples PPG		1024 samples PPG	
Segment No	PRD % Stego	PRD % Extracted	PRD % Stego	PRD % Extracted
1	0.4544	0.2595	0.0404	0.7343
2	0.6934	0.5405	0.8482	0.1833
3	0.7405	0.0453	0.7469	0.7005
4	0.7767	0.2447	0.5891	0.0523
5	0.5680	0.6192	0.7677	0.6313
6	0.4002	0.7714	0.4271	0.3396
7	0.6855	0.2165	0.7851	0.9262
8	0.2179	0.6260	0.0574	0.7311
9	0.7998	0.4791	0.0275	0.9694
10	0.4535	0.6849	0.3306	0.1428
11	0.9688	0.0857	0.6948	0.3534
12	0.5542	0.6500	0.7085	0.3554
13	0.2996	0.4096	0.1598	0.6140
14	0.6693	0.5232	0.4327	0.6156

D. Comparison With Current Work

To the best of our knowledge, steganography with Walsh-Hadamard transformation has never been used before in biomedical signals field. Nevertheless, there is a recent ad-

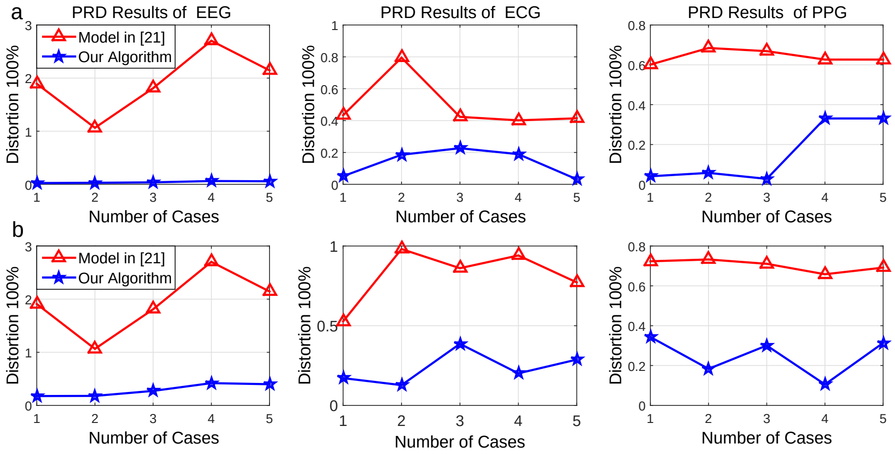


Fig. 11. Comparison of various obtained distortion outcomes of 1024 readings of biomedical signals (EEG, ECG and PPG) between our technique vs the algorithm in [21]. (a) PRDs for normal signal vs and Stego form, (b) PRDs between the normal vs retrieved form.

vance introduced model [21] that has a close approach with unlike transformation technique. Therefore, our algorithm is assessed with the work in [21] where the authors introduced a steganographic model to embed patients' data in their gathered ECG samples utilizing wavelet transformation. There are three main improvements in our algorithm:

1) After testing both techniques using various types of biomedical signals samples (See Fig 11), the observation was that our technique has much lower and constant deformation than the algorithm in [21]. This is because (i) their algorithm has been designed and experimented only on ECG signals, whereas our algorithm is a generic model which is designed and tested with various biomedical signals (e.g. ECG, EEG and PPG), and (ii) only less important coefficients are utilized in the embedding operations in our algorithm, but all tree sub-bands values (coefficients) are used in their model.
2) Despite the advantages of wavelet transformation that is used in the model [21] to hide more data, producing the multidimensional sub-bands wavelets tree is a relatively expensive process regarding both time (i.e. quadratic complexity) and operations (i.e. based on multiplications) [27]. On the other hand, our algorithm relies on a much lighter and fast transformation technique (i.e. Fast Walsh-Hadamard) in terms of time (i.e. linearithmic complexity $n \log n$) and operations (i.e. based on additions and subtractions) [17]. Also, from the experimental viewpoint, Fig. 12 clearly shows that our algorithm needs much less time to complete the transformation and the hiding processes than their model in [21].
3) Both techniques use a security key to cipher the confidential information; However, our algorithm is stronger in terms of generating and managing the random hiding order matrix. This is because in our model, this $3D$ matrix is dynamically generated on the fly using the key, whereas in the model [21] the hiding order is statically stored in a form of 2D matrix of size 128×32 which obviously consumes the storage space in addition to the security risk of storing and managing this matrix.

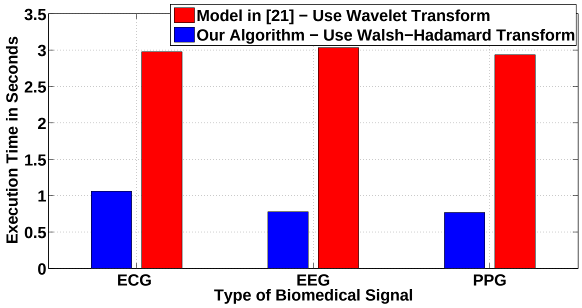


Fig. 12. The required time to hide patient's sensitive information in their collected biomedical signals (i.e. ECG, EEG and PPG) of length 10 seconds.

VI. Conclusion

In this work, a new robust 3D steganographic based Walsh-Hadamard algorithm has been introduced to preserve patients' private data in PoC systems by key-driven random dynamic embedding inside transferred biomedical signals on a bit level. This ensures (1) a robust end-to-end privacy preservation for confidential information, and (2) solid authenticity evidence for the normal signals. To guarantee the maximum embedding level, FWHT is applied to transform the signals into a frequency based coefficients. To ensure the lowest distortion, only unimportant coefficients are utilized. To strengthen the security, a key is utilized to (i) only cipher the private information, (ii) reformat the FWHT values into a random $3D$ template, and (iii) produce a dynamic hiding sequence as $3D$. The resultant distortion has been thoroughly measured at all stages - the original, the stego, and the extracted forms - using a widely-known measurement called PRD. After extensive experiments on three different types of signals (i.e. ECG, PPG and EEG) it has been proven that our algorithm has little impact on the genuine signals ($< 1\%$). The security assessments also emphasize that unauthorised extraction of the hidden bits within a rational time is extremely infeasible.

References

[1] D. Venkateswarlu et al. e health networking to cater to rural health care and health care for the aged. In *e-Health Networking, Application and Services, 2007 9th International Conference on*, pages 273–276. IEEE, 2007.

[2] L. Yuan-Hsiang et al. A wireless pda-based physiological monitoring system for patient transport. *Information Technology in Biomedicine, IEEE Transactions on*, 8(4):439–447, 2004.

[3] H. Fei et al. Privacy-preserving telecardiology sensor networks: toward a low-cost portable wireless hardware/software codesign. *Information Technology in Biomedicine, IEEE Transactions on*, 11(6):619–627, 2007.

[4] J. Zhou et al. Securing m-healthcare social networks: Challenges, countermeasures and future directions. *Wireless Communications, IEEE*, 20(4), 2013.

[5] Centers For Disease Control, Prevention, et al. Hipaa privacy rule and public health. guidance from cdc and the us department of health and human services. *MMWR: Morbidity and Mortality Weekly Report*, 52(Suppl. 1):1–17, 2003.

[6] Colin Thomson. The regulation of health information privacy in australia. www.nhmrc.gov.au/_files_nhmrc/publications/attachments/nh53.pdf, 2004.

[7] L. Wei-Bin and L. Chien-Ding. A cryptographic key management solution for hipaa privacy/security regulations. *Information Technology in Biomedicine, IEEE Transactions on*, 12(1):34–41, 2008.

[8] H. Wang et al. Resource-aware secure ecg healthcare monitoring through body sensor networks. *Wireless Communications, IEEE*, 17(1):12–19, 2010.

[9] D. Algarin et al. A security framework for xml schemas and documents for healthcare. In *Bioinformatics and Biomedicine Workshops (BIBMW), 2012 IEEE International Conference on*, pages 782–789. IEEE, 2012.

[10] M. Li et al. Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. *Parallel and Distributed Systems, IEEE Transactions on*, 24(1):131–143, 2013.

[11] R. Lu et al. Spoc: A secure and privacy-preserving opportunistic computing framework for mobile-healthcare emergency. *Parallel and Distributed Systems, IEEE Transactions on*, 24(3):614–624, 2013.

[12] S. Katzenbeisser and M Petkovic. Privacy-preserving recommendation systems for consumer healthcare services. In *Availability, Reliability and Security, 2008. ARES 08. Third International Conference on*, pages 889–895. IEEE, 2008.

[13] H. Perl et al. Fast confidential search for bio-medical data using bloom filters and homomorphic cryptography. In *E-Science (e-Science), 2012 IEEE 8th International Conference on*, pages 1–8. IEEE, 2012.

[14] O. Kocabas et al. Assessment of cloud-based health monitoring using homomorphic encryption. In *ICCD*, pages 443–446, 2013.

[15] A Ikuomola and O Arowolo. Securing patient privacy in e-health cloud using homomorphic encryption and access control.

[16] F. Petitcolas et al. Information hiding-a survey. *Proceedings of the IEEE*, 87(7):1062–1078, 1999.

[17] J. Fino and V. Algazi. Unified matrix treatment of the fast walsh-hadamard transform. *IEEE Transactions on Computers*, 25(11):1142–1146, 1976.

[18] K Zheng and Xu Qian. Reversible data hiding for electrocardiogram signal based on wavelet transforms. In *Computational Intelligence and Security, 2008. CIS'08. International Conference on*, volume 1, pages 295–299. IEEE, 2008.

[19] H Golpira and H Danyali. Reversible blind watermarking for medical images based on wavelet histogram shifting. In *Signal Processing and Information Technology (ISSPIT), 2009 IEEE International Symposium on*, pages 31–36. IEEE, 2009.

[20] S. Kaur et al. Digital watermarking of ecg data for secure wireless commuication. In *Recent Trends in Information, Telecommunication and Computing (ITC), 2010 International Conference on*, pages 140–144. IEEE, 2010.

[21] A. Ibaida and I. Khalil. Wavelet based ecg steganography for protecting patient confidential information in point-of-care systems. *IEEE transactions on bio-medical engineering*, 60:3322–3330, 2013.

[22] A. Cheddad et al. Digital image steganography: Survey and analysis of current methods. *Signal processing*, 90(3):727–752, 2010.

[23] Y. Huang et ali. Steganography integration into a low-bit rate speech codec. *IEEE Transactions on Information Forensics and Security*, 7(6):1865–1875, Dec 2012.

[24] Q. Cheng and T. S. Huang. An additive approach to transform-domain information hiding and optimum detection structure. *IEEE Transactions on Multimedia*, 3(3):273–284, Sep 2001.

[25] A. Abuadbba and I. Khalil. Wavelet based steganographic technique to protect household confidential information and seal the transmitted smart grid readings. *Information Systems (2014).*, 2014.

[26] A. Abuadbba et al. Robust privacy preservation and authenticity of the collected data in cognitive radio networkwalsh–hadamard based steganographic approach. *Pervasive and Mobile Computing (2015).*, 2015.

[27] A Akansu and P Haddad. *Multiresolution signal decomposition: transforms, subbands, and wavelets*. Academic Press, 2000.

[28] H Harmuth. Applications of walsh functions in communications. *Spectrum, IEEE*, 6(11):82–91, 1969.

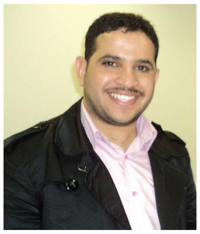
[29] K. Beauchamp et al. *Applications of Walsh and Related Functions: With an Introduction to Sequency Theory*. Academic press New York, 1984.

[30] Tom Beer. Walsh transforms. *American Journal of Physics*, 49(5):466–472, 1981.

[31] David Salomon. *Data compression: the complete reference*. Springer, 2004.

[32] A. Goldberger et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000 (June 13). Circulation Electronic Pages: http://circ.ahajournals.org/cgi/content/full/101/23/e215 PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[33] Henri Begleiter. Neurodynamics laboratory, state university of new york health center. archive.ics.uci.edu/ml/datasets/EEG+Databasel, 1999.



Alsharif Abuadbba is currently working toward the Ph.D. degree in computer science at RMIT University, Melbourne, Australia. He received the Masters degree in computer science from RMIT (2013). Based on the results and achievements during his Master and PhD study, he received several awards at RMIT University: Significant achievement by Golden Key organisation in 2012, Boeing postgraduate scholarship in 2013, Recognition of contribution in 2013, Best master student in 2014 and publication award in 2015. His research interests include big data security, steganography, signal processing, high sensor streams and IoT.

Ibrahim Khalil is an associate professor in School of Computer Science & IT, RMIT University, Melbourne, Australia. Ibrahim obtained his Ph.D. in 2003 from the University of Berne in Switzerland. He has several years of experience in Silicon Valley. He worked for EPFL and University of Berne in Switzerland and Osaka University in Japan. His research interests are in scalable computing in distributed systems, m-health, e-health, wireless and body sensor networks, biomedical signal processing, network security and remote healthcare.