IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 8, 2021, accepted December 9, 2021, date of publication December 23, 2021, date of current version January 5, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3138135

A New Approach for Secure Cloud-Based Electronic Health Record and Its Experimental Testbed

JUSAK JUSAK[1,2], (Member, IEEE), SEEDAHMED S. MAHMOUD[3], (Senior Member, IEEE), ROY LAURENS[4], (Member, IEEE), MUSLEH ALSULAMI[5], AND QIANG FANG[3], (Member, IEEE)

[1]Department of Computer Engineering, Dinamika University, Surabaya 60298, Indonesia
[2]School of Science and Technology, James Cook University, Singapore 387380
[3]Department of Biomedical Engineering, Shantou University, Shantou 515063, China
[4]Department of Computer Science, University of Central Florida, Orlando, FL 32816, USA
[5]Department of Information Systems, Umm Al-Qura University, Makkah 24382, Saudi Arabia

Corresponding author: Seedahmed S. Mahmoud (mahmoud@stu.edu.cn)

This work was supported by the 2020 Li Ka Shing Foundation Cross-Disciplinary Research Grant under Ref: 2020LKSFG04C.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Rector of Dinamika University under Application No. 065/UDK/8.5/X/2021.

ABSTRACT The tremendous development of the Internet of Things (IoT) technology in the last decades has fostered advancement in automatic medical assistive devices to support the existing Electronic Health Record (EHR) system. As an integral part of the EHR electronic model, public cloud servers store medical data. Unfortunately, public cloud servers are prone to security and privacy breach. This paper introduces a novel non-cryptographic approach to preserve electrocardiograph (ECG) data confidentiality and integrity in the EHR environment. The main objective of the proposed anonymization algorithm is to obscure the patient's cardiac information during transmission and to protect information stored in the cloud database. Although we focus on ECG data, generalization to other types of clinical data can be derived using the proposed method. Performance evaluation of the proposed scheme showed that the algorithm conceals both fiducial and non-fiducial features of the data. Therefore, confidentiality feature is preserved. This paper examined confidentiality of the anonymized data using the Percentage Residual Difference (PRD) and investigated the integrity of the reconstructed data in terms of cross-correlation. Security analysis carried out using the PRD, brute force attack, and performance comparison between the proposed algorithm and existing methods. Evaluation showed that the proposed scheme offers a secure non-cryptographic model for transmission and storing clinical data in the cloud. Moreover, in terms of processing time, the proposed algorithm is ten times faster than the existing wavelet packet method when processing long ECG data, 65,536 sample points. In a real-time experimental testbed, the implemented proposed system was successful.

INDEX TERMS Eelectrocardiography, electronic health record, fast Fourier transform, information security.

I. INTRODUCTION

With the advent of Internet of Things (IoT) technology and the development of medical Internet-based applications and devices, we have witnessed remarkable progress over the last two decades [1]–[6]. For example, according to the Global Market Insights, the Electronic Health Records (EHR) market size in 2018 was valued at over 25.5 billion USD around the world and it is forecasted that the market size will become 38 billion USD in 2025 [7].

The proliferation of the current information and communication technology has changed the way medical data recorded and stored. These medical data may include several personal medical data such as: medical histories and medication and laboratory test results. Additionally, it contains vital information such as an electrocardiograph (ECG),

The associate editor coordinating the review of this manuscript and approving it for publication was Ilsun You.

phonocardiograph (PCG), electroencephalogram (EEG), blood pressure, respiration rate, and body temperature (the so-called human vital signs). The vital signs data were previously recorded in paper-based format. The transformation from paper-based electronic documents has gradually become a compulsory requirement to provide efficient and flexible healthcare services. EHRs have several advantages, including enabling data sharing across various stakeholders for future clinical analysis and decision making, easy data access at any time and anywhere, reduced time for locating physical records, reduced physical storage, and less medical costs [8].

Further, the widespread use of cloud computing technology for a broad range of applications is another key factor that has attributed to the fast growth of EHR in the healthcare industry. There are several features that make cloud computing different from traditional computing techniques, including on-demand self-service, broad network access, resource pooling, rapid elasticity, and measure service [9], [10]. These features make cloud services an important part of the deployment of the EHR, since they reduce costs associated with storing and maintaining data, improve accuracy of processing, and allow data to be exchanged among stakeholders.

While the cloud computing paradigm offers tremendous change to the way EHRs are implemented, it also poses particular security and privacy threats [11], [12]. Due to the loss of full control over their clinical data in this cloud network model, users are more concerned about security and privacy. The escalating threats to this cloud system may include spoofing identity, tampering with clinical data, and the disclosure of sensitive information 13]. Moreover, cloud virtualization that enables multiple users to share the same physical storage can lead to a problem of multi-tenancy, whereas malicious users may access data that belongs to other users [9], [17]. In addition to the possibility of external threats, special attention needs to be paid to internal threats from person with authority over data such as database administrators, software engineers, programmers, or probably key managers.

Several attempts have been made to prevent security breaches in the cloud-based EHR system. However, the existing security and privacy techniques do not offer adequate protection for healthcare services [13]–[15]. For example, the well-accepted methods such as attribute-based encryption (ABE), identity-based encryption (IBE), searchable encryption, symmetric key encryption (SKE), were found to be inefficient due to their computational complexity [12]. In contrast, the public key encryption (PKE) techniques have a relatively slow operation due to larger key sizes.

Recently, blockchain technology has emerged as an alternative method for protecting medical records and other sensitive healthcare-related information [16]–[19]. As the number of users of the blockchain grows linearly, it forms a growing distributed ledger of records. Each node in this ledger structure is commonly called a block. Each block is digitally signed by cryptographic hashes and validated by network peers. Due to this chain structure, blockchain can be a secure way of storing medical records [20]. Nevertheless, blockchain is still in its initial phase and lack defined standards [21].

In this paper, we introduce a new non-cryptographic scheme to secure clinical data. Although we focus on electrocardiograph (ECG) data processing, generalization to other types of clinical data such as PCG and EEG data, can be derived using the proposed method. The main objective of the proposed algorithm is to obscure patient's cardiac information during transmission of ECG data and to protect information storing in the cloud database of the EHR environment. ECG data have been used for object evaluation because they embody cardiac health information of a patient and that information is highly unique for every individual [22]. As a result of this discovery, several companies are currently exploring ECG biometric that will be used for personal attribute authentication and plan to deploy it for enterprise applications [23], [24]. The unique features, however, also make ECG data vulnerable to malicious attacks and make it crucial for any service providers that handle clinical data to provide maximum security.

The main contributions of this paper can be summarized as follows:

i We introduce a practical non-cryptographic approach to preserve ECG data confidentiality and integrity in the EHR environment. It should be noted that the approach can convey any types of clinical data, not just ECG data.

ii The ECG data is anonymized using signal processing algorithms, and then stored in the cloud system as a secure data. Therefore, the proposed algorithm conceals the cardiovascular features of a patient fulfilling the *Health Information Protection and Privacy Act* (HIPPA) regulations (1996) [25].

iii We establish a reconstruction algorithm to retrieve the original ECG data from the anonymized ECG.

iv We have implemented the proposed system in a real-time testbed e.g., wireless sensor node and ECG signal processing module to evaluate the practicality and attractiveness of the algorithm to support the existing EHR system.

The proposed algorithm is based on our previous development of the anonymization algorithm in [26] with significant improvement in this paper to enhance the efficiency and robustness of the algorithm. In common medical terms, anonymization refers to a process that removes personal data, e.g., name, address, postcode, so that a data subject can no longer be identified directly or indirectly [27]. In this paper, a similar term, anonymization will be used to describe an algorithm to obfuscate the structure of ECG data so that it cannot be identified without reconstruction. The term was originally used in [28].

The rest of the paper is organized as follows. In the first section, we elaborated on the crucial need for security and privacy in the cloud-based EHR system followed by main contributions of this work. Next, we will explore the

existing anonymization algorithms and evaluate their performance in Section II. Section III describes the detail of the proposed algorithm including the anonymization and reconstruction methods as well as security analysis of the proposed algorithm. Evaluation of the algorithm and its experimental testbed will be carried out in Section IV and finally, conclusions will be drawn in the last section.

II. RELATED WORKS

A. REVIEW ON ANONYMIZATION METHODS

Several studies for securing ECG data with anonymization techniques were widely available in the literature. In this section, we will first discuss two approaches to ECG anonymization. Both algorithms decompose the ECG data using wavelet packet techniques. Later, we will explain the fast Fourier Transform (FFT) algorithm that consumes less energy in real-time implementation.

The authors of [28] suggested using the wavelet packet method to transform the ECG data from the time domain into the frequency domain and decompose the data to subband coefficients. In this method, lower frequency components of the data were then removed and replaced with zeros that corresponded to the distorted coefficients. Subsequently, all coefficients from the ECG data were reconstructed, including those associated with the zeros in the lower subband and those associated with higher frequency coefficients, and transformed back into the time domain to yield anonymized ECG data. Here, the removal of the lower subband coefficients in the frequency domain distorted the structure of the ECG data. Thus, the anonymized ECG data strongly differed from the original. It was done purposefully in order to deform the time domain structure of the original ECG data. Furthermore, for distribution of the data across the networks, the camouflaged ECGs were sent over the Internet to a certain medical center database, while the low frequency subband was sent to an authorized person as a secret key. Finally, the original ECG was recovered by combining the secret key and the distorted ECG data on the receiver side using the reconstruction method. Careful examination in [29] showed that the method in [28] does not fully conceal the fiducial features of the ECG. The reason is that the RR- interval is still present in the anonymized ECG data and can still be identified easily; consequently, heart rate variability of a patient can be revealed using this anonymized data.

A short review about the fiducial and non-fiducial features of ECG data is as follows. The fiducial-based feature representation exploits the characteristic points on the ECG data to reveal amplitude, distance, envelope, slope, time/interval, and area features. The characteristic points are referring to the locations of the peaks and boundaries of P, Q, R, S, and T waves on the ECG waveform. In contrast, the non-fiducial-based feature representation extracts the distinctive information within the ECG by way of the autocorrelation coefficient, Fourier coefficient, and wavelet coefficients.

As a result of the anonymization, it was shown in [28] that the waveforms of the anonymized ECG data were very similar to the original ECG data. Therefore, malicious users might be able to recognize the anonymized data without any difficulties and use it for their purpose. Fortunately, despite the drawback conceived in the algorithm, it showed some significant results. For example, the examination in the paper revealed that the size of the secret key achieved only 5.8% of the original ECG data size. Secondly, the algorithm established the use of compression and encryption techniques to secure the secret key before distribution.

Subsequently, an algorithm that was introduced in [29] showed significant performance improvement over the previous work by using a slightly different approach. The security algorithm made use of the generalized wavelet packet method to decompose the ECG data in several subbands encompassing low-frequency components to high-frequency components. In contrast to the work in [28], the paper claimed that the proposed ECG anonymization method could successfully obfuscate intrinsic features such as fiducial and non-fiducial features. The algorithm had thoroughly examined over normal and abnormal ECG data.

The algorithm in [29] removed the lower subband of ECG data points and treated them as a secret key. This secret key was distributed separately to a medical center server whereas the anonymized ECG data was transformed to time domain and stored in a public server. At the receiver end, only authorized persons with the ability to access the secret key and the reversible function would have the authority to reconstruct the original ECG from the anonymized ECG. The performance of the proposed framework had been evaluated using fiducial features such as the cross-correlation analysis, power spectral density, and the percentage residual difference (PRD). The paper argued that the reconstructed ECG data was highly correlated with the original one. In other words, the proposed method has successfully achieved a lossless reconstruction of the ECG data and proved its robustness. However, the real-time practicality of both works in [28] and [29] have not been examined and tested in a hardware testbed.

The use of the wavelet transforms to manipulate the ECG data as in the described papers leads to the increment of overall computation complexity. This is mainly due to the ECG data decomposition and reconstruction processes with the help of the wavelet-packet algorithm that consumes almost 90% of all anonymization processing time. A study by way of computer simulation in [26] showed that replacing the wavelet decomposition and reconstruction procedures using the fast Fourier transform (FFT) method could achieve the speed of processing 5 times faster than the wavelet-packet based algorithm. For example, this study showed that the proposed method could anonymize ECG data with a length of 16,384 points in 3 ms only, while in contrast to that, the wavelet-packet transform as in [29] required approximately 33 ms to complete the whole anonymization process using the same evaluated ECG data. Therefore, the proposed ECG security method in [26] could be considered as the most suitable algorithm (compared to the existing ones) for implementation of the whole set of systems in the IoT environment,

where some constraints like power and computation limitations could be substantial factors.

B. EVALUATION ON THE EXISTING METHODS

In this sub-section, the anonymization of ECG data based on the FFT described in [26] will be investigated thoroughly. The investigation mainly focuses on the weaknesses of the algorithm in order to find an appropriate solution that will mitigate the vulnerability that exists.

Assume a time-domain ECG data sequence, $x[n]$, represented by a mathematical form $\{x[n]: n = 1, 2, \ldots, N\}$, where N is the ECG sequence length. This ECG sequence is then transformed into its frequency domain using the fast Fourier transform (FFT) algorithm. The mathematical form of the frequency domain signal is expressed as $\{X[k]: k = 1, 2, \ldots, N-1\}$. The frequency domain form is written in capital letters as is typical of many literary works. Fig. 1 depicts the block diagram of the anonymization algorithm developed in [26].

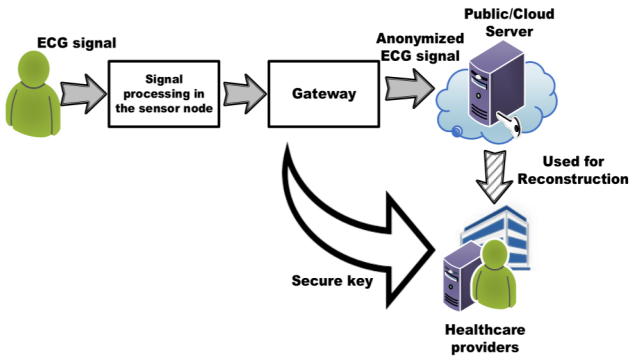

FIGURE 1. ECG data anonymization approach [26], [29].

The most important part of this method is the partition of the frequency domain ECG signal into two subbands, i.e., the low-frequency subband and the high-frequency subband as described in (1),

$$X[k] = \{X_1[0, \cdots, P], X_2[P+1, \cdots, Q]\}. \quad (1)$$

In (1), $X_1[0, \ldots, P]$ represents the low-frequency component of the ECG data and $X_2[P+1, \ldots, Q]$ represents the high-frequency component.

The structure of the ECG data was then modified using the two steps. Firstly, the algorithm removed $X_1[0, \ldots, P]$ from $X[k]$ and called it a secret key, $\kappa[k]$. This secret key was subsequently encrypted, compressed, and sent as a secure key to an authorized healthcare provider database (see Fig. 2). Secondly, the algorithm manipulated the high-frequency component, $X_2[P+1, \ldots, Q]$ with a reversible function, $\Omega[k]$ according to (2),

$$\overline{X}_2[k] = \{X_2[k] * \Omega[k]: P+1, \cdots, Q\} \quad (2)$$

where $X_2[k]$ and $\Omega[k]$ are an element-wise multiplication and $\Omega[k]$ is a vector of

$$\Omega[k] = \{\kappa[k] + offset: k = 0, \cdots, P\} \quad (3)$$

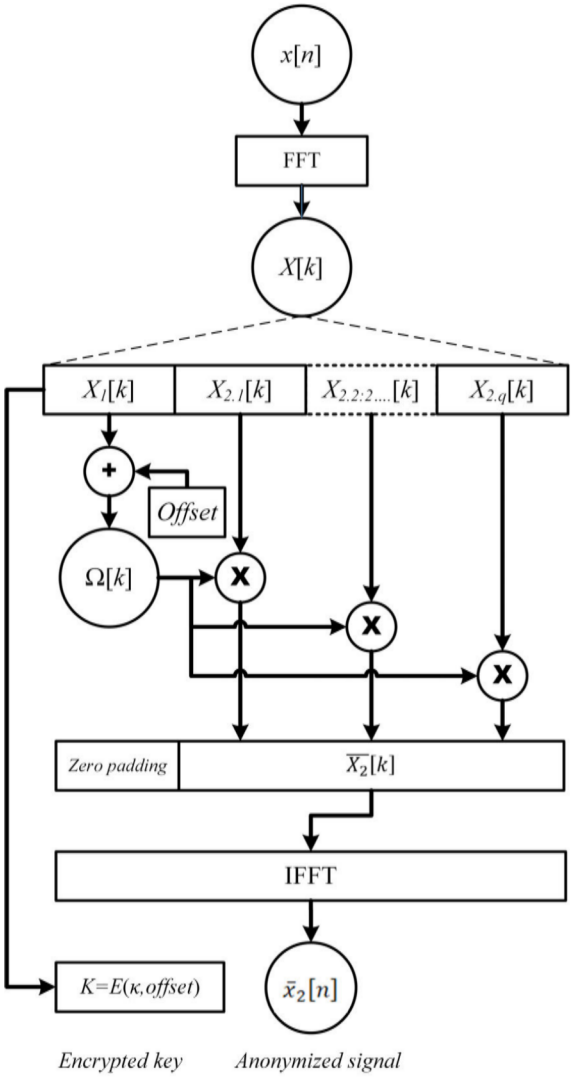

FIGURE 2. Block diagram of the anonymization algorithm in [26].

where the *offset* value was defined as $offset = |\min(\kappa)| + \eta$ and element η is a constant value to prevent division by zero. The anonymized signal, $\bar{x}_2[n]$, was finally produced after taking inverse FFT (IFFT) of the $\overline{X}_2[k]$.

It should be noted that the element-wise multiplication in (2) can be replaced by an element-wise division. This is because both multiplication and division inherit the same computational complexity notated by $O(n^2)$.

In its implementation, the anonymized signal, $\bar{x}_2[n]$, shown in Fig. 2 was transmitted and stored in a public cloud database whereas K was stored separately in the healthcare provider server. Therefore, a potential attacker who is somehow able to get access to this public database can only see the anonymized ECG data, not the original data. To retrieve the original ECG data on the receiver side, an authorized medical doctor/specialist performed a reconstruction method using the secret key taken from a healthcare provider's server and the anonymized signal taken from a public cloud database.

Careful examination of (2) shows that the element-wise multiplication of vector $X_2[k]$ by a vector $\Omega[k]$ requires both vectors to have the same size. However, we see in (3) that the size of vector $\Omega[k]$ is in fact smaller than the size of vector $X_2[k]$. Therefore, the vector $\Omega[k]$ is mandatorily repeated several times until it has the same size as $X_2[k]$ to maintain the valid operation of (2).

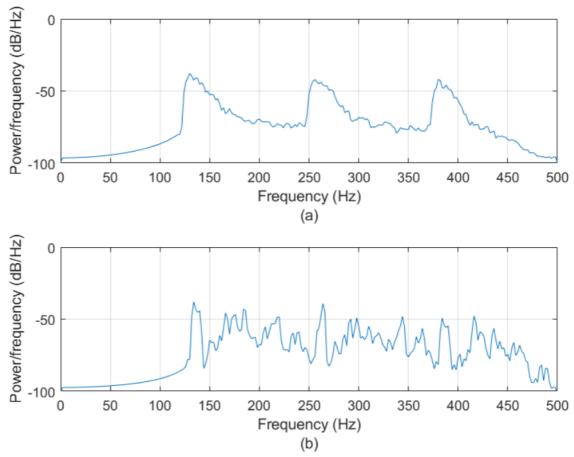

FIGURE 3. Power Spectral Density (PSD) of the anonymized ECG: (a) element-wise multiplication, (b) element-wise division.

Our further studies showed that the modification of the $X_2[k]$ by way of element-wise multiplication or division of the repetitive vector $\Omega[k]$ has eventually produced an obvious vulnerability for the anonymized ECG data itself. This defect is clearly shown in Fig. 3 whereas the power spectral density peaks appear every 125 Hz for both element-wise multiplication and element-wise division operations.

Fig. 3 shows the frequency response in terms of power spectral density of the anonymized ECG data (i.e., patient245, signal s0474 from PTB database) with sampling frequency $f_s = 1,000$Hz. The frequency response of the ECG in Fig. 3a has been modified using an element-wise multiplication of vector $X_2[k]$ and vector $\Omega[k]$ as shown in (2). On the other hand, Fig. 3b depicts the frequency response of the anonymized ECG data that has been modified using element-wise division. The two figures obviously show repetition of the magnitude peak and subband of the frequency response every 125 Hz period. Although signal repetition in Fig. 3b is more subtle than the one in Fig. 3a., careful examination shows that there is a clear individual peak started at every 125 Hz span (i.e., 125 Hz, 250 Hz, 375 Hz) followed by several ripples.

Frequency domain recurrence in Fig. 3 directly associates with the repetition of the vector $\Omega[k]$ as a result of the operation in (2) and (3). Therefore, a malicious attacker can potentially detect the structure of vector $\Omega[k]$ and use it to exploit the ECG anonymization due to this vulnerability. An attacker who has access to the secret key, can easily interpret and reconstruct the anonymized ECG data by simply repeating the secret key every 125 Hz period followed by a reverse function operation of the vector $\Omega[k]$ and vector $X_2[k]$, subsequently.

III. THE PROPOSED METHOD

Modification of the existing anonymization algorithm will be presented in this sub-section. Besides removing frequency domain repetition, the new algorithm should comply with sensor devices limitation, such as low energy consumption and low computational processing. Therefore, vector operation and other mathematical operations in the proposed algorithm were set to obey such behavior while at the same time improve the security of data transmission. Fig. 4 displays a block diagram of the proposed method.

A. THE PROPOSED ANONYMIZATION ALGORITHM

Assume ECG data sequence in the form of $\{x[n]: n = 0, \ldots, N-1\}$. In the first step of the algorithm, we apply the FFT to the ECG data sequence to obtain the frequency domain representation that is represented by $\{X[k]: k = 1, 2, \ldots, N-1\}$, where N is the length of the ECG data sequence.

Secondly, the frequency domain segmentation takes place after transformation from time to frequency domain in the first step. At this stage, we split the frequency domain vector, $X[k]$, into two subbands, i.e., $X_1[k]$ and $X_2[k]$. The first subband, $X_1[k]$ represents low-frequency components of the ECG data whereas $X_2[k]$ signifies high-frequency components. Segmentation of this frequency domain ECG data can be seen in (4),

$$X[k] = \{X_1[k], X_2[k]\}. \tag{4}$$

To eliminate the recurrence behavior of the existing algorithm as shown in Fig. 3, we employ individual operation for each segment in the vector $X_2[k]$ without repeating vector $\Omega[k]$. Suppose the segmentation of $X_2[k]$ is represented by (5),

$$X_2[k] = \{X_{2.1}[k], X_{2.2}[k], \cdots, X_{2.r}[k] : r = 1, 2, \cdots, R\} \tag{5}$$

where R is the number of segments or sub-subbands created from $X_2[k]$.

Next, consider $\ddot{X}[k]$ as a Root Mean Square (RMS) of the ECG data computed in the frequency domain and denoted as in (6).

$$\ddot{X}[k] = \sqrt{\sum_k \left| \frac{X[k]}{N} \right|^2} \tag{6}$$

The RMS represents the overall energy level contained in the ECG data across a frequency range and will be used to define the *offset* values in the anonymization process. Utilization of RMS as in (6) will guarantee non-zero values for the *offset* as the total energy is preserved in the signal. Nevertheless, several experiments showed that this

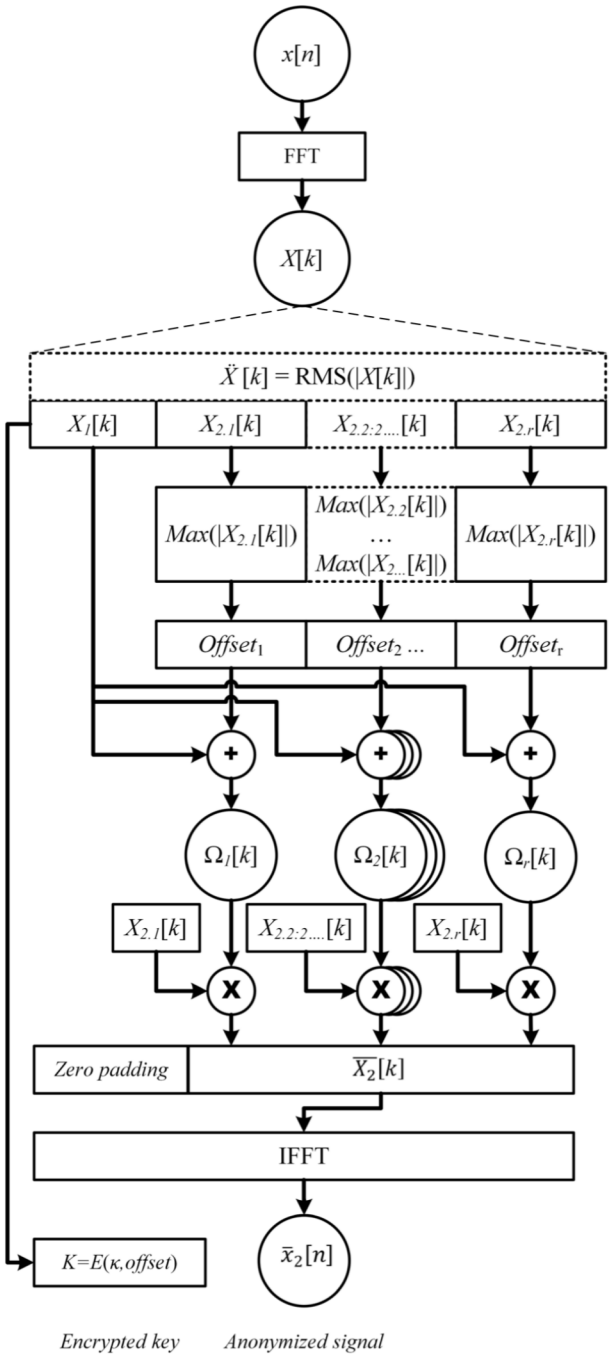

FIGURE 4. Block diagram of the proposed anonymization algorithm.

RMS value requires a scaling operation to obtain acceptable *offset* values for the anonymization process. Therefore, the *offset* values are defined according to (7),

$$offset_r = \frac{\ddot{X}[k]}{\max(|X_{2.r}[k]|) + \eta}: \quad r = 1, 2, \cdots, R \tag{7}$$

The element η in (7) is a constant to prevent division by zero in both the anonymization and reconstruction processes.

Subsequently, the vector $\Omega[k]$ can now be redefined as in (8)

$$\Omega_r[k] = \{\kappa[k] + offset_r : r = 0, \cdots, R\} \tag{8}$$

where $\kappa[k]$ is a secret key which has been defined above as $\kappa[k] = X_1[0, \ldots, P]$.

We define the core of the anonymization process as an element-wise multiplication between $X_{2.r}[k]$ and $\Omega_r[k]$ as shown in (9).

$$\ddot{X}_{2.r}[k] = \{X_{2.r}[k] * \Omega_r[k] : r = 1, 2, \cdots, R\} \tag{9}$$

Element-wise multiplications in (9) serve as a reversible operation in the anonymization procedure. Consequently, it will become a division between $\tilde{X}_{2.r}$ and $\Omega_r[k]$ in the reconstruction process to retrieve the original data back.

The individual segment multiplication in (9) results in the modification of each subband of the ECG data where the secret key performs as a modifier. After the multiplication of each segment, the process continues to the next step that merges each segment into one single vector, i.e. $\overline{X}_2[k]$. It becomes:

$$\overline{X}_2[k] = \left\{\overline{X}_{2.1}[k], \overline{X}_{2.2}[k], \cdots, \overline{X}_{2.r}[k] : r = 1, 2, \cdots, R\right\} \tag{10}$$

In the final step, the inverse fast Fourier transform (IFFT) algorithm performs its action to transform the modified frequency domain ECG data, $\overline{X}_2[k]$, into its time-domain data as in (11). The result is the anonymized ECG data,

$$\bar{x}_2[k] = IFFT\left(\overline{X}_2[k]\right). \tag{11}$$

Fig. 4 shows that the proposed algorithm produces two sets of vectors, i.e., the secret key, $\kappa[k]$ stacked with the *offset*, and the anonymized data, $\bar{x}_2[n]$. It is advisable that for security purposes the two vectors should be stored in different locations/cloud systems. Furthermore, the key security will be achieved by encrypting the secret key and the *offset* vectors according to (12),

$$K = E(\kappa, \Phi), \tag{12}$$

where E is the encryption operator and Φ is a collection of offset defined in (13),

$$\Phi = \{offset_1, offset_2, \cdots, offset_r : r = 1, 2, \cdots, R\}. \tag{13}$$

After encryption, the K can be securely distributed to the healthcare providers' storage while at the same time the anonymized data stored in the public cloud server. See Fig. 1 for clarity. Both data are stored with a unique identification number and specific metadata for each ECG data.

B. THE PROPOSED RECONSTRUCTION ALGORITHM

An expert or doctor resided in a particular healthcare provider needs to employ a reconstruction algorithm to retrieve the original ECG data based on the anonymized ECG stored in the digital storage. Here we elaborate on the reconstruction algorithm as a backward procedure for the previous anonymization algorithm.

On the first step, the device on an expert or a doctor side requires to download both the encrypted secret key and Φ from the healthcare provider storage and the anonymized

ECG from the public cloud server. It is then followed by decryption of the secret key and Φ according to (14),

$$(\kappa, \Phi) = D(K), \tag{14}$$

where $D(.)$ is the decryption operator.

Secondly, the algorithm transforms the anonymized ECG, $\overline{x}_2[n]$, from the time domain to frequency domain representation using the well-known FFT algorithm. As a result, we get the segment $\overline{X}_2[k]$ that contains the modified frequency domain ECG as defined in (10).

In order to retrieve the high frequency subband ECG data, $X_2[k]$, it is compulsory to do a reverse operation of $\overline{X}_2[k]$ shown in (9) towards the vector $\Omega_r[k]$. Therefore, when vector operation in (9) performs element-wise multiplication, the reconstruction algorithm should enforce element-wise division as shown in (15). Conversely, when vector operation in (9) takes element-wise division, the reconstruction algorithm in (15) should impose an element-wise multiplication operation.

$$X_{2.r}[k] = \left\{ \frac{\overline{X}_{2.1}[k]}{\Omega_r[k]} : r = 1, 2, \cdots, R \right\} \tag{15}$$

The last procedure in the reconstruction algorithm is combining the reconstructed vector $X_2[k]$ and the secret key $\kappa[k] = X_1[k]$ to become $\tilde{X}[k]$ as shown in (16),

$$\tilde{X} = \{X_1[k], X_2[k]\} \tag{16}$$

Subsequently, an inverse FFT operation should be applied to (16) to obtain the estimation of the time domain original ECG data, $\tilde{x}[n]$.

C. SECURITY ANALYSIS

Within the EHR system, we include security aspects that emphasize confidentiality and integrity in our proposed algorithm to ensure that the ECG data is secure throughout the cycle. Confidentiality ensures protection of the ECG data from being exposed to unauthorized individuals or parties whereas integrity refers to protection of the data from being modified intentionally or unintentionally.

The confidentiality aspect is shown by relationship between the original and the anonymized data in terms of the Percentage Residual Difference (PRD). The PRD has been used by several papers [26], [29] to quantify the distinction between the original ECG data and the anonymized ECG data. It is defined according to

$$\text{PRD} = \sqrt{\frac{\sum_{i=1}^{N} (x[i] - \overline{x}_2[i])^2}{\sum_{i-1}^{N} x^2[i]}}, \tag{17}$$

where $x[i]$ denotes the original ECG data, $\overline{x}_2[i]$ is the anonymized ECG data, and $i = 1, \ldots, N, N$ is the total number of samples in the ECG data. A careful examination of formula (17) will reveal that PRD $= 0$ when the two time series are identical. Therefore, PRD > 0 measures the degree of distortion among the two data sequences.

The PRD of the ECG data with identification number bs10089603 as in Fig. 6 for different values of η is shown in Table 1. Our experiment shows that the PRDs are larger than 0 for all η. Based on this table, it can be interpreted that the anonymized ECG is significantly different from the original ECG data. The malicious attackers who forcefully gain access to this anonymized data from the public cloud servers will find it nearly impossible to unravel the fiducial features of the ECG data. According to this analysis, the proposed algorithm protects the confidentiality of the data well.

TABLE 1. PRD of ECG data (bs10089603) for variation of η.

Value of Parameter η	PRD
$\eta = 0.01$	28.82
$\eta = 0.1$	3.10
$\eta = 0.3$	1.48
$\eta = 0.7$	1.18
$\eta = 1.0$	1.13

Despite confidentiality, brute force attacks are commonly used to compromise data. The current strong computing power enables this attack to guess passwords or encryption keys by guessing combinations of characters.

In our proposed scheme, the secret key, $\kappa[k]$, comprises of analog numbers with certain length. For example, a key length of 1024 contains 1024 floating-point numbers representing the $X_1[k]$ vector as in (4). Computers normally comply with the IEEE 754 standard to adequately store this secret key into the single precision (32 bit) or double precision (64 bit) digital form. Consequently, for the case of the single precision format and a key length of 1024, a secret key in our model can be represented by 32.768 bits length in the IEEE 754 standard. Therefore, it may take $2^{32.768}$ combinations to brute force the secret key.

We examine processing time for brute forcing the secret key with a key length of 1024. However, due to limitation in our computer, we considered a brute force simulation by injecting as many as 10^9 key combinations (it approximately equals to 2^{30} combinations) out of $2^{32.768}$ combinations that ran for several types of Microprocessors, i.e., Intel Core i7 memory 16GB, Intel Core i7 memory 8GB, and Intel Core i5 memory 12 GB. Fig. 5 shows that the slowest processor took approximately 1976.2 s (32.94 minutes), whereas the fastest processor spent approximately 1031.2 s (17.2 minutes) to complete the simulation. Therefore, assuming the processing time grows linearly with the number of trials for brute force attack, it will take approximately $1031.2 \times 2^{32.738}$s for the fastest computer in the study to attempt all combinations. In other words, it is extremely difficult to accomplish this task using today's fastest processor within acceptable time.

To conclude the security analysis of the proposed algorithm, Table 2 presents performance comparison study among the existing methods for electronic healthcare cloud security and the proposed algorithm in terms of their characteristics, strength, and weakness.

IV. RESULTS AND DISCUSSIONS

This section provides an evaluation of the proposed algorithm for securing ECG data both in transmission and storing in the cloud database. The first evaluation includes the algorithm ability to obscure fiducial and non-fiducial features and its ability to maintain ECG data integrity. We then compare the processing time of the algorithm with the similar function of the existing methods, i.e., the wavelet packet-based and the FFT-based algorithms. Secondly, we will elaborate on a real-time implementation of the proposed algorithm in a small device. The hardware implementation of the proposed algorithm serves as an experimental testbed that observes its applicability in the future.

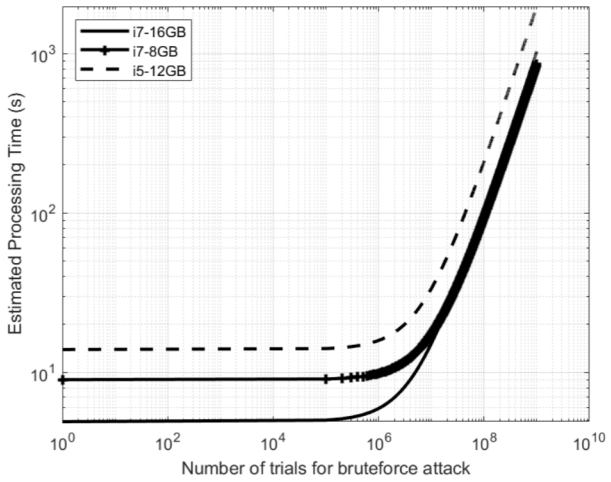

FIGURE 5. Brute force simulation using 10^9 key combinations on several computers powered with Intel Core i7 memory 16GB, Intel Core i7 memory 8GB, and Intel Core i5 memory 12 GB.

A. RESULTS AND ANALYSIS

Simulations and comparisons with existing methods were used to observe the performance of the proposed algorithm. Both time domain representation and frequency domain representation of the ECG data were evaluated in terms of power spectral density (PSD).

The ECG data were collected from a real-time measurement of 32 healthy volunteers with careful supervision from a medical doctor (volunteers' mean age is 23 ± 2 years). The process of capturing ECG signals in this study was a non-invasive process, thus it did not involve any harmful procedure. However, all volunteers consented before the signal was retrieved. Each segment of the ECG data has 10 seconds duration with a sampling frequency of 1000 Hz. Therefore, the number of samples in each segment is 10,000 samples. To ensure the broad application of the proposed algorithm, we also applied the algorithm to anonymize ECG data obtained from the publicly available PTB database [30].

In this evaluation, the size of the secret key as in (1) was set to $P = 1,024$, associated with a frequency range of 0 Hz to 125 Hz. Based on application of (4) to the ECG data, the frequencies between 0 Hz and 125 Hz are removed and retained for use as the secret key. Meanwhile, the frequency components larger than 125 Hz will be anonymized and uploaded to a public cloud server after being transformed into its time-domain data. In the anonymization phase, the ECG data was adjusted using (7) to (9) with a constant value of $\eta = 0.3$. This value was chosen to provide moderate PRD shown in Table 1.

Fig. 6 shows a time-domain representation of original ECG data from a subject with identification number: bs10089603 and its anonymized ECG. It can be seen in Fig. 6b that the algorithm has removed fiducial features of the original ECG wave, i.e., the P, Q, R, S, and T peaks. The peaks have been concealed by high fluctuations of signals.

Fig. 7 (a) and (b) exhibit frequency domain representation in terms of Welch's power spectral density (PSD) estimation for both the original ECG and the anonymized data, respectively. Both diagrams show the PSD in dB/Hz as a function of frequency ranging from 0 Hz to one-half the sampling rate, i.e., 500 Hz. It can be seen in Fig. 6b that the non-fiducial features of the ECG data have been removed utterly by the anonymization process of the proposed algorithm. For example, repetition of the peaks and subbands displayed in Fig. 3 has completely disappeared. Additionally, multiplication in (9) gives an effect on leveraging the magnitude of the PSD of all frequencies above 125Hz. Therefore, modification in the frequency domain is directly associated with the presence of the high-frequency components in the anonymized ECG as shown in Fig. 6b.

Fig. 6b also infers that the original ECG data cannot be interpreted properly after being anonymized. The anonymized data appear to be random and unstructured. If a malicious attacker successfully penetrates a cloud system to acquire the anonymized ECG data, he/she cannot elucidate the structure of the anonymized ECG data without reconstructing it. The anonymized data bears meaningless information in the absence of the reconstruction algorithm and the secret key. Due to the secret key being stored in a different database system, e.g., in the healthcare provider's server, it will be difficult for attackers to access both servers. Therefore, the proposed scheme guarantees a secure and robust model for protecting patient's ECG data.

Fig. 8a shows the reconstructed ECG data and Fig. 8b depicts the cross-correlation between the original ECG data and the reconstructed data. The cross-correlation at lag 0 is 1 that clearly indicates the reconstructed signal is exactly the same as the original ECG. It proves that the proposed reconstruction algorithm explained in III.B works remarkably well to retrieve the original ECG data without losing any information. In this way, the proposed algorithm guarantee integrity of patient's ECG data during transmission and storing in the public cloud database.

Fig. 9 shows the anonymization process of the ECG data when the algorithm was executed using element-wise division in (9). Fig. 9a reveals that the amplitude of the anonymized signal is relatively small due to the element-wise division. However, it can also be noticed that the fiducial structures

TABLE 2. Performance Comparison of Electronic Healthcare Security Methods [8]–[13].

Scheme	Characteristics	Strength	Weakness
Cryptographic Techniques			
Symmetric Key Encryption	• Employing the same shared key for encryption and decryption. • Common algorithms: Advanced Encryption Standard (AES), Data Encryption Standard (DES), RC4, and Blow Fish.	Computationally less intensive, which produces small file size. Therefore, it has advantages such as fast transmission and small storage space.	• Requirement of a shared secret key necessitate the present of key management for distribution and storing keys. Therefore, reliability and security such system is demanding. • The shared key method reduces the anonymity security principle. • Inflexible access control and inability to manage multiple user roles.
Public Key Encryption	• Applying two separate keys for encryption and decryption process comprises of public key and private key. • Common algorithms: RSA (Rivest-Shamir-Adleman) and Elliptic Curve Cryptography (ECC).	• Utilization of unique public and private keys removes unnecessary exchange key between users. • Possibility to establish authentication on sender side by integrating digital signature techniques.	• Computationally more costly than the symmetric key encryption. • Require large key size to provide adequate security. • Prone to man-in-the-middle attack in the communication of public keys. • Requires involvement of Certification Authority (CA) to digitally sign and secure the public keys.
Attribute-Based Encryption	• A type of public key encryption in which decryption of ciphertext depends on predefined set of attributes of the user. • Two approaches widely used in ABE are Key-Policy ABE and Chipertext-Policy ABE.	• The algorithm provides fine-grained and dynamic access control. • The algorithm offers scalable data sharing as access guaranteed only for those recipients that match the predefined attributes.	• Due to its computational complexity, it raises a challenge for implementation ABE in EHR systems. • It inherits complexity in key management and access control policy when the number of attributes grow exponentially.
Non- Cryptographic Techniques			
Access-Control Mechanisms	• Access-control mechanisms enforce policy-based authorization infrastructure to protect privacy by restricting access to the data. • Examples of access-control mechanisms including Mandatory Access Control (MAC), Role-Based Access Control (RBAC), Attribute-Based Access Control (ABAC), and Identity-Based Access Control (IBAC).	• Deployment of access control by using authentication and authorization to restrict access and operation of data.	• These methods require a large number of rules for decision making. They mainly become inefficient and expensive process in a dynamic environment.
The proposed method	• We enforce the proposed anonymization algorithm to obscure sensitive information during transmission and to protect information storing in the cloud database.	• The proposed method exploits both time and frequency domains structure of the data to secure information. This feature reveals significant different with other cryptographic and non-cryptographic techniques. • It does not require large key size to provide adequate security.	• Although it inherits totally different concept of key with others, the proposed system requires a key management repository for distribution and storing keys.

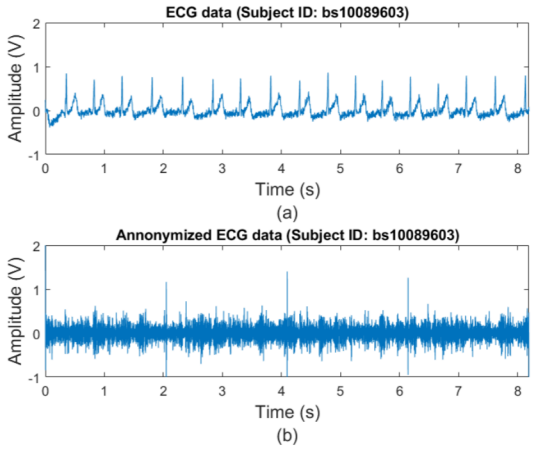

FIGURE 6. Anonymization by employing element-wise multiplication in (9). (a) Original ECG data (Subject ID: bs10089603) and (b) its anonymized data.

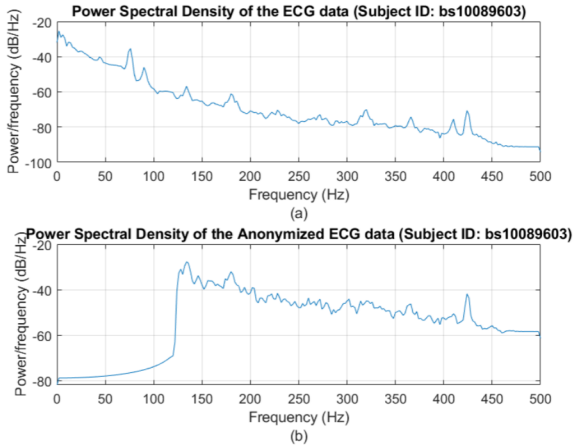

FIGURE 7. Power spectral density (PSD) of the original ECG data (Subject ID: bs10089603) and PSD of its anonymized data.

of the ECG data are successfully disguised by the algorithm. For example, the R-to-R peaks of the ECG data can not be detected from the figure. Fig. 9b shows the frequency response of the anonymized ECG and Fig. 9c shows the reconstructed data. A strong correlation between the original and the reconstructed ECG is indicated by a value of 1 at lag 0 in Fig. 9d. Moreover, a comparison between Fig. 9a and Fig. 6b concludes that element-wise division operation in (9) generally produces anonymized ECG data with lower average amplitude than the operation of element-wise multiplication.

To observe the computational complexity of the algorithm, we examined the processing time between the proposed algorithm and the existing anonymization schemes [26], [29]. We have used the ECG data that were obtained from the PhysioNet PTB Database [30] (i.e., patient180, signal s0475). Fig. 10 draws a comparison graph of the processing time (ms) as a function of ECG data sequence length ($^2\log(N)$) among several anonymization methods. The figure clearly shows that the processing time of the proposed algorithm is comparable

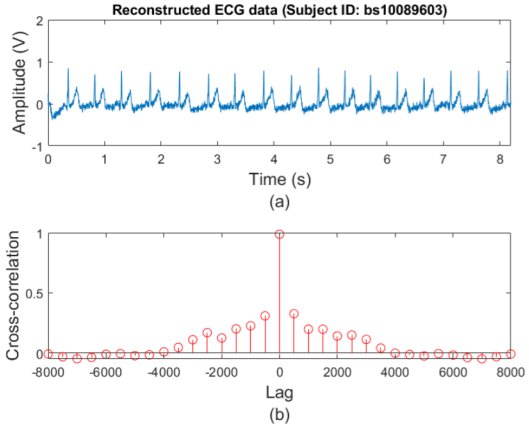

FIGURE 8. Reconstructed ECG data (Subject ID: bs10089603) and the cross-correlation between the original ECG and its reconstructed ECG data.

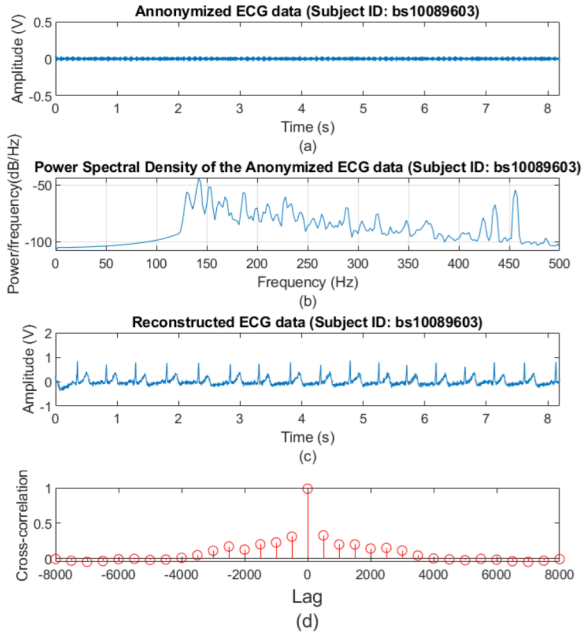

FIGURE 9. Anonymization by employing element-wise division in (9). (a) Anonymized ECG data (Subject ID: bs10089603), (b) PSD of the anonymized data, (c) Reconstructed ECG data, and (d) the cross-correlation between the original ECG data and its reconstructed ECG data.

to the FFT-based anonymization algorithm [26]. However, when ECG data sequence length is increasing up to $N = 2^{16} = 65,536$ data, the processing time of the proposed algorithm is slightly longer than the anonymization algorithm in [26]. This is true because the algorithm requires time to compute the RMS in (5) and the offset values in (7) whereas the FFT-based algorithm does not undergo those processes.

In contrast, the wavelet packet-based anonymization algorithm spends the longest processing time among others. For example, the wavelet-based algorithm has approximately 10 times slower process than the proposed algorithm for processing with 65,536 points of ECG data. Surprisingly, for processing with $N = 2^{12} = 4,096$ data, the proposed algorithm

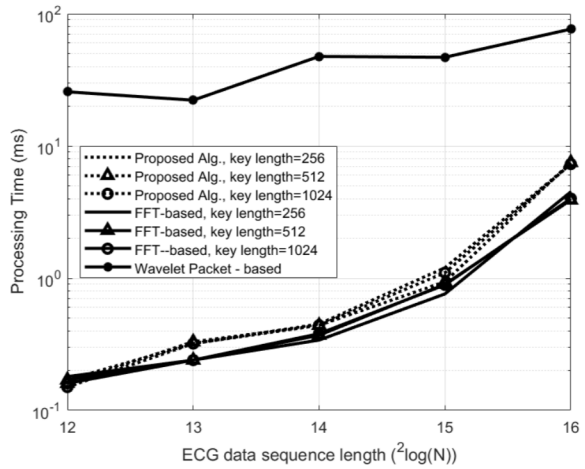

FIGURE 10. Processing time of the proposed algorithm for several key lengths compared to the existing anonymization algorithms including (FFT-based) in [26] and wavelet packet-based algorithm in [29].

TABLE 3. Averaged processing time of the anonymization algorithms.

Algorithm	Averaged Processing Time (ms)				
	$N=$ 4,096	$N=$ 8,192	$N=$ 16,384	$N=$ 32,768	$N=$ 65,536
Wavelet packet –based	25.7	22.2	47.6	46.9	76.6
FFT – based					
Key length=256	0.18	0.24	0.34	0.76	4.5
Key length=512	0.17	0.24	0.37	0.90	3.9
Key length=1024	0.16	0.24	0.38	0.89	4.0
Proposed algorithm					
Key length=256	0.17	0.32	0.45	1.2	7.3
Key length=512	0.16	0.33	0.44	0.94	7.5
Key length=1024	0.15	0.32	0.44	1.1	7.3

*N denotes the ECG sequence length.

performed 150 times faster than the wavelet packet-based algorithm processing time. The detail of the averaged processing time over 1000 trials for all schemes is shown in Table 3.

B. EXPERIMENTAL TESTBED

To conclude this section, we describe the experimental testbed development of the proposed scheme. The ECG data from an observed person were captured using a three-electrode ECG sensor that was directly connected to an integrated single-lead ECG signal conditioning module, i.e., AD8232 as shown in Fig. 11 whereas the circuit diagram connection between the AD8232 and the microcontroller unit is illustrated in Fig. 12. The AD8232 module provides high signal gain ($G = 100$) with dc blocking capabilities. It is also equipped with a two-pole adjustable high pass filter for reducing motion artifacts, a three-pole adjustable low pass filter for eliminating any additional noise, and a right leg drive (RLD) amplifier. The RLD is commonly used to eliminate interference noise from any biological signals. Volunteers were positioned in the seated posture as shown in Fig. 13 with the ECG electrode clamps positioned at the left arm, right arm, and right leg to obtain single-lead ECG data.

Referring to Fig. 12, the output pin of the AD8232 module was connected to the analog input of a microcontroller for analog to digital conversion. After this process, the microcontroller transmitted the data to the ECG security processing module. In this paper, we made use of the Arduino Uno microcontroller board that has an embedded 10-bit analog to digital conversion (ADC) module which converted analog input from the AD8232 directly to digital form. At this point, the single-lead ECG signal could be directly monitored using a display module attached to the board.

The ECG security processor in Fig. 11 performs the anonymization algorithm. In terms of hardware specification, this processor can be anything ranging from single-board computers like a Raspberry Pi or a dedicated computer/server that is able to receive multiple connections from several sensor nodes. Due to this reason, we decided to separate the sensor node and the ECG security processing. The scheme allows the ECG security processor to execute the proposed algorithm for several ECG signals from different sensor nodes. At the same time, it can maintain mobility activities of the sensor nodes when each of the sensor nodes is equipped with wireless connection capabilities like Bluetooth or WiFi. As part of our experiment, we utilized a Raspberry Pi 3 single-board computer with a processor speed of 1.2GHz and Random-Access Memory (RAM) of 1GB, which is more than enough to run the anonymization algorithm. Using Raspberry Pi has another advantage in our case since it has onboard Bluetooth and 802.11 wireless LAN connections, which should support scalability and mobility for the nodes in the future.

The ECG security processor produces the anonymized ECG data as in (11), a sequence of the secret key in (8), and an *offset* vector, Φ, in (13). Because the secret key and the *offset* vector were left unprotected by the algorithm, we instructed the processor to run the AES-256 encryption to encrypt the secret key and the vector Φ altogether.

Finally, the gateway received the anonymized ECG data, the encrypted secret key, and the *offset* vector as shown in the rightmost of Fig. 11. Upon receiving these data, the gateway transmitted the anonymized ECG data to a public cloud server and at the same time transmitted the encrypted secret key and *offset* vector to the healthcare provider server as shown in Fig. 1.

On the receiver side, an authorized medical doctor or expert equipped with a mobile application could display and observe the reconstructed ECG data after reconstruction. Reconstruction's backend process is described as follows. Firstly, the healthcare provider's server performed decryption of both the secret key and *offset* vector. Secondly, the server downloaded the anonymized data from public cloud database. Finally, it executed the reconstruction algorithm for the anonymized data using the decrypted secret key and *offset* vector.

Fig. 14 presents a mobile phone display of ECG data from a subject with identification number ss25029503. The left

part of the figure displays the anonymized ECG data taken from the public cloud server without reconstruction. This figure is an example of ECG data that was being exposed by an attacker that was somehow able to break through the cloud server security and got access to the data. The figure shows that the mobile phone displays a mere noisy like data instead of the expected ECG data. Only the authorized medical doctor or expert who has the secret key and the *offset* vector (stored in the healthcare provider server) can

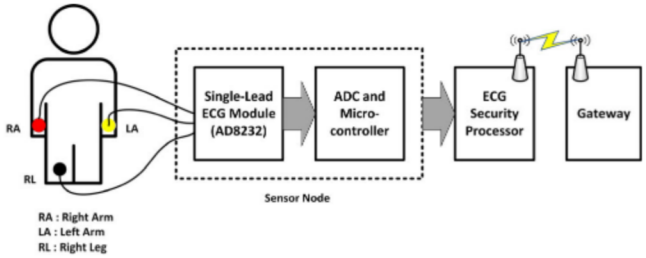

FIGURE 11. Block diagram of the single-lead ECG recording and processing.

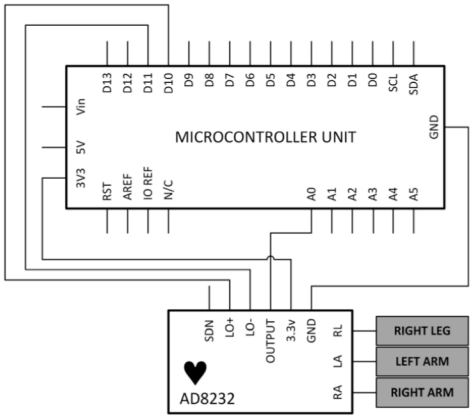

FIGURE 12. Circuit diagram of the sensor node.

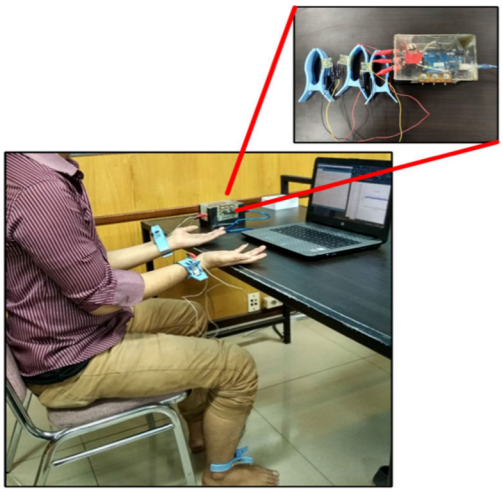

FIGURE 13. ECG data acquisition in the seated posture and the sensor node (inset).

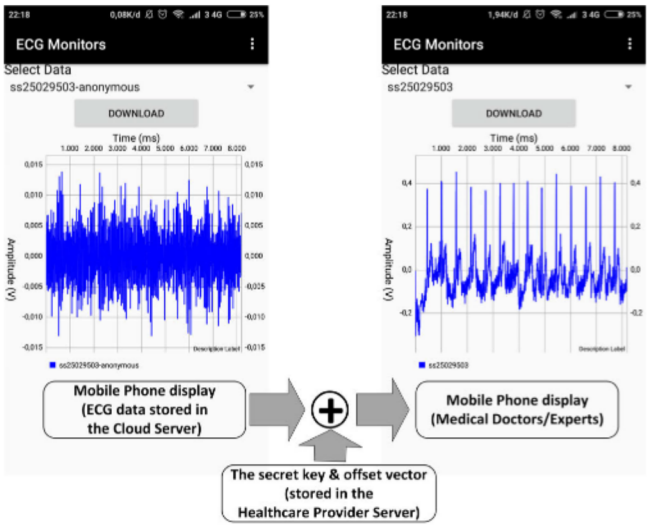

FIGURE 14. Mobile phone display of the ECG data (ss25029503). The anonymized ECG data stored in the cloud server (left) and the reconstructed ECG data displayed in the medical doctors/experts mobile phone (right).

TABLE 4. Averaged processing time of anonymization and reconstruction algorithms.

Key Length (bit)	Averaged Processing Time (ms)	
	Anonymization	Reconstruction
256	2.6	2.6
512	2.6	2.5
1024	2.4	2.3
2048	2.7	2.4
4096	2.6	2.4
Average	**2.58**	**2.44**
Confidence Interval (95%)	**2.44 – 2.72**	**2.30 – 2.58**

successfully obtain, reconstruct, and display the original ECG data. The reconstructed ECG data is shown in the right part of Fig. 14.

To conclude the evaluation, we measured averaged processing time for the anonymization and reconstruction algorithms independently to acquire clear figure of the algorithm time response in a real environment. Table 4 shows the averaged processing time (over 100 trials) that appears to be constant against variation of the key length. The results are in agreement with the results of our previous examination in Table 3. Due to the efficacy of the FFT and iFFT algorithms used in the proposed model, the single-board computer would not be able to detect small differences in processing times between key lengths. Furthermore, the table reveals that the anonymization and reconstruction phases consume similar amount of time. The only difference is that the anonymization algorithms was performed in a single-board computer (see the ECG security processor in Fig. 11) whereas the reconstruction algorithms was executed in the healthcare provider's server.

V. CONCLUSION

This paper established a non-cryptographic approach to preserve ECG data confidentiality and integrity in the EHR environment. The proposed algorithm has been shown to perform well to preserve confidentiality of patient's cardiac information (both fiducial and non-fiducial features) during transmission of ECG data and to protect information storing in the cloud database. We have examined confidentiality of the anonymized data using the PRD. In order to prove the integrity of the reconstructed ECG data, we have evaluated both the original and the reconstructed data in terms of the cross-correlation. On the other hand, security analysis has been carried out using PRD, brute force attack, and performance comparison between the proposed algorithm and the existing methods. Based on this evaluation, it can be concluded that the proposed scheme offers a secure and robust model for protecting patient's privacy in both transmission and storing patient's clinical data in the cloud. Additionally, performance evaluation in terms of processing time proved that the proposed algorithm approximately 10 times faster than the existing wavelet packet-based algorithm for processing with 65,536 points of ECG data and it approximately 150 times faster for processing with 4,096 points of data. Finally, the experimental testbed has been accomplished to show that the proposed scheme works well to support the existing EHR system. Evaluation on the processing time of both the anonymization and reconstruction algorithms shows advantages of the proposed model in a way that the processing speed is not affected by variation of the key length.

REFERENCES

[1] S. B. Baker, W. Xiang, and I. Atkinson, "Internet of Things for smart healthcare: Technologies, challenges, and opportunities," *IEEE Access*, vol. 5, pp. 26521–26544, Nov. 2017, doi: 10.1109/ACCESS.2017.2775180.

[2] Y. Yin, Y. Zeng, X. Chen, and Y. Fan, "The Internet of Things in healthcare: An overview," *J. Ind. Inf. Integr.*, vol. 1, pp. 3–13, Mar. 2016, doi: 10.1016/j.jii.2016.03.004.

[3] J. Jusak and I. Puspasari, "Wireless tele-auscultation for phonocardiograph signal recording through zigbee networks," in *Proc. IEEE Asia Pacific Conf. Wireless Mobile (APWiMob)*, Aug. 2015, pp. 95–100, doi: 10.1109/APWiMob.2015.7374939.

[4] S. M. Riazul Islam, D. Kwak, M. Humaun Kabir, M. Hossain, and K.-S. Kwak, "The Internet of Things for health care: A comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015, doi: 10.1109/ACCESS.2015.2437951.

[5] R. Prakash, S. V. Girish, and A. B. Ganesh, "Real-time remote monitoring of human vital signs using Internet of Things (IoT) and GSM connectivity," in *Proc. Int. Conf. Soft Comput. Syst.*, 2016, pp. 47–56, doi: 10.1007/978-81-322-2674-1_5.

[6] A. Limaye and T. Adegbija, "HERMIT: A benchmark suite for the Internet of Medical Things," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 4212–4222, Oct. 2018, doi: 10.1109/JIOT.2018.2849859.

[7] S. Ugalmugel and R. Swain, "Electronic health records by product, by application, by end-use, industry analysis report, country outlook, application potentials, price trends, competitive market share & forecast, 2019–2025," Global Market Insights, Selbyville, DE, USA, Tech. Rep. GMI3130, Mar. 2019.

[8] A. Abbas and S. U. Khan, "A review on the state-of-the-art privacy-preserving approaches in the e-health clouds," *IEEE J. Biomed. Health Inform.*, vol. 18, no. 4, pp. 1431–1441, Jul. 2014, doi: 10.1109/JBHI.2014.2300846.

[9] Z. Xiao and Y. Xiao, "Security and privacy in cloud computing," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 843–859, 2nd Quart., 2013, doi: 10.1109/SURV.2012.060912.00182.

[10] (2017). *Security Guidance for Critical Areas of Focus in Cloud Computing v4.0. Cloud Security Alliance*. Accessed: Oct. 26, 2020. [Online]. Available: https://cloudsecurityalliance.org/research/guidance/

[11] N. A. Azeez and C. Van der Vyver, "Security and privacy issues in e-health cloud-based system: A comprehensive content analysis," *Egyptian Inform. J.*, vol. 20, no. 2, pp. 97–108, Jul. 2018, doi: 10.1016/j.eij.2018.12.001.

[12] S. Chenthara, K. Ahmed, H. Wang, and F. Whittaker, "Security and privacy-preserving challenges of e-Health solutions in cloud computing," *IEEE Access*, vol. 7, pp. 74361–74382, 2019, doi: 10.1109/ACCESS.2019.2919982.

[13] P. Sun, "Security and privacy protection in cloud computing: Discussions and challenges," *J. Netw. Comput. Appl.*, vol. 160, Jun. 2020, Art. no. 102642, doi: 10.1016/j.jnca.2020.102642.

[14] A. Mehmood, I. Natgunanathan, Y. Xiang, H. Poston, and Y. Zhang, "Anonymous authentication scheme for smart cloud based healthcare applications," *IEEE Access*, vol. 6, pp. 33552–33567, 2018, doi: 10.1109/ACCESS.2018.2841972.

[15] J. L. Fernández-Alemán, I. C. Señor, P. Á. O. Lozoya, and A. Toval, "Security and privacy in electronic health records: A systematic literature review," *J. Biomed. Informat.*, vol. 46, no. 3, pp. 541–562, Jun. 2013, doi: 10.1016/j.jbi.2012.12.003.

[16] F. Tang, S. Ma, Y. Xiang, and C. Lin, "An efficient authentication scheme for blockchain-based electronic health records," *IEEE Access*, vol. 7, pp. 41678–41689, 2019, doi: 10.1109/ACCESS.2019.2904300.

[17] A. Shahnaz, U. Qamar, and A. Khalid, "Using blockchain for electronic health records," *IEEE Access*, vol. 7, pp. 147782–147795, 2019, doi: 10.1109/ACCESS.2019.2946373.

[18] A. Farouk, A. Alahmadi, S. Ghose, and A. Mashatan, "Blockchain platform for industrial healthcare: Vision and future opportunities," *Comput. Commun.*, vol. 154, pp. 223–235, Mar. 2020, doi: 10.1016/j.comcom.2020.02.058.

[19] A. Tandon, A. Dhir, A. K. M. N. Islam, and M. Mäntymäki, "Blockchain in healthcare: A systematic literature review, synthesizing framework and future research agenda," *Comput. Ind.*, vol. 122, Nov. 2020, Art. no. 103290, doi: 10.1016/j.compind.2020.103290.

[20] A. Hasselgren, K. Kralevska, D. Gligoroski, S. A. Pedersen, and A. Faxvaag, "Blockchain in healthcare and health sciences—A scoping review," *Int. J. Med. Informat.*, vol. 134, Feb. 2020, Art. no. 104040, doi: 10.1016/j.ijmedinf.2019.104040.

[21] E. Chukwu and L. Garg, "A systematic review of blockchain in healthcare: Frameworks, prototypes, and implementations," *IEEE Access*, vol. 8, pp. 21196–21214, 2020, doi: 10.1109/ACCESS.2020.2969881.

[22] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: A new approach in human identification," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 3, pp. 808–812, Jun. 2001, doi: 10.1109/19.930458.

[23] A. Condon and G. Willatt, "ECG biometrics: The heart of data-driven disruption," *Biometric Technol. Today*, vol. 2018, no. 1, pp. 7–9, 2018, doi: 10.1016/S0969-4765(18)30011-0.

[24] M. Merone, P. Soda, M. Sansone, and C. Sansone, "ECG databases for biometric systems: A systematic review," *Expert Syst. Appl.*, vol. 67, pp. 189–202, Jan. 2017, doi: 10.1016/j.eswa.2016.09.030.

[25] *HIPAA Security Series: 1 Security 101 for Covered Entities*, Centers for Medicare Medicade Sercices, US Department of Health and Human Services, Washington, DC, USA, 2007, pp. 1–11, vol. 2.

[26] J. Jusak and S. S. Mahmoud, "A novel and low processing time ECG security method suitable for sensor node platforms," in *Int. J. Commun. Netw. Inf. Secur.*, vol. 10, no. 1, pp. 213–222, 2018.

[27] *Health Informatics—Pseudonymization*, document ISO 25237:2017, 2017.

[28] F. Sufi, S. S. Mahmoud, and I. Khalil, "A novel wavelet packet-based anti-spoofing technique to secure ECG data," *Int. J. Biometrics*, vol. 1, no. 2, pp. 191–208, Aug. 2008, doi: 10.1504/IJBM.2008.020144.

[29] S. S. Mahmoud, "A generalised wavelet packet-based anonymisation approach for ECG security application," *Secur. Commun. Netw.*, vol. 9, no. 18, pp. 6137–6147, Dec. 2016, doi: 10.1002/sec.1762.

[30] *The PTB Diagnostic ECG Database (Physionet)*. Accessed: Jun. 2020. [Online]. Available: http://www.physionet.org/physiobank/database/ptbdb

JUSAK JUSAK (Member, IEEE) received the B.S. degree in electrical engineering from Brawijaya University, Malang, Indonesia, in 1996, and the Ph.D. degree in electrical engineering from the Royal Melbourne Institute of Technology (RMIT) University, Melbourne, Australia, in 2006.

From 2009 to 2011, he was a Postdoctoral Research with Massey University, Palmerston North, New Zealand, working in a Next Generation Networks research project supported the Telecom New Zealand. Between 2011 and 2021, he was a Senior Lecturer with the Computer Engineering Department, Dinamika University, Surabaya, Indonesia. He is currently with James Cook University Singapore. His research interests include signal processing for wireless communication networks, biomedical signal processing, and Internet of Things for medical applications and its security.

Dr. Jusak was a recipient of Ph.D. research award, in 2004. He received several national level research competitive grants, from 2011 to 2020, and the best paper award in the International Conference on Information Technology Applications and Systems, in 2018.

SEEDAHMED S. MAHMOUD (Senior Member, IEEE) received the B.S. degree (Hons.) in medical instrumentation from Gezira University, Sudan, in 1996, the M.S. degree in electronic systems engineering from University Putra Malaysis, Malaysia, in 1998, and the Ph.D. degree in electrical engineering from RMIT University, Melbourne, Australia, in 2004.

He was appointed as an Associate Professor by Shantou University, in 2019. He is the part of the research and teaching faculty with the Department of Biomedical Engineering, Shantou University. He has eight years of teaching experience and has taught various courses in electrical engineering, including signal processing in biomedical engineering with the Department of Electrical and Electronic Technology, Lincoln College International, Saudi Arabia. Prior to this, he was a Research Fellow with RMIT University. He has 11 years of professional working experience at Digital Signals Division of Future Fibre Technologies (FFT) Ltd., Melbourne. As the Leader of the research team at FFT, he developed a variety of innovative algorithms and identification technologies for FFT's optical fiber intrusion detection system. His work on machine learning for perimeter intrusion detection system has led to three international patents. He has been involved in a number of research projects, has a number of patents, published more than 20 journal articles which have been cited more than 40 times, and published one book and three book chapters. He is also serving in the Editorial Board for the *Journal of the Computer Science* (JCS) and a reviewer for a number of IEEE and Elsevier journals.

ROY LAURENS (Member, IEEE) received the B.S. degree in computer science from Sekolah Tinggi Teknik Surabaya, Indonesia, in 1996, and the M.S. degree in information network from Carnegie-Mellon University, USA, in 1998. He is currently pursuing the Ph.D. degree in computer science with the University of Central Florida, USA. He is also a Distance-Learning Lecturer with the Department of Computer Engineering, Dinamika University, Surabaya. His research interests include payment card fraud and network security.

Since 2012, he has been the System Architect for the payment card fraud prevention at Sola Fide, Inc., in Tampa, FL, USA. He is also the director of the several non-profit organizations in the Central FL area. He has a U.S. patent for his work in mobile and wireless environment. His research interests include automating fraud detection using both rule-based machine learning and anomaly detection to identify fraudulent transaction while reducing customer friction.

MUSLEH ALSULAMI received the B.Sc. degree in computer science from Imam University, Saudi Arabia, in 2004, the M.Sc. degree in information technology, in 2010, and the Ph.D. degree in information systems from Monash University, Australia, in 2017. He is currently an Assistant Professor in information systems with Umm Al-Qura University. His current research interests include enterprise resources planning (ERP), cloud ERP, ERP life cycle, ERP implementation stakeholders' conflicts, digital transformation in government organization, software quality, and human–computer interaction.

QIANG FANG (Member, IEEE) received the B.S. degree in applied physics form Tsinghua University, Beijing, China, in 1991, and the Ph.D. degree in biomedical engineering from Monash University, Melbourne, Australia, in 2000. He is currently the Founding Chair of the Biomedical Engineering Department and a Full Professor with Shantou University, China. Before, he moved to Shantou University, in 2017, he was a permanent Academic Staff of RMIT University, Melbourne, for 15 years. His research interests include intelligent and miniaturized medical instrumentation, wearable and implantable body sensor networks, and pervasive computing technologies.