Received March 8, 2019, accepted April 8, 2019, date of publication April 12, 2019, date of current version April 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2910753

Unlock Your Heart: Next Generation Biometric in Resource-Constrained Healthcare Systems and IoT

NIMA KARIMIAN[1], (Member, IEEE), MARK TEHRANIPOOR[2], (Fellow, IEEE), DAMON WOODARD[2], (Senior Member, IEEE), AND DOMENIC FORTE[2], (Senior Member, IEEE)

[1]Department of Computer Engineering, San Jose State University, San Jose, CA 95192, USA
[2]Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611, USA

Corresponding author: Nima Karimian (nima.karimian@sjsu.edu)

This work was supported in part by the U.S. Army Research Office grant under award number Grant W911NF16-1-0321.

ABSTRACT With the emergence of the Internet-of-Things, there is a growing need for access control and data protection on low-power, pervasive devices. Key-based biometric cryptosystems are promising for IoT due to its convenient nature and lower susceptibility to attacks. However, the costs associated with biometric processing and template protection are nontrivial for smart cards, and so forth. In this paper, we discuss the cost versus the utility of biometric systems and investigate frameworks for improving them. We propose the noise-aware biometric quantization framework (NA-IOMBA) capable of generating unique, reliable, and high entropy keys with low enrollment times and costs with several experiments. First, we compare its performance with IOMBA and one-class-SVM on multiple biometric modalities, including popular ones (fingerprint and iris) and emerging cardiovascular ones (ECG and PPG). The results show that NA-IOMBA outperforms them all and that ECG provides the best trade-off between reliability, key length, entropy, and implementation cost. Second, we examine the impact on key reliability with ECGs obtained at different sessions and trained with a different number of heartbeats. Finally, implementation results show that incorporating noise models with NA-IOMBA reduces power and utilization overhead by more than 60% by adapting the pre-processing, feature extraction, and post-processing modules.

INDEX TERMS Internet of Things, ECG, biometric, quantization, PPG, noise, healthcare, resource-constrained, security, access control.

I. INTRODUCTION

Internet of Things (IoT) has come in order to demonstrate the widespread of sensors and actuators to create a ubiquitous embedded architecture where machine-to-machine, human-to-machine, human-to-thing, etc. interactions are commonplace. Nowadays, IoT technology has pervaded into our everyday life through the smartphone, smart office, autonomous vehicle, smart homes, smart transportation and potentially cost-effective approach to the Internet in order to monitor physical systems and critical infrastructure such as remote healthcare application, airplane, etc. Our healthcare is increasingly being made smarter by the patient needs through embedded sensors which integrated into the wearable medical sensors and modern communication. This device can enable continuous and remote monitoring of people health at a low cost. Wearable technologies such as Fitbit, Apple Watch, and Microsoft Band are becoming part of everyday life and they have been integrated into a person's daily routine which is making it possible to remotely monitor a patient's health with low-cost sensors. For example, the Microsoft Band can be used to monitor heart rate, 3-axis accelerometer, skin temperature, and galvanic skin response that can be useful for people who are actively involved in data collection in diagnosing to finding the best feasible treatment solution. Apple watch series 4 have added electrodes on the digital crown and on the electrical heart rate sensor on the back of the watch that can measure heart activity (ECG signal) to detect abnormalities

The associate editor coordinating the review of this manuscript and approving it for publication was Adnan M. Abu-Mahfouz.

in the rhythm. In recent studies [1], [3] on the effectiveness of telehealth, they found that telehealth can be useful to track patients with higher exposure to prevent demanding conditions to save their lives. Moreover, they can reduce the hospitalization cost, the length of time that patients need to stay in the hospital, and mortality due to heart failure [1].

As IoT based Healthcare devices and applications are expected to monitor patient by gathering the vital data and transmitting it to other sources for remote monitoring. If they fail, the patient's life is at risk. Examples include pacemakers. Thus, it is expected that security and patient privacy have higher priority in driving such technologies. In addition, such IoT medical devices may be joined to global information networks for their access anytime, anywhere. Therefore, the IoT healthcare may be a target of adversarial and can put the patient's life at risk. Moreover, the National Healthcare Anti-Fraud Association estimates the loss to health care to be about billions of dollars annually in many countries [4], [5] based on fraud, waste, and abuse in healthcare. In order to assure the healthy operation of a Healthcare insurance system, patient verification can prevent false data injection for access control.

One important issue related to IoT-based healthcare application is an incorrect patient identification that leading to inaccurate treatment solution on the patient (such as wrong drug/dose/time/procedure). Therefore, the process of avoiding this happen is costly and have a significant impact on the healthcare provider [7], [14]. In addition, one of the crucial demands in healthcare that is how we can transmit data in a safe way. As it turns out, the existing access control structure relies heavily on traditional authentication, a new range of security attacks becomes feasible. Biometric technology-based patient recognition can significantly provide the desired levels of speed and reliability. Biometrics can also be considered more seamless and convenient, especially for continuous authentication. That being said, it has already been demonstrated that many of the most popular biometric modalities (iris, face, fingerprint, and speech) [12], [13], [17] has been used to overcome some of the aforementioned limitations.

Therefore, there is a need to provide a balance between medical data security and resource consumption of IoT wearable device. In recent years, the objectives of electrocardiogram (ECG) and photoplethysmogram (PPG) monitoring have gone beyond more heart rate and rhythm measurement to the analysis of chronic illnesses including diabetes, heart disease, and sleep disorders among others [18], [19]. Recently, it has been shown that the cardiovascular signal from person-to-person is unique and may be distinctive enough for biometric applications [17], [20]. In order to minimize product costs, remote monitoring of IoT based healthcare is likely to come equipped with low-power computing devices. This implies that biometric and security technologies should consider the constraints of IoT healthcare device. However, biometric protection becomes challenging since the template is frequently compromised, therefore, key derivation functions are used to derive cryptographic keys to protect sensitive information. In addition, IoT device is resource constrained, thus reducing storage complexity are needed. As a result, embedding identity information into a binary template can be applied to reduce the storage and matching complexity. However, biometric key generation may suffer from environment/measurement noise, internal instability, and so on which could result in errors during key generation process and serious security and privacy threat such as intrusion attack and cross-matching the template [48], [27] and make it impossible for even the system owner to access/use the system. Due to a large intra-subject variance during multiple acquisitions of the same biometric trait, it is a challenging task in order to overcome this issue for biometric template protection. To mitigate this challenging task, we propose a novel biometric key generation from noisy data to mitigate or eliminate the intra-subject variance while preserving privacy and generating long keys by developing a statistical approach.

There has been a number of technique deployed for biometrics key generation using in the literature. Monrose *et al.* [42] and Teoh *et al.* [43] applied a scheme that quantizes each background feature space into two intervals where each interval is labeled with bit "0" or "1" based on a fixed threshold. A feature value that falls into an interval is mapped to the corresponding 1-bit output label. Kelkboom *et al.* [2] developed a framework to estimate the genuine and imposter bit error probability mass functions by measuring hamming distance. Drozdowsk *et al.* [44], the authors applied deep learning for feature extraction followed by encoding schemes (Linearly Separable Subcode) which exhibit full-ideal and near-ideal separability capabilities, respectively. Chen *et al.* [36], [45] proposed a generic bit allocation algorithm scheme using pairwise adaptive phase quantization and long-short pairing strategy. Lim *et al.* [37], [46] developed a DROBA-based approach, to improve the biometric performance of the binarized representation of facial features based on bit statistics (reliability and discriminability). Osadchy and Dunkelman [31], explore the existing security and privacy of feature extraction and binarization processes where they found that the most important part of biometric protection is how to extract features with high accuracy. Kaur and Khanna [40] proposed cancelable biometric template protection to address security and privacy. In 2019, cancelable ECG biometric has been proposed by Wu *et al.* [6] to address the biometric privacy properties such as revocability, unlinkability, and irreversibility.

One of the most important requirements in bio key is the entropy [8] in order to guarantee resistance against attacks. However, it is a challenging task to simultaneously achieve high key entropy and high key stability. Usually, error correction code (ECC), and fuzzy extraction will be applied [9], [11]. However, fuzzy extraction has potential drawbacks including high computational cost, increased area overhead, and vulnerability to side channel attacks [10] which again it is not feasible for IoT application. The best method to protect user biometrics is a hash function. However, since the biometric is noisy, therefore, the hash function cannot be directly used. To address this issue, the quantization technique where it can map noisy samples to a unique vector then protected by hash function. There is limitation associated with existing quantization algorithm, where the formulation is a generic characteristic and because of that, the attacker has appropriate knowledge of transformation algorithm and parameters, which resulted in linkability of the system. In fact, most of the limitation due to reliability and entropy of the system. In addition, the binarization technique is generic, however, our approaches adapt to the population itself (to preserve entropy) while also being adaptable to individual users (to improve reliability), without revealing any information about the system's users.

The difference between this work and the previous work is that the tunable parameters which allow an amount of intra-user variability among multiple acquisitions of the same biometric trait are dynamically updated by optimizing local minimum of the proposed algorithm. The major goal of this work is to demonstrate how the proposed algorithm can address the intra-variability of ECG-based biometric while multiple biometric modalities such as PPG, iris, and fingerprint is conducted. We also conducted the proposed algorithm with the challenging ECG database with very noisy signal under multiple data acquisition session. Moreover, none of the existing work implements their proposed algorithm, in this work, our algorithm for the first time has been implemented to evaluate it in terms of the feasibility of the IoT application.

In this paper, we focus on how to incorporate a new biometric key generation/authentication framework that provides secure and reliable access control in resource-constrained environments for remote healthcare application. Our main contributions can be summarized as follows:

1) **Noise-aware Feature Quantization**: We propose noise aware interval optimized mapping bit allocation (NA-IOMBA) to select and quantize biometric features on a user-to-user basis. In other words, only the most robust features are selected for each user, thus avoiding unnecessary post-processing costs. Noise models are used to predict the impact of noise and further reduce error correction costs and enrollment times.
2) **Multiple Modalities**: We perform experiments on four biometrics modalities (ECG, PPG, fingerprint and iris scan) to demonstrate the effectiveness of NA-IOMBA. Experimental results show that NA-IOMBA increases the length, entropy, and robustness of keys generated from multiple biometric modalities. Further, ECG-based authentication performs the best among the candidates. Thus, we take a closer look at its performance throughout the rest of the paper.
3) **Synthetic ECG signals:** One of the biggest challenges in electrocardiogram (ECG) based authentication and/or key generation lies in how to obtain noisy ECG data. Put simply, it is impossible to take raw ECG measurements from a large population under all possible test conditions. In this paper, we discuss how dynamical models can be used to generate synthetic ECG signals for common noise sources as well as activity/stress.
4) **Biometric Authentication Cost and Enrollment Time**: By incorporating noise models into feature selection, certain denoising/filtering steps can be avoided by NA-IOMBA. Implementation results from Xilinx Zynq-7000 show that, even with less resources (65% power and 62% utilization), NA-IOMBA offers higher reliability. In addition, the number of enrollment samples and enrollment time can be reduced since biometric noise and different conditions are modeled in NA-IOMBA. This behavior is confirmed with experiments using multi-session data and training data from non-ideal conditions.
5) **Machine Learning vs. Quantization**: Machine learning results are compared to our quantization schemes on multiple biometric modalities. We demonstrate that the accuracy of the proposed approach performs just as good, and in most cases better than machine learning based authentication.
6) **Security Analysis based on Entropy and Key Randomness**: To evaluate the correctness and security of our proposed biometric key generation, we inspect it using a set of standard statistical suite test (NIST). The results of these test on our proposed method passes the randomness test.

The rest of the paper is structured as follows. In Section II, the noise-aware quantization framework is described. Denoising overhead reduction, ECG synthetic model, ECG noise modeling, and ECG stress modeling, are described in Section III. Section IV presents NA-IOMBA case studies including the comparison of multiple biometric modalities, ECG long-term feasibility, the impact of heartbeats on training, a fixed threshold in IOMBA, and machine learning vs quantization approach. The analysis of noisy ECG signal, FPGA implementation of our approach, and reduction in enrollment times are presented in Section IV. Finally, the conclusion is drawn in Section VI.

II. NOISE-AWARE QUANTIZATION FRAMEWORK

A. IOMBA BASED BIOMETRIC KEY GENERATION

Biometric systems that operate using any single biometric characteristic have several limitations such as noise in sensed data, intra-class variations (the biometric data acquired from an individual during authentication may be very different from the data that was used to generate the template during enrollment), and distinctiveness (biometric trait is expected to vary significantly across individuals). Noise in biometric authentication and recognition applications can be tolerated to some extent; however, not a single error can be tolerated in key generation applications. To address this limitation, we have previously developed the interval optimized mapping bit allocation (IOMBA) scheme for biometric key generation [21].

In short, IOMBA tunes the biometric key generation process to each user rather than relying on a generic approach for all users. The performance is controlled by two parameters, α and β, which define the entropy and reliability of generated keys respectively. IOMBA eliminates features from the space that has low entropy and quantizes features with larger entropy and lesser noise into more bits. Then, since noise will impact each user's biometric differently, the features that are more sensitive to noise (i.e., unreliable for key generation) are removed on a user-to-user basis. As a result, the length of keys generated varies based on the α and β parameters as well as from user to user.

The major steps of the IOMBA are described as follows.

1) DATA PRE-PROCESSING

The signals from the population are pre-processed to remove noise followed by feature extraction. The feature elements from the same location are extracted from the population and normalized into a standard normal distribution. The same normalization parameters are later exploited to normalize the corresponding feature elements of each subject. Further, a decorrelation step can be applied to the distributions as well. Note that our approach can work with any biometric provided it produces statistically independent and Gaussian features in some representation - any feature that does not meet these requirements will be discarded.

2) IOMBA MARGIN CALCULATION FROM POPULATION STATISTICS

IOMBA quantizes each feature into a different number of bits. 2 bit quantization is illustrated in Fig. 1. The population probability density function (PDF) of a feature is shown in blue. To illustrate the reliability calculation, distributions for a single feature from three subjects are overlaid on the same plot. Figure 1, illustrates the overlap between a user's feature PDF and the T threshold. The amount of overlap indicates the amount of error that we can expect if the feature is chosen for the user's key. As can be seen in Fig. 1, '01' is encoding of the features considered when it is measured to be on the right of T. However, due to noise, there will be an error encoding the feature if it appears on the left of T. To mitigate this problem, the maximum allowable overlap for a reliable feature as β is implemented in our algorithm. Thus, β controls the probability of error in (or reliability of) the key generation. Following equations are computed for the two bit encoding case, where, only features that satisfy the following constraints can be selected for key generation.

$$\int_{T}^{\infty} PDF_{\text{pop},f} \leq \beta, \quad \text{if } \mu_{f,i} < T \tag{1}$$

$$\int_{-\infty}^{T} PDF_{\text{pop},f} \leq \beta \cap \int_{0}^{\infty} PDF_{f,i} \leq \beta, \quad \text{if } T < \mu_{f,i} < 0 \tag{2}$$

where $\mu_{f,i}$ denotes the mean for feature f of subject i, and β is the maximum allowable bit-error probability. Note that similar constraints can be specified for $\mu_{f,i} > 0$, but are withheld for brevity. Note based on our definitions above, the constraints can also be written as

$$\mu_{f,i} \leq \mu_{00}, \quad \text{if } \mu_{f,i} < T \tag{3}$$

$$\mu_{\text{left},01} \leq \mu_{f,i} \leq \mu_{\text{right},01}, \quad \text{if } T < \mu_{f,i} < 0 \tag{4}$$

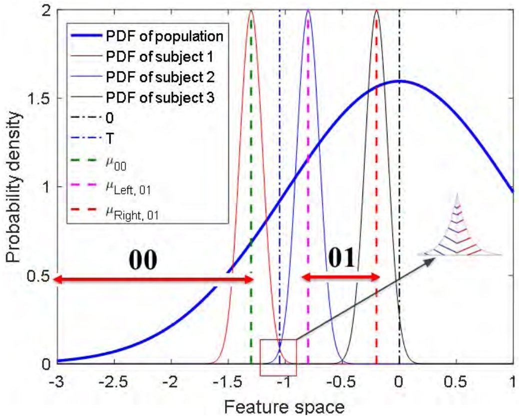


FIGURE 1. **Illustration showing how IOMBA optimizes quantization of a single feature into two bits.**

Moreover, the entropy should be large enough to guarantee resistance against attacks. To address this problem,α is implemented into this algorithm to achieve high entropy. Let P_{00} and P_{10} denote the probability of a feature falling into bins '00' and '10' respectively. Therefore, we have the following equation:

$$P_{00} = \int_{-\infty}^{\mu_{00}} PDF_{\text{pop},f}\,dx \tag{5}$$

$$P_{01} = \int_{\mu_{left,01}}^{\mu_{right,01}} PDF_{\text{pop},f}\,dx \tag{6}$$

The ideal case is $P_{00} = P_{01}$ where key bits are equally distributed. However, it may be too restrictive in practice. As a result, it may not be possible to find thresholds and parameters to fulfill this condition, resulting in very few features for key generation. To relax this constraint, α along with the following constraint has been implemented as the entropy parameter

$$\frac{P_{01}}{P_{00}} \leq \alpha \tag{7}$$

Note that the above formulation applies to the two bit case, but this approach can also be extended to three bits, four bits, etc. In practice, for each feature one can quantize to the maximum number of bits based on the input reliability and entropy parameters, β and α. Fig. 1 illustrates these values on the negative side of the x axis. The positive boundaries and thresholds can be simply computed by mirroring the negative values onto the positive side against the y axis. These boundaries and thresholds guarantee that enrolled and regenerated key bits are statistically random and reliable.

3) ENROLLMENT FOR KEY AND HELPER DATA GENERATION

Essentially, IOMBA personalizes a biometric system to each user. For each subject, the key generation framework utilizes the above boundaries to determine whether each feature element is good for generating key bits or not. The least reliable features of a user which do not fulfill the above constraints are discarded. The helper data for each subject consists of the following: (i) the index of the reliable features selected for the user, (ii) the number of bits each feature can be quantized into, and (iii) the normalization parameters for each feature. If an error still exists, then error correction based helper data and an error correction module can also be added in the next step.

4) KEY REGENERATION

The user presents his/her biometric, features are extracted, and the helper data stored on the system is used to eliminate the unreliable features and then quantize the reliable features to regenerate the key.

In IOMBA, feature extraction, helper data, and key regeneration steps are different for each user. Reconfigurable hardware is, therefore, a natural candidate for implementing IOMBA because each step can be tailored uniquely to the user, thereby avoiding certain processing costs. Results show that the proposed approach for ECG key generation achieves 28% improvement of reliability and four times longer key size in the worst case scenario compare to our previous work [21].

B. NOISE AWARE IOMBA (NA-IOMBA)

In the original IOMBA, the standard deviation in user PDFs was fixed for each feature in a worst-case manner. Feature selection was therefore pessimistic and resulted in shorter keys. In fact, when the biometric data is not noisy, IOMBA approach works quite well. Estimating the standard deviation, especially for continuous biometrics (keystroke dynamics, ECG, etc.), is nontrivial since the biometric would need to be collected at all conditions and types of noise. In most applications, such as IoT, the enrollment process would be too long for users to tolerate. In addition, the impact of noise is substantially affected by the type and amount of pre-processing. In resource-constrained scenarios, it would be better to eliminate pre-processing steps which are costly and energy consuming to perform. To accommodate these issues, we propose the noise-aware IOMBA framework in this section and demonstrate its benefits in Section IV. NA-IOMBA is a variant of IOMBA that incorporates models to predict the impact of different noise sources, noise scales, and pre-processing steps on biometric key generation technique. In short, all IOMBA steps are performed with one major change; Margins and boundaries are recomputed based on more accurate estimates of user feature standard deviation. Basically, features that are modeled as less (more) susceptible to noise will, therefore, be given smaller (larger) margins than IOMBA. To better understand NA-IOMBA, we have illustrated it in Fig. 2. First, optimal margins and thresholds are calculated based on inputs α (reliability parameter), β (entropy parameter) and population statistics of a biometric sample. Next, an artificial noise will be added to the biometric sample in order to capture the worst case scenario of a noisy biometric condition. Then, the IOMBA margin and boundaries will be updated based on the new input data (e.g. noisy biometric samples). The margins and thresholds for each feature are then used to select the most reliable features from a user's biometric sample and quantize it into a high entropy key. The indexes of the selected features, margin, and boundaries are stored as helper data for later usage (authentication or key generation phase).

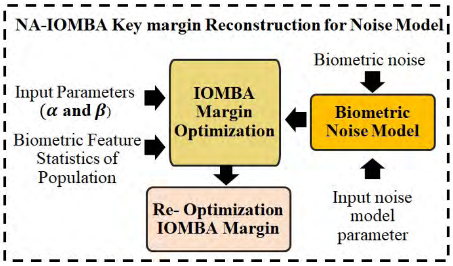

NA-IOMBA Key margin Reconstruction for Noise Model
Input Parameters (α and β)
Biometric Feature Statistics of Population
IOMBA Margin Optimization
Biometric noise
Biometric Noise Model
Input noise model parameter
Re- Optimization IOMBA Margin

FIGURE 2. Block diagram show for NA-IOMBA margin reconstruction key calculation from noisy biometric.

In order to determine a noise model in NA-IOMBA (see Fig. 2), N_e and N_v has been considered as an enrollment and verification noise measurement, respectively. The synthetic noise model is employed as verification noise. Synthetic noise which enables us to avoid exhaustive measurement of noises for the assessment in verification step. We also assume that N_e and N_v are mutually independent where the σ_e and σ_v denoted as standard deviation of measurement noise for enrollment and verification, respectively. Thus, we can assume $X = S + N_e$ and $Y = S + N_v$ are noisy biometric signals in enrollment and verification if we considered S is clean signal. Since the noise are mutually independent [2], the standard deviation of X and Y is indicated by $\sigma_X = \sigma_S + \sigma_e$ and $\sigma_Y = \sigma_S + \sigma_v$. Thus the verification biometric sample Y is related to the enrollment measurement as $Y = \lambda X + R$ where the λ is the input noise model parameter and R is zero-mean biometric noise, independent of X as indicated in Fig. 2. Please note that physionet bank database provided ECG and PPG noises, therefore ECG/PPG noise model has been generated in this work to evaluate the proposed work under the worst case scenario. However, we could not find any noise database for iris and fingerprint, therefore, random noise with specific different SNR has been added to the feature vectors.

The following are the impacts and benefits of this approach:

- **Impact on key length**: If a margin for a feature increases, it could result in the feature **being selected** and/or longer bit lengths compared to IOMBA. If a margin for a feature shrinks, it could result in the feature **no longer being selected** and/or shorter bit lengths compared to IOMBA.

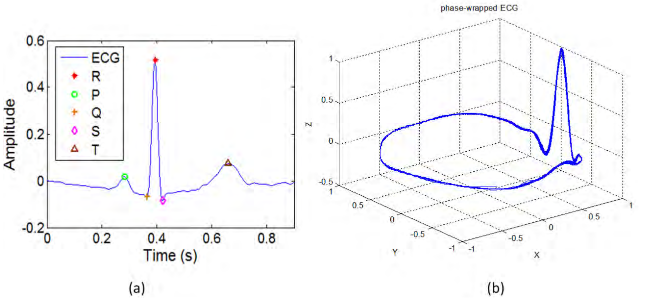


FIGURE 3. **Typical normal ECG signal. (a) One beat normal ECG signal with fiducial point, (b) trajectories several cycles of the ECG phase-wrapped in the Cartesian coordinates.**

- **Impact on key reliability**: If noise samples or models are accurate, the reliability of key should improve regardless of whether or not the key length shrinks/grows.
- **Impact on cost and enrollment time**: There are three ways that overheads can be reduced. First, error correction costs tend to increase nonlinearly. By improving key reliability, error correction hardware can be substantially reduced. Second, certain denoising/filtering steps can be removed provided that NA-IOMBA accurately estimates the noise appearing in features without them. Third, the number of enrollment samples and enrollment time can be reduced if noise over time and different conditions can be modeled. This can be particularly important for continuous physiological signals, like ECG and PPG, which can be impacted by so many different conditions, e.g., exercise, stress, and food/drink/drug consumption.

III. INCORPORATION OF REAL NOISE MODELS

In this next section, we discuss how NA-IOMBA can be further improved by incorporating actual noise models. We take ECG as an example since there are a variety of models available in the literature for it. In order to determine the sensitivity of ECG key generation based on these feature extraction, the noisy ECG signal with different variances (SNR) is applied. To view the impact of each noise source, synthetic ECGs are generated and not pre-processed to remove the noise.

A. ECG SYNTHETIC MODEL

We adopt the non-linear dynamical model proposed by McSharry *et al.* [22] to extract parameters from an ECG and generate synthetic ECGs. The aim of this approach is to provide a standard realistic ECG signal with known characteristics from ECG (Fig. 3 (a)), which can be generated with specific statistics thereby facilitating the performance evaluation of a given technique. McSharry *et al.*'s model uses three ordinary differential equations. It consists of a circular limit cycle of unit radius in the (x, y) plane around which the trajectory is pushed up and down as it approaches the P, Q, R, S and T points in the ECG. The dynamic state equations proposed by McSharry et al. can also be transformed into polar coordinates as follows [23]

$$\begin{cases} \frac{dr}{dt} = r(1-r) \\ \frac{d\theta}{dt} = \omega \\ \frac{dx}{dt} = -\sum_{i\in P,Q,R,S,T} a_i \Delta\theta_i exp[-\frac{\Delta\theta_i^2}{2b_i^2}] - (z - z_0) \end{cases} \tag{8}$$

The first equation in Eq.(8) shows the circular behavior of the generated trajectory by the model. second and third equations in Eq.(8) are independent from r, making the first equation redundant. Therefore, the first equation may be excluded as it has no effect on the synthetic ECG. To estimate the dynamic model parameters for the given ECG, mean and variance of the phase-wrapped ECG is calculated for all phases between $-\pi$ and π which are depicted in Figure 3.

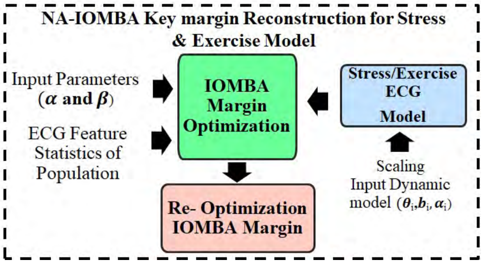

NA-IOMBA Key margin Reconstruction for Stress & Exercise Model
Input Parameters (α and β)
ECG Feature Statistics of Population
IOMBA Margin Optimization
Stress/Exercise ECG Model
Scaling Input Dynamic model (θ_i,b_i,α_i)
Re- Optimization IOMBA Margin

FIGURE 4. **Block diagram show for NA-IOMBA margin reconstruction key calculation from stressed ECG.**

B. ECG NOISE MODELING

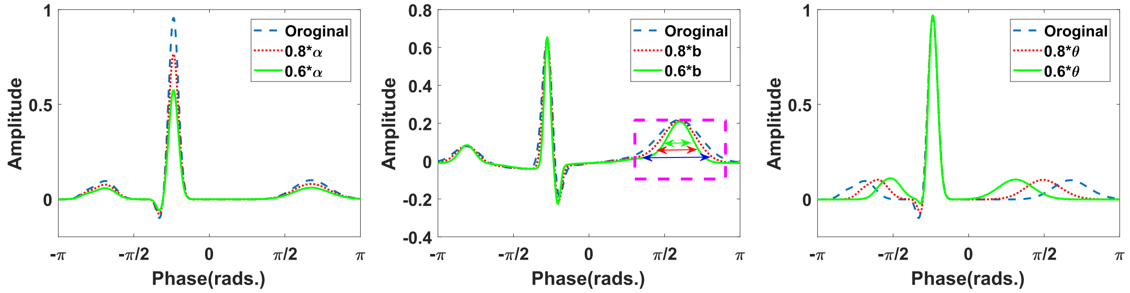
The three main types of noise sources in raw ECG signals are (1) muscle artifacts (MA) which occur due to electrical activity of muscles; (2) baseline wander (BW) caused by body movement; and (3) electrode movement (EM) due to poor contact to the sensor. We adopt the noise model from [24], [25].

Fig. 4 presents the experimental protocol of NA-IOMBA scheme for noisy ECG signals. In noisy ECG key margin reconstruction, dynamic model parameters $(\theta_i, \alpha_i, b_i)$ from original ECG signal are considered as the input of the synthetic ECG module. Then, synthetic ECG noise with desired SNR is employed to add into the clean (synthetic) ECG.

Time-varying auto-regressive (AR) parametric models are applied to generate realistic ECG noise which follows the non-stationary characteristics and the spectral shape of real noise. The parameters of this model are trained by using real noises such as NSTDB [26]. To estimate the time-varying AR parameters, a standard Kalman Filter (KF) is used [28]. For the time series of y_n, a time-varying AR model of order p can be written as follows:

$$y_n = -\sum_{i}^{p} a_n(i)y(n-i) + v_n \tag{9}$$

where v_n is the input white noise and $a_n(i)(i = 1, \ldots, p)$ coefficients are the p time-varying AR parameters at the time instance of n. We refer the reader to [24] for more details of calculating AR parameters. Having the time-varying AR



(a) (b) (c)

FIGURE 5. Impact of ECG signal by decreasing dynamical model parameters: (a) α parameters associated with amplitude, (b) b parameters, and (c) θ parameters associated with interval and heart rate.

TABLE 1. High-level summary analysis for multiple biometric modalities.

	Databases	Train / Test sizes	Pre-processing	Feature Extraction
ECG	PTB	52*1000/1560*1000	FIR filter, R peak det., Segmentation	NCN
Multi-session ECG	ECGID	90*500/2700*500	FIR filter, R peak det., Segmentation	NCN
PPG	Capnobase	42*700/ 840*700	Butterworth, Peak det., Segmentation	DWT
Iris	CASIAV1	108*4800/756*4800	Localization	Gabor wavelet
Finger	FVC2004-DB3	100*6056/800*6056	Normalization, Orientation	Gabor wavelet

model, we later generate synthetic BW, EM, and MA noise by the proposed method with a different signal-to-noise ratio (SNR). Since the sampling rate of original source noises and ECG signals are 360 and 1000 Hz, the synthetic noises are re-sampled to 1000 Hz. Since we expect more noise in verification step, we therefore assume that $\sigma_v > \sigma_e$ (see Section II-B). Thus, re-optimization of IOMBA margins module in NA-IOMBA determines new margins based on feedback from this assessment. For our approach, we consider the mixed noises with SNR=5dB for ECG margin reconstruction.

C. ECG STRESS MODELING AND REDUCING ENROLLMENT TIMES

Another concern that restricts the use of ECG for biometric authentication is the variability of heart signal within the subjects. Heart rate varies with an individual's physiological and mental conditions. Stress, excitement, exercise, and other working activities may have an impact on the heart rate and can elevate it. These variations are likely to affect the reliability of ECG based key generation or authentication. A previous study [29] about the influences of physical exercise indicates that the ECG morphology is affected by exercise/stress. In other words, each peak (P, QRS, T) in the ECG may increase/decrease in amplitude, temporal location, etc. To cover the impact of stress/exercise on the reliability with different scenarios, we vary the dynamical model parameters by type ($\theta_i, \alpha_i, b_i \forall i$) (Eq. 8) and analyze the impact of each part of ECG waveform. In Fig. 4 (b), we illustrate how noise in ECG features from stress/exercise can be handled using the model in NA-IOMBA. In order to assess the impact of stress in NA-IOMBA, NA-IOMBA trains itself with the help of the information from the standard deviation of stress ECG signal model and re-optimizes IOMBA margins.

Each dynamical model parameter is scaled by a factor $(0.9-0.5)$. Fig. 5(a-c) show how the ECG changes when scaling α, b, and θ parameters respectively. As shown in Fig. 5 (a), the α parameter controls amplitudes of each component of the ECG waveform. In contrast, the onset and offset of ECG waveform and interval duration are associated with scaling b and θ parameters. We intend to investigate the impact of dynamic parameters on each ECG waveform component to analyze the impact of stress & exercise on the reliability of IOMBA/NA-IOMBA.

IV. NA-IOMBA CASE STUDIES

The improvements gained by the proposed approach will be initially demonstrated in this section.

A. COMPARISON OF MULTIPLE MODALITIES USING IOMBA AND NA-IOMBA

In this section, we present a comprehensive performance evaluation of biometric-based key generation. We apply our approaches (IOMBA & NA-IOMBA) on four biometric modalities: ECG, PPG, iris, and fingerprint. Table 1 shows the methodologies, databases, and train/test sizes, that have been employed for multiple biometric modalities.

1) ELECTROCARDIOGRAM (ECG)

ECG is a recording of the electric potential, generated by the electric activity of the heart. The ECG recordings of 52 subjects from the PTB database [15] are used in this paper. We employ low and high pass finite impulse response (FIR) filters with cut off frequencies 1Hz-40Hz to eliminate noise associated with an ECG signal (see Fig. 6). Normalize-Convoluted Normalize (NCN) is used as the feature extraction technique [24].

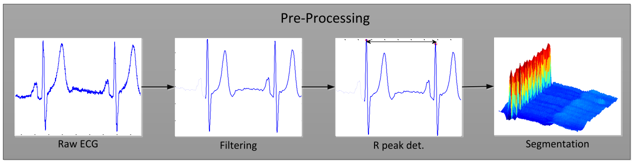
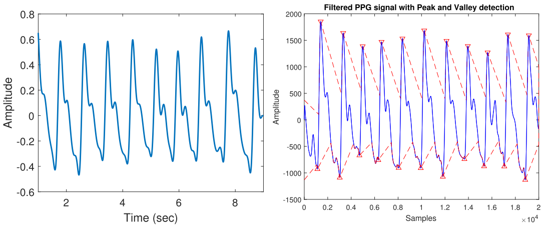


FIGURE 6. ECG Pre-processing. Raw data is acquired from database and then butterworth filtered; the individual heartbeat waveforms are segmented by their R peaks.



(a) (b)

FIGURE 7. Plots illustrating PPG signal from the database for (a) filtered PPG signal where the baseline and other artificial noise sources have been removed. (b) Extracted systolic and diastolic peaks by using pan tompkins algorithm.

2) MULTI-ECG

The ECG-ID database at the PhysioNet [15] was used as a multi-ECG in our experiments. Each raw ECG record was acquired for about 20 seconds with a sampling rate of 500 Hz and 12-bit resolution. First two records acquired from the same day were used for each subject. The database consists of 310 one-lead ECG recording sessions obtained from 90 volunteers during a resting state. The number of sessions for each volunteer varied from 2 to 20 with a time span of 1-day to 6-months between the initial and last recordings. This study utilizes the same pre-processing aforementioned for ECG PTB database.

3) PHOTOPLETHYSMOGRAM (PPG)

The photoplethysmogram (PPG) is a biomedical signal that estimates volumetric blood flow changes in peripheral circulation using low-cost and simple LED-based devices typically placed on the fingertips. In order to evaluate the efficiency of the PPG biometric authentication based on IOMBA and NA-IOMBA, a publicly available Capnobase dataset [30] with 42 subjects was used.

1) **Pre-processing:** Typically PPG signal is interfered by several noise sources including baseline wander (BW), motion artifact (MA), and respiration. To remove this artifact, we applied a third order Butterworth band pass filter with cutoff frequency 1Hz-5Hz as can be seen in Figure 7. Then, filtred PPG signal is passed by Pan Tompkins algorithm to extract systolic and diastolic peaks.
2) **Feature Extraction** In this section the feature extraction methods based on non-fiducial approach which will be discussed below.
3) **PPG Non-Fiducial Features:** Since the PPG signal is effected by noise, we applied non-fiducial feature extraction based on overall morphology of waveform rather than specific fiducial points. To this end, wavelet transform technique which is a very popular technique for biomedical signal processing due to the fact that it is lightweight and capable of providing time and frequency information simultaneously. The PPG signal will be passed through a series of low and high pass filters by decomposing it into various scales.

The discrete wavelet transform (DWT) is defined by

$$y[n] = \sum_{k=-\infty}^{\infty} s[k]\psi[n-k] \tag{10}$$

where the $s[k]$ indicates the PPG signal. Typically, DWT is derived by mother wavelet $\psi(t)$ which is expanded by value $s = 2^j$, translated by constant $\tau = k \times 2^j$, and normalized, where the j, k are integers. A wavelet defined by the solution of a dilation follows [47]

$$\psi_{j,k}[t] = \frac{1}{\sqrt{s}}\psi[\frac{t-\tau}{s}] = \frac{1}{\sqrt{2^j}}\phi[2^{-j}t-k] \tag{11}$$

where, j is the dilation parameter, or the visibility in frequency, and k is the parameter about the position.

The wavelet coefficients can be obtained by taking the inner product:

$$V_\phi[j_0, k] = \frac{1}{\sqrt{M}}\sum_n PPG[n]\phi_{j_0,k}[n] \tag{12}$$

$$W_\psi[j_0, k] = \frac{1}{\sqrt{M}}\sum_n PPG[n]\psi_{j,k}[n] \quad j_0 \leq k \tag{13}$$

where $\phi_{j_0,k}[n]$ and $\psi_{j,k}[n]$ are discrete functions. $\{\phi_{j_0,k}[n]\}_{k\in z}$ and $\{\psi_{j,k}[n]\}_{(j,k)\in z^2, j\leq j_0}$ are orthogonal to each other. Equation 12 represents approximation coefficients (CA) while Equation 13 demonstrates detailed coefficients (CD). In this paper, CA and CD are used as the non-fiducial feature vectors. We found that Coiflet mother wavelet is the best for key generation.

4) IRIS SCAN

The iris is called the colored ring around the eye pupil. According to research, the human eye is one of its most unique characteristic that can be used for biometric recognition. To evaluate the iris key generation based on IOMBA and NA-IOMBA, we first take the iris images from available CASIAv1-Interval iris database [16]. In the pre-processing stage, Canny edge detection is used to enhance the iris outer boundary due to the eyelid or eyelashes. Detecting the boundary of iris and sclera is applied for segmentation. Then, the iris is converted to a 2D matrix represented by (r, θ) which is the polar position of the original pixel of image. Finally, two-dimensional Gabor Filters are utilized for feature extraction [33]. Open-source OSIRIS package is used in this work [49]

TABLE 2. **IOMBA/NA-IOMBA results for four biometric modalities.**

		Key length			Reliability			Min-entropy			EER		
		Max	Ave	Min	Max	Ave	Min	Max	Ave	Min	Max	Ave	Min
ECG	NA-IOMBA	1247	953	784	100	98.76	96.17	1.00	0.984	0.905	3.84	1.25	0
	IOMBA	976	668	512	100	94.92	80.021	1.00	0.982	0.890	19.98	5.08	0
Multi-ECG	NA-IOMBA	851	409	153	100	98.47	91.61	1.00	0.913	0.630	8.40	1.43	0
	IOMBA	513	339	127	99.95	92.27	77.95	1.00	0.880	0.397	23.10	7.75	0.06
PPG	NA-IOMBA	195	107	17	100	99.04	91.11	0.99	0.863	0.623	8.90	0.97	0
	IOMBA	175	114	14	100	84.54	79.32	0.98	0.825	0.603	20.69	15.46	0
Iris	NA-IOMBA	1136	556	128	100	98.36	93.42	1.00	0.938	0.787	6.61	1.70	0
	IOMBA	204	66	23	98.92	90.88	72.67	1.00	0.869	0.756	27.30	9.12	1.09
Finger	NA-IOMBA	3567	1004	321	99.54	98.76	91.72	1.00	0.822	0.491	8.31	1.96	0.65
	IOMBA	1144	497	202	98.01	82.05	75.38	0.99	0.757	0.326	24.42	17.95	2.01

5) FINGERPRINT

Fingerprint is a unique pattern of ridges and valleys that have been used widely in biometric application. The fingerprint used for biometric key generation in the study is taken from FVC2004 database [34]. This data set, containing 8 images of 110 users. Images are aligned according to a standard core point position, in order to avoid a one-to-one alignment. In this paper, the Gabor filter is used to directly extract fingerprint features from gray level images [35]. The raw measurements contain two categories: the squared directional field in both x and y directions, and the Gabor response in 8 orientations.

The quality of generated keys is compared by four evaluation criteria: reliability, entropy, key length, and equal error rate (EER). The metrics used for each are discussed below and a brief comparison of the above biometric modalities based on IOMBA and NA-IOMBA is provided in Tables 2. In the table, 'max', 'ave', and 'min' columns correspond to the highest value (best case) achieved among all users, the average of keys across the users, and the lowest value obtained among all users. Note that for this initial comparison, the noise model in NA-IOMBA is adopted from the standard deviation of enrollment measurements and the standard deviation is adjusted on a per feature basis. A more elaborate model will be used for ECGs in the next section.

6) RELIABILITY

Reliability of key generation represents the stability of keys over time. If all bits generated by the biometric of an individual are equal to the key produced in enrollment, it can be considered reliable. We adopt the reliability metrics from [21]. As can be seen in Table 2, improvements in reliability are achieved by applying the NA-IOMBA technique. Average and worst cases improve by 2% and 9.7% on average for all modalities compared to IOMBA. Among all modalities, fingerprint attains the largest percentage of improvements (3.8% and 26.9% on average and worst cases). However, ECG has the best performance for both NA-IOMBA and IOMBA.

7) ENTROPY

To measure key randomness, we calculate the min-entropy. We adopt the min-entropy metrics from [21]. As shown in Tables 2, the min-entropy of ECG signal is higher than iris, PPG and fingerprint, however, under NA-IOMBA technique, there is a huge entropy improvement for iris, fingerprint and PPG compared to IOMBA results. For example, the min-entropy is not only improved by 35% at the minimum case for fingerprint based on NA-IOMBA but also increased by 8% on average.

8) KEY LENGTH

Since certain features may be reliable for some users and unreliable for others, our approach will only use reliable features from each individual. Thus, the key length per person may change. As can be indicated in Table 2, the key length of ECG, PPG, iris, and fingerprint based on IOMBA are 668, 114, 66, and 835, respectively. When NA-IOMBA is applied, the average key length for ECG, PPG, iris, and fingerprint increases by 30%, 6%, 88%, and 27%. Fingerprint obtains the largest key for both IOMBA and NA-IOMBA while PPG obtains the smallest.

a: EQUAL ERROR RATE (EER)

EER is a step in the biometric security system that determined by the threshold values for its false reject rate (FRR) and false accept rate (FAR). FAR refers to the rate at which an impostor user incorrectly identified as a genuine user, while FRR refers to the rate at which a genuine user incorrectly identified as an impostor user. An ideal biometric system, the EER is close to zero. To better understand how we can calculate EER from our IOMBA/NA-IOMBA, Fig. 8 is demonstrated. As can be seen in this figure, the probability density function (PDF) of the impostor is plotted with the red solid line and PDF of genuine is plotted with the blue solid line. Minimizing the EER is equivalent to minimizing two areas shaded as indicated in Fig. 8 corresponds to the FRR and FAR. Improving the biometric system authentication performance requires diminishing in this area.

According to Poh and Bengio [38], we can consider $P(y|x \in x_G)$ as a score's probability density function for the genuine user G and $P(y|x \in x_I)$ similarly for the impostor user I. Since these PDFs are Gaussian [21], FRR and FAR can be defined as follows [38]:

$$FRR(\theta) = \int_{-\infty}^{\theta} P(y|x \in x_G)dy$$

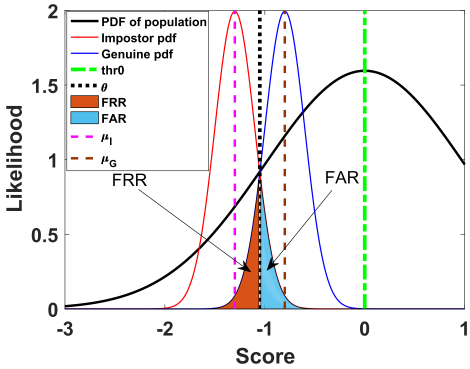


FIGURE 8. Averaging score distributions for calculating equal error rate.

$$
\begin{aligned}
&= \int_{-\infty}^{\theta} \frac{1}{\sigma_G\sqrt{2\pi}} \exp \frac{-(y-y_G)^2}{2\sigma_G^2} dy \\
&= \frac{1}{2} + \frac{1}{2} erf \frac{\theta-\mu_G}{\sigma_G\sqrt{2}} \qquad (14)
\end{aligned}
$$

$$
\begin{aligned}
FAR(\theta) &= \int_{\theta}^{\infty} P(y|x \in x_I) dy \\
&= 1 - \int_{-\infty}^{\theta} P(y|x \in x_I) dy \\
&= 1 - (\frac{1}{2} + \frac{1}{2} erf \frac{\theta-\mu_I}{\sigma_I\sqrt{2}}) \\
&= \frac{1}{2} - \frac{1}{2} erf \frac{\theta-\mu_I}{\sigma_I\sqrt{2}} \qquad (15)
\end{aligned}
$$

where the mean value and standard deviation of scores corresponding to the μ_G and σ_G for genuine user G and similarly μ_I and σ_I for the impostor user I. That being said, as we mentioned earlier, in biometric system authentication the goal is to minimize EER and the minimal error occurs when $\text{FAR}(\theta) = \text{FRR}(\theta) = \text{EER}$, therefore:

$$
EER = \frac{1}{2} - \frac{1}{2} erf \frac{\mu_G - \mu_I}{(\sigma_I + \sigma_G)\sqrt{2}} \qquad (16)
$$

For one session ECG signal (PTB database), we have constructed an average of 668 and 953 key bits with 5.08% and 1.25% EER based on IOMBA and NA-IOMBA. We achieve an EER around 15.46% based on IOMBA for PPG database while by incorporating NA-IOMBA model, the EER is decreased by 93%. In addition, the EER is decreased by 81% and 89% for fingerprint and iris database after employing NA-IOMBA approach.

9) MACHINE LEARNING VS QUANTIZATION

The key element of traditional biometrics authentication system is driven by machine learning, deep learning that makes it possible to drive the decision making processes based on given input data from original biometric templates or features that has been extracted during pre-processing to deal with the intra-class variation of biometric measurement. As indicated earlier, we propose a noise aware quantization approach to enhance the system security and user privacy on the resource-constrained IoT application. Existing quantization techniques invariably result in loss of some discriminatory information leading to lower recognition performance. Moreover, we compare the performance of IOMBA/NA-IOMBA with other state-of-art methods for biometric authentication. In order to compare our approach with machine learning, we apply a one-class support vector machine (SVM) technique for classification as a biometric authentication matching. Similar to most of the existing works for biometric authentication, we evaluated the accuracy and EER by applying machine learning techniques. Receiver operating characteristic (ROC) curves are shown in Fig. 9 for the biometric modalities considered in this paper. As shown in this figure, the average accuracy of ECG, multi-ECG, PPG, Fingerprint, and Iris are 98.35%, 94.09%, 98.26%, 92.63%, and 86.90%; the average EERs are 1.75%, 7.99%, 1.81%, 8.83%, and 20.68%, respectively; while performing five-fold cross-validation.

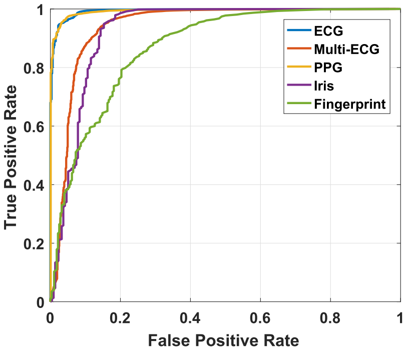


FIGURE 9. ROC curves for different biometric modalities.

Interestingly, we were able to approach a performance of 98.76%, 98.47%, 99.04%, 98.36%, and 98.76% in accuracy and 1.25%, 1.43%,.97%, 1.7%, and 1.96% in EER based on NA-IOMBA method for ECG, Multi-ECG, PPG, fingerprint, and iris, respectively, which is significantly better than machine learning results based on Table 2. In other words, we perform as well and in many cases better than off-the-shelf machine learning. In addition, quantization is a key step in signal/image system, especially in the big data age, because quantization not only can decrease storage costs but also can accelerate the detection speed in a large-scale database. Finally, machine learning techniques require many training samples per subjects in order to determine user-specific feature while in NA-IOMBA, we have shown that it is independent of training samples.

10) COST

The effectiveness of biometric technology is dependent on how and where it is used. Each biometric modality has its own strengths and weaknesses. Today, an ECG or a PPG sensor costs around $20 when ordered in large quantities, thus has a marginal cost of embedding into a biometric system. However, fingerprint and iris scan costs about $70 and $280,

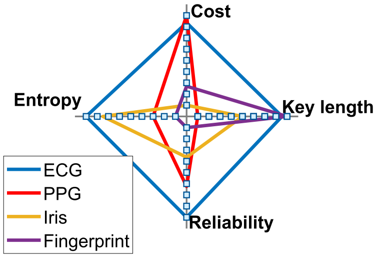


FIGURE 10. Star graph for comprehensive comparison among four biometrics using NA-IOMBA.

respectively. Note that the hardware cost is normalized into 1 in order to make it simpler to consider as metrics. In that case, if the value is lower than 1; meaning a more expensive sensor.

Figure 10 ranks four common technologies (ECG, PPG, iris scan, finger scan) according to four criteria: reliability, entropy, key length, and hardware cost. The maximum point in each length indicates the best candidate for that specific criteria. As can be seen in Figure 10, the average reliability of ECG is 99.76% belong to the maximum point of the plot while the average reliability of iris is 98.36% which belongs to the minimum point of the plot. In addition, the entropy of ECG is on the maximum point of the plot. For the cost, PPG is the best choice among all biometric modalities. Furthermore, the key length of the fingerprint is higher than other biometric modalities which made it become at top of the plot, although the ECG signal is following this with a small margin. ECG appears to be the best candidate for all the criteria when applying NA-IOMBA. However, it is worth noting that ECG still suffers from several other issues (impact of noises, stress condition, and aging) that need to be tackled in order to make this candidate an even stronger selection. The rest of this paper more closely examines the impact of noise on ECG with and without incorporation of noise models.

11) ECG LONG-TERM FEASIBILITY

We studied the impact of long-term variability of ECG signal called multi-session ECG signal from various days and obtained the number of training and testing heartbeats on ECG biometric performance. To this end, we have considered the case whereas the training feature sets contain subsets of ECG beats that has been extracted from a session, while the testing data come from another session. In this paper, the effect of changes on the ECG signal over time has been investigated. As can be seen in Table 2, the performance of ECG based on IOMBA technique has degraded while NA-IOMBA were incorporated to reduce the loss of performance. These results demonstrate that the evolution of the multi-session ECG signal is generated in a long time interval based on NA-IOMBA approach. All in all, results seem to confirm the long-term stability of reliability and EER. This is an essential condition since the threshold and boundaries in NA-IOMBA and IOMBA are the pre-configured parameter of the system, so if these results vary with time, the system will be incompatible and less useful for IoT healthcare applications.

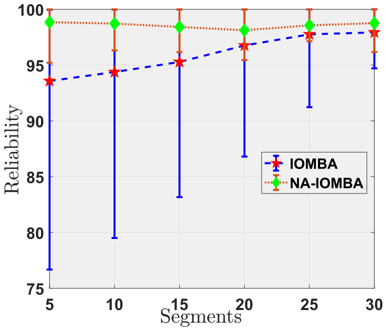


FIGURE 11. Reliability of IOMBA vs NA-IOMBA with different numbers of heart beats used during training.

12) IMPACT OF TRAINING SET SIZE

To understand the utility of ECG as a biometric, we examine the ECG biometric performances using a different number of training heartbeats. This helps to understand the sample size effect on ECG biometric performance for both IOMBA and NA-IOMBA. Fig. 11 shows the key reliability for testing on a session, when training is performed on a different number of heartbeats. In general, the variation in key reliability increases as the number of samples in the training increases. This is in agreement with our expectation. The average key reliability is increased when the number of training heartbeats or segments increases for IOMBA technique. However, the average key reliability for NA-IOMBA is already saturated at 5 samples. We observe the NA-IOMBA is independent of training heartbeats; hence enrollment time is reduced.

The maximum, average, and minimum accuracy value for a training sample size $N = 5$ are 99.95%, 93.57%, and 76.67% for IOMBA model; and 100%, 98.84%, and 95.21% for NA-IOMBA. As can be seen in Fig. 11, the accuracy of key based on IOMBA is improved by using 10 samples in the training, where 99.96%, 94.37%, 79.50% accuracy has been obtained. As we expected, the accuracy of keys did not change significantly by incorporating NA-IOMBA model. Finally, at the maximum number of training sample $N = 30$, 99.96%, 94.37%, 79.50% accuracy for IOMBA, 100%, 97.93%, and 94.71% for NA-IOMBA are achieved. As shown in Fig. 11, the NA-IOMBA performance is independent of the number of samples in a training set. This characteristic helps the biometric system to become faster and more robust.

V. CASE STUDIES ON ECG USING NOISE MODELS

In this section, we follow the approach described in Section III be incorporating ECG noise models. Figure 12 is a box plot showing the reliability rate versus input signal-to-noise ratio (SNR) for ECG based on IOMBA without denoising filters, IOMBA with denoising filters, and NA-IOMBA without denoising filters. The SNR during the noisy segments was set to 30dB, 20dB, 10dB, 5dB, 0dB, and −5dB separately. Figures 12 (a-d) indicate the impact of

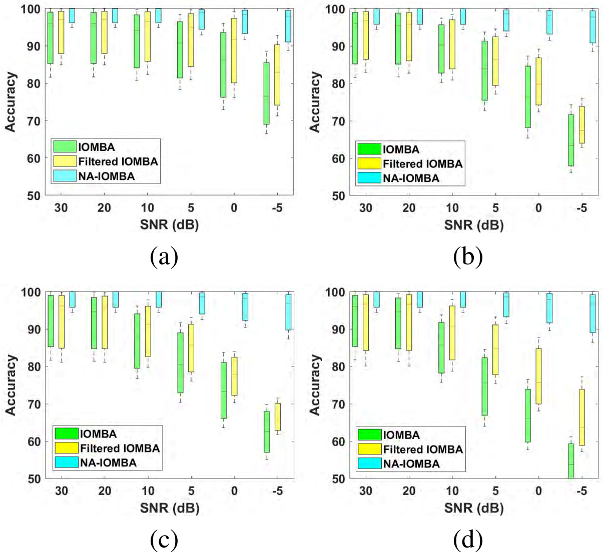


FIGURE 12. IOMBA, filtered IOMBA, and NA-IOMBA Keys reliability rate vs input SNR; impact of (a) BW noise, (b) EM noise, (c) MA noise, and (d) mixed noises, on the reliability.

each noise source (BW, EM, and MA, and mixed of them) on the reliability. In the context of ECG noise levels, the lower SNR provides more fluctuation on ECGs (higher intra-class variation). Intuitively, there appears to be an inverse relationship between the level of generated noise and reliability. Among all these noise sources, MA and EM are the strongest noise sources and have an enormous impact on key reliability when IOMBA is applied with and without filtering. As one would expect, IOMBA with filtering obtains better key reliability than without. However, for NA-IOMBA, the reliability is never less than 96.7% even at worst case (mixed noise with -5dB). In contrast, there is considerable degradation beyond 20dB by using IOMBA with/without filtering (63% reliability). As mentioned earlier, ECC increases nonlinearly with the number of errors. NA-IOMBA has very high reliability compared to IOMBA, and therefore ECC will inherently consume less overhead. The cost reduction is discussed further below.

A. FPGA IMPLEMENTATION OF AN IOMBA

In this paper, finite impulse response (FIR) is designed using Simulink in the Xilinx System Generator. The Xilinx System Generator tool is a high-level tool for designing high-performance DSP systems and enables us to integrate Xilinx with Simulink. To implement noise reduction using FIR filter, an FDA tool has been applied to design a filter for required specifications. Pan Tompkins algorithm is applied for detecting ECG R peak and segmentation. Finally, the NCN feature extraction technique has been considered for key generation. Table 3 shows implementation of the ECG key generation using IOMBA with filtering and NA-IOMBA without filtering on the Xilinx Zynq-7000. In IOMBA case, 11% of total flip-flops (FF), 20% of all available Look-up tables (LUTs), and 71% of the DSP slice are used while in NA-IOMBA consumes only 1% of FFs, 4% of LUTs, and 10% of DSP are utilized. In addition, IOMBA consumes 113 mW power while NA-IOMBA consumes only 39 mW power. As a result, by saving overall overhead while applying NA-IOMBA, we are able to add ECC in the IoT devices to reconstruct the errors. In fact, NA-IOMBA allows hardware to adapt pre-processing, feature extraction, post-processing, and error correction overheads on a user-to-user basis.

TABLE 3. Hardware utilization report on the Zynq-7000 SoC XC7Z020.

Resource	IOMBA Usage	NA-IOMBA Usage	Available
Flip Flops	11830	1064	106400
Look-Up Tables	10386	2232	53200
DSP Slices	157	21	220
I/O	32	20	200

B. STRESS/EXERCISE RESULTS

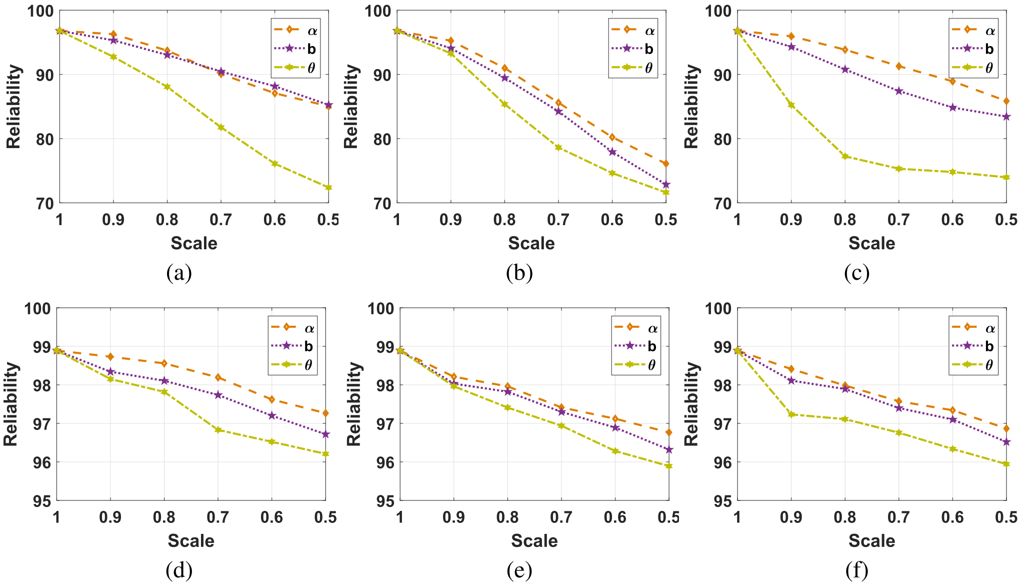
Fig. 13((a-c) indicate the reliability of each dynamic parameters on the P, QRS, and T waves based on IOMBA, respectively. The reliability of NA-IOMBA is shown in Fig. 13(d-f) where IOMBA margins are re-optimized assuming dynamic parameters scaled to 0.7. For IOMBA, the T wave is impacted by all parameters (α, b, and θ). At lowest scale value (0.5), the minimum reliability of dynamic parameters of T wave is 71.61% while for P and QRS wave are 72.37%, and 73.96% respectively. θ has a larger impact than other parameters for P and QRS waves because it causes distortion in ECG time intervals (distances between peaks). Even though there is too much degradation on the reliability for IOMBA technique, the performance of NA-IOMBA in stress/exercise situation is much higher than IOMBA. The minimum reliability of dynamic parameter on P, QRS, and T waves (Fig. 13(d)) are improved by 25%. Note, however, that the key length is sacrificed by 56% (420 average key bits) to obtain this improvement. We expect that this will be long enough for most cryptographic applications.

VI. SECURITY ANALYSIS

Frankly speaking, measuring the entropy of key bits is one measures the strength of a cryptographic, which quantifies the amount of uncertainty in the key from an adversary's standpoint [39]. We also evaluate the security requirements for biometric key generation based on key randomness.

A. KEY RANDOMNESS BASED ON NIST TEST

The Bio-key generated through our scheme need appear as random to an adversary which has access to the auxiliary information. To evaluate the randomness, the NIST statistical test suite is applied [41]. The NIST Test Suite (NTS) is a statistical test consisting of different types of tests to evaluate the randomness of binary sequences. Each statistical test is employed to calculate a p-value that shows the randomness of the given sequences based on that test. If a p-value for a test is determined to be equal to 1, then the sequence appears to have perfect randomness. A p-value $\geq$ 0.01



(a) (b) (c)

(d) (e) (f)

FIGURE 13. Impact of stress/exercise on the reliability of; IOMBA for changing (a) P wave, (b) QRS wave, (c) T wave; NA-IOMBA for changing, (d) P wave, (e) QRS wave, and (f) T wave of ECG.

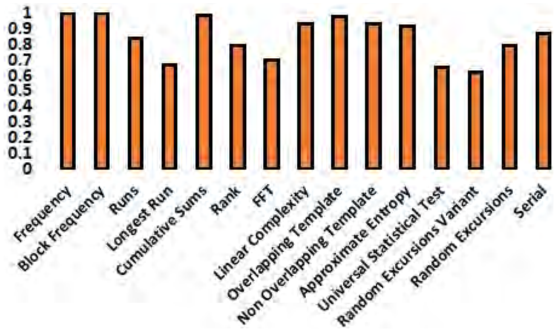


FIGURE 14. NIST statistical tests Suite results for the randomness tests of proposed biometric key generation.

(normally 1%) means that sequence would be considered to be random with confidence of 99%. The results of 15 performed NIST tests on our proposed key generation are shown in Fig. 14 that showing proposed key generation passes the randomness tests.

B. ENTROPY ANALYSIS

For the highest level of security, the keys used should be drawn from a uniform distribution since this will result in the greatest number of brute force attempts needed to identify the correct key by an attacker. Entropy is one of the most widely used measures of such randomness or unpredictability in keys. The entropy should be large enough to guarantee resistance against attacks. In this paper, each key has been generated from our proposed algorithm and the randomness of each bit is estimated by calculating the min-entropy, which is the most conservative way of estimating entropy or unpredictability. In this paper, the min-entropy of a feature k is calculated as follows

$$H_{\infty}(k) = -\xi \log_2(\underset{i}{max}\{P_i(k)\}) \tag{17}$$

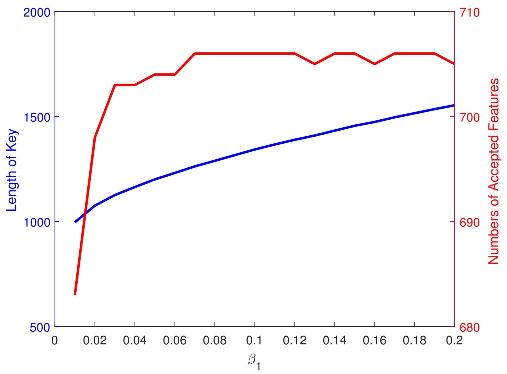


FIGURE 15. The key bits length and number of accepted features for the template.

where $P_i(k) = Pr(X = i)$ for the kth feature of a subject. Note that $i \in \{0, 1\}$,$\{00, 01, 10, 11\}$, and $\{000, 001, 010, 011, 100, 101, 110, 111\}$ for features that are quantized to one bit, two bits, and three bits respectively from our proposed NA-IOMBA and IOMBA. ξ is a normalizing parameter set equal to 1, $\frac{1}{2}$ and $\frac{1}{3}$ for 1-bit, 2-bit, and 3-bit cases. Note that the maximum min-entropy (1) occurs when $P_i = \frac{1}{2^n} \forall i$, where the n indicates number of bits for quantization. The results of entropy for each biometric modalities are measured in Table 2. From these results which are close to ideal case (1), it does not depict the vulnerability of our proposed system.

C. TEMPLATE IRREVERSIBLE

The requirement of a template being irreversible is that the dimensionality of feature set should be smaller than dimensional subspace [6]. In our approach, each user has its own template due to a user-specific algorithm. In other words, none of all feature vector will be selected for key generation.

As can be seen in Fig 15, the number of feature sets that have been accepted in our algorithm versus the number of total feature sets is different. In addition, if a feature x has been selected in subject one does not guarantee that the same feature will be selected for other subjects. Moreover, there is no correlation between any two templates, thus will reside in different subspace and they are distinguishable from each other.

VII. CONCLUSIONS

In this paper, the interval optimized mapping bit allocation (IOMBA) scheme for the key generation was improved by incorporating noise models. It was demonstrated that keys generated from ECG, PPG, iris, and fingerprint by noise-aware IOMBA are more reliable, longer, and higher entropy than noise-free IOMBA. Furthermore, by using more advanced noise models for ECG, overhead from denoising filters and error correction could be further reduced by 62% without additional enrollment measurements. Moreover, we analyze our model under the multiple-session ECG signal where a single session is used to train our model and testing data come from different sessions. Besides, we also compared the performance of our noise-aware biometric quantization framework with other state-of-art machine learning techniques. In the future, we intend to develop an end-to-end framework ensuring to protect our model from vulnerabilities to attacks. In addition, revocability, unlinkability, and irreversibility will be an interesting and challenging work that deserves our collective efforts in the future.

REFERENCES

[1] M.-H. Lin, W.-L. Yuan, T.-C. Huang, H.-F. Zhang, J.-T. Mai, and J.-F. Wang, "Clinical effectiveness of telemedicine for chronic heart failure: A systematic review and meta-analysis," *J. Invest. Med.*, vol. 65, no. 5, pp. 899–911, 2017.

[2] E. J. Kelkboom, G. G. Molina, J. Breebaart, R. N. Veldhuis, T. A. Kevenaar, and W. Jonker, "Binary biometrics: An analytic framework to estimate the performance curves under gaussian assumption," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 40, no. 3, pp. 555–571, May 2010.

[3] M. Clarke, J. de Folter, V. Verma, and H. Gokalp, "Interoperable end-to-end remote patient monitoring platform based on IEEE 11073 PHD and ZigBee health care profile," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 5, pp. 1014–1025, May 2018.

[4] R. J. Bolton and D. J. Hand, "Statistical fraud detection: A review," *Stat. Sci.*, pp. 235–249, 2002.

[5] U. Srinivasan and B. Arunasalam, "Leveraging Big Data Analytics to Reduce Healthcare Costs," *IT Prof.*, vol. 15, no. 6, pp. 21–28, Nov. 2013.

[6] S.-C. Wu, P.-T. Chen, A. L. Swindlehurst, and P.-L. Hung, "Cancelable biometric recognition with ECGs: Subspace-based approaches," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1323–1336, May 2019.

[7] R. K. Michaels *et al.*, "Achieving the national quality forum's 'never events': Prevention of wrong site, wrong procedure, and wrong patient operations," *Ann. Surg.*, vol. 245, no. 4, p. 526, 2007.

[8] S. Pirbhulal, H. Zhang, W. Wu, S. C. Mukhopadhyay, and Y.-T. Zhang, "Heartbeats based biometric random binary sequences generation to secure wireless body sensor networks," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 12, pp. 2751–2759, Dec. 2018.

[9] Z. Jin, A. B. J. Teoh, B.-M. Goi, and Y.-H. Tay, "Biometric cryptosystems: A new biometric key binding and its implementation for fingerprint minutiae-based representation," *Pattern Recognit.*, vol. 56, pp. 50–62, Aug. 2016.

[10] D. Karakoyunlu and B. Sunar, "Differential template attacks on PUF enabled cryptographic devices," in *Proc. IEEE Int. Workshop Inf. Forensics Secur.*, Dec. 2010, pp. 1–6.

[11] G. Zheng, W. Li, and C. Zhan, "Cryptographic key generation from biometric data using lattice mapping," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, vol. 4, Aug. 2006, pp. 513–516.

[12] P. Tome, J. Fierrez, R. Vera-Rodriguez, and M. S. Nixon, "Soft biometrics and their application in person recognition at a distance," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 3, pp. 464–475, Mar. 2014.

[13] L. Best-Rowden and A. K. Jain, "Learning face image quality from human assessments," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 12, pp. 3064–3077, Dec. 2018.

[14] D. Thakkar. *Use of Biometrics for Accurate Patient Identification*. Accessed: Jul. 1, 2016. [Online]. Available: https://www.bayometric.com/biometrics-accurate-patient-identification

[15] *Physiobank*. Accessed: Aug. 1, 2018. [Online]. Available: http://physionet.org

[16] *CASIA-IrisV1*. Accessed: Aug. 1, 2018. [Online]. Available: http://biometrics.idealtest.org

[17] X. Nie, X. Li, Y. Chai, C. Cui, X. Xi, and Y. Yin, "Robust image fingerprinting based on feature point relationship mining," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 6, pp. 1509–1523, Jun. 2018.

[18] K. Uemura, T. Kawada, C. Zheng, and M. Sugimachi, "Less invasive and inotrope-reduction approach to automated closed-loop control of hemodynamics in decompensated heart failure," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 8, pp. 1699–1708, Aug. 2016.

[19] M. Sugimachi and K. Sunagawa, "Bionic cardiology: Exploration into a wealth of controllable body parts in the cardiovascular system," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 172–186, 2009.

[20] E. J. da Silva Luz, G. J. P. Moreira, L. S. Oliveira, W. R. Schwartz, and D. Menotti, "Learning deep off-the-person heart biometrics representations," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 5, pp. 1258–1270, May 2018.

[21] N. Karimian, Z. Guo, M. Tehranipoor, and D. Forte, "Highly reliable key generation from electrocardiogram (ECG)," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 6, pp. 1400–1411, Jun. 2017.

[22] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289–294, Mar. 2003.

[23] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, "A nonlinear Bayesian filtering framework for ECG denoising," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2172–2185, Dec. 2007.

[24] N. Karimian, F. Tehranipoor, Z. Guo, M. Tehranipoor, and D. Forte, "Noise assessment framework for optimizing ecg key generation," in *Proc. IEEE Int. Symp. Technol. Homeland Secur. (HST)*, Apr. 2017, pp. 1–6.

[25] R. Sameni, C. Jutten, and M. B. Shamsollahi, "Multichannel electrocardiogram decomposition using periodic component analysis," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 8, pp. 1935–1940, Aug. 2008.

[26] G. B. Moody, W. K. Muldrow, and R. G. Mark, "A noise stress test for arrhythmia detectors," *Comput. Cardiol.*, vol. 11, no. 3, pp. 381–384, 1984.

[27] C. Li and J. Hu, "A security-enhanced alignment-free fuzzy vault-based fingerprint cryptosystem using pair-polar minutiae structures," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 3, pp. 543–555, Mar. 2016.

[28] R. Sameni, G. D. Clifford, C. Jutten, and M. B. Shamsollahi, "Multichannel ECG and noise modeling: Application to maternal and fetal ECG signals," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, 2007, Art. no. 043407.

[29] P. A. Friedman, K. E. Bennet, C. J. Bruce, and V. K. Somers, "Non-invasive monitoring of physiological conditions," U.S. Patent 9 907 478, Mar. 6, 2018.

[30] W. Karlen, S. Raman, J. M. Ansermino, and G. A. Dumont, "Multiparameter respiratory rate estimation from the photoplethysmogram," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 7, pp. 1946–1953, Jul. 2013.

[31] M. Osadchy and O. Dunkelman, "It is all in the system's parameters: Privacy and security issues in transforming biometric raw data into binary strings," *IEEE Trans. Dependable Secure Comput.*, to be published.

[32] N. Karimian, M. Tehranipoor, and D. Forte, "Non-fiducial PPG-based authentication for healthcare application," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Inform.*, Feb. 2017, pp. 429–432.

[33] J. Daugman, "How iris recognition works," in *The Essential Guide to Image Processing*. Amsterdam, The Netherlands: Elsevier, 2009, pp. 715–739.

[34] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain, "FVC2004: Third fingerprint verification competition," in *Biometric Authentication*. London, U.K.: Springer, 2004, pp. 1–7.

[35] A. Ross, A. Jain, and J. Reisman, "A hybrid fingerprint matcher," *Pattern Recognit.*, vol. 36, no. 7, pp. 1661–1673, Jul. 2003.

[36] C. Chen, R. N. Veldhuis, T. Kevenaar, and A. Akkermans, "Multi-bits biometric string generation based on the likelihood ratio," in *Proc. 1st IEEE Int. Conf. Biometrics, Theory, Appl., Syst.*, Sep. 2007, pp. 1–6.

[37] M.-H. Lim, A. B. J. Teoh, and K.-A. Toh, "Dynamic detection-rate-based bit allocation with genuine interval concealment for binary biometric representation," *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 843–857, Jun. 2013.

[38] N. Poh and S. Bengio, "Why do multi-stream, multi-band and multi-modal approaches work on biometric user authentication tasks?" in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 5, May 2004, p. V–893.

[39] L. Ballard, S. Kamara, and M. K. Reiter, "The practical subtleties of biometric key generation," in *Proc. USENIX Secur. Symp.*, 2008, pp. 61–74.

[40] H. Kaur and P. Khanna, "Random distance method for generating unimodal and multimodal cancelable biometric features," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 3, pp. 709–719, Mar. 2019.

[41] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, and E. Barker, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," NIST, Gaithersburg, MD, USA, Tech. Rep., 2010.

[42] F. Monrose, M. K. Reiter, Q. Li, and S. Wetzel, "Cryptographic key generation from voice," in *Proc. IEEE Symp. Secur. Privacy*, May 2000, pp. 202–213.

[43] A. B. J. Teoh, D. C. L. Ngo, and A. Goh, "Personalised cryptographic key generation based on facehashing," *Comput. Secur.*, vol. 23, no. 7, pp. 606–614, 2004.

[44] P. Drozdowsk, F. Struck, C. Rathgeb, and C. Busch, "Benchaarking binarisation schemes for deep face templates," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 191–195.

[45] C. Chen and R. Veldhuis, "Binary biometric representation through pairwise adaptive phase quantization," *EURASIP J. Inf. Secur.*, vol. 2011, no. 1, 2011, Art. no. 543106.

[46] M.-H. Lim, A. B. J. Teoh, and K.-A. Toh, "An efficient dynamic reliability-dependent bit allocation for biometric discretization," *Pattern Recognit.*, vol. 45, no. 5, pp. 1960–1971, 2012.

[47] S. Mallat, *A Wavelet Tour of Signal Processing*. Amsterdam, The Netherlands: Elsevier, 1999.

[48] A. Nagar, K. Nandakumar, and A. K. Jain, "Biometric template transformation: A security analysis," *Proc. SPIE*, vol. 7541, Jan. 2010, Art. no. 75410O.

[49] N. Othman, B. Dorizzi, and S. Garcia-Salicetti, "OSIRIS: An open source iris recognition software," *Pattern Recognit. Lett.*, vol. 82, pp. 124–131, 2016.

NIMA KARIMIAN (S'15–M'18) received the Ph.D. degree from the University of Connecticut, in 2018. He is currently an Assistant Professor of computer engineering with San Jose State University. He has authored several papers in International Joint Conference on Biometrics (IJCB), ICASSP, TBME, TVLSI, BHI, and CODES+ISSS. His research interests include machine learning (ML), deep learning, pattern recognition, biometrics authentication and identification, optimization problem, signal processing, the Internet-of-Things (IoT), and hardware security primitives. He was a recipient of several best paper awards from venues such as International Joint Conference on Biometrics (IJCB) for the IAPR TC4 Best Student Paper Award, the Best Technical Paper Award in 30th International Conference on VLSI Design (VLSID), and the Best Poster Award from FICS Research Conference on Cybersecurity.

MARK TEHRANIPOOR (S'02–M'04–SM'07–F'18) received the Ph.D. degree from the University of Texas at Dallas, Richardson, TX, USA, in 2004. He was the Founding Director of the Center for Hardware Assurance, Security, and Engineering and the Comcast Center of Excellence for Security Innovation Centers, University of Connecticut. He is currently the Intel Charles E. Young Preeminence Endowed Professor of Cybersecurity with the University of Florida, Gainesville, FL, USA. He is also serving as the Co-Director of the Florida Institute for Cybersecurity Research. He has authored or coauthored over 300 journal articles and refereed conference papers and has given over 150 invited talks and keynote addresses. He has authored or coauthored six books and 11 book chapters. His current research interests include hardware security and trust, supply-chain security, and VLSI design, test, and reliability. He is a Golden Core Member of the IEEE and a member of the Association for Computing Machinery (ACM) and the ACM Special Interest Group on Design Automation. He was a recipient of several best paper awards, including the 2008 IEEE Computer Society (CS) Meritorious Service Award, the 2009 National Science Foundation CAREER Award, the 2012 IEEE CS Outstanding Contribution, and the 2014 Multidisciplinary University Initiative Award. He serves on the Program Committee of over a dozen of leading conferences and workshops. He served as the Program Chair of the 2007 IEEE Defect-Based Testing Workshop and the 2008 IEEE Defect and Data Driven Testing (D3T) Workshop, the Co-Program Chair of the 2008 International Symposium on Defect and Fault Tolerance in VLSI Systems (DFTS), the General Chair for D3T 2009 and DFTS 2009, and the Vice General Chair for the IEEE North Atlantic Test Workshop 2011. He cofounded the IEEE International Symposium on Hardware-Oriented Security and Trust (HOST) and served as the HOST-2008 and HOST-2009 General Chair. He is currently serving as a founding EIC for the *Journal on Hardware and Systems Security* (HaSS) and an Associate Editor for JETTA, JOLPE, IEEE TVLSI, and ACM TODAES.

DAMON WOODARD (S'01–M'04–SM'11) received the B.S. degree in computer science and computer information systems from Tulane University, in 1997, the M.E. degree in computer science and engineering from Pennsylvania State University, in 1997, and the Ph.D. in computer science and engineering from the University of Notre Dame, in 2005.

From 2006 to 2014, he was an Assistant Professor and then an Associate Professor with the School of Computing, Clemson University, Clemson, SC, USA. In 2015, he joined the Faculty of the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA, as an Associate Professor. He is also a member of the Florida Institute for Cybersecurity (FICS) Research. His research interests include biometrics, machine learning, pattern recognition, natural language processing, and signal/image analysis.

Dr. Woodard is an ACM Senior Member and the National Academy of Science Kavli Frontier Fellow. He currently serves as the Vice President of Finances for the IEEE Biometrics Council and a Program Chair for the 10th IEEE International Conference on Biometrics: Technology, Applications, and Systems (BTAS 2019).

DOMENIC FORTE (S'09–M'13–SM'18) received the B.S. degree in electrical engineering from the Manhattan College, Riverdale, NY, USA, in 2006, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland at College Park, College Park, MD, USA, in 2010 and 2013, respectively. He is currently an Assistant Professor with the Electrical and Computer Engineering Department, University of Florida, Gainesville, FL, USA. His current research interests include the domain of hardware security, including the investigation of hardware security primitives, hardware Trojan detection and prevention, electronics supply-chain security, anti-reverse engineering, and biometrics security. He was a recipient of the NSF CAREER Award, the Early Career Award for Scientists and Engineers (ECASE-Army) by Army Research Office, and the George Corcoran Memorial Outstanding Teaching Award by the Electrical and Computer Engineering Department, University of Maryland. His work has also been recognized through several best paper awards and nominations from venues, such as the HOST, IJCB, DAC, and AHS. He is also serving on the Organizing Committees of HOST and AsianHOST as well as the Technical Program Committee of several other top conferences. He is currently serving as an Associate Editor for the *Journal of Hardware and Systems Security*.