Is Your Smartband Smart Enough to Know Who You Are: Towards Continuous Physiological Authentication in The Wild

Deniz Ekiz, Yekta Said Can, Yağmur Ceren Dardağan and Cem Ersoy

Abstract—The use of cloud services that process privacy-sensitive information such as digital banking, pervasive healthcare, smart home applications require an implicit continuous authentication solution which will make these systems less vulnerable to the spoofing attacks. Physiological signals can be used for continuous authentication due to their personal uniqueness. Wrist-worn wearable devices such as smartwatches and smartbands became ubiquitous. They are equipped with photoplethysmogram (PPG) sensors which enable to extract heart rate variability (HRV) features. In this study, we show that these devices can be used for continuous physiological authentication, for enhancing the security of the cloud, edge services, and IoT devices. A system that is suitable for the smartband framework comes with new challenges such as relatively low signal quality and artifacts due to placement which were not encountered in full lead ECG systems. After the artifact removal, cleaned physiological signals are fed to the machine learning (ML) algorithms. In order to train our machine learning models, we collected physiological data using off-the-shelf smartbands and smartwatches in a real-life event. Performance evaluation of different machine learning algorithms shows that HRV is a strong candidate for continuous unobtrusive implicit physiological authentication.

Index Terms—smartband, smartwatch, heart rate variability, continuous authentication

I. INTRODUCTION

IMPLICIT continuous authentication is required for cloud oriented services which grant access to the privacy sensitive information domains such as mobile banking, pervasive healthcare [1], [2]. Smartphones, computers, smartwatches and Internet-of-Things (IoT) devices become more dependent on these services. It is expected that the number of IoT devices will be more than 75 Billion in 2025 [3]. However, these services are vulnerable to attacks once users authenticate. For example, a smartphone can be forgotten logged-in, the privacy-sensitive services and information can be stolen by the attackers. One simple mechanism can be asking a password to the user frequently. However, this is annoying for the consumer experience of these services. Continuous authentication should be implicit. Face-based systems can be tricked by using presence attacks, such as printing the face of the victim on a paper. Storing the face pictures of the users also create a privacy concern [4]. Furthermore, fingerprint which is another prominent traditional biometrics modality along with the face-based systems can be easily manipulated [5] and fails on liveness detection tests. On the other hand, biosignals are difficult to temper with and they inherently have liveness detection feature [6]. Heart activity is unique to individuals and biosignal authentication research has started investigating this signal [6]. One of the most important property of this signal is the Heart Rate Variability (HRV). Although research in the past mostly focused on the connection between HRV and different types of health disorders [7], the validity of using HRV for biometric recognition is supported by the fact that the physiological and geometrical differences of the heart in different individuals display certain uniqueness in their HRV features [8].

High-end wearable systems are expensive and provide low comfort for the users, which limit their wide range application. Recently, smart bands and smartwatches became adopted by consumers. These devices are equipped with a rich set of sensors such as accelerometer, heart rate monitor and skin conductance. These advances create an opportunity to build a continuous implicit authentication system. However, these devices are prone to activity related errors [9] unlike full lead ECG systems. Modality specific artifact detection and removal mechanism should be developed for accurate measurements. A solution suitable for IoT connected devices should be context-independent because every service may require different types of behavior. Therefore, systems that only work in certain scenarios such as while typing or walking are very limited in terms of the application area.

Deniz Ekiz, Yekta Said Can, Yağmur Ceren Dardağan and Cem Ersoy were with the Department of Computer Engineering, Boğaziçi University, Turkey.
E-mail: deniz.ekiz@boun.edu.tr
Manuscript received XXX, XXX
The physiological parameters like hidden heart-related biometrics are more suitable for this purpose due to their uniqueness and activity independence [10], [11]. We propose an unobtrusive, low cost, activity-independent continuous authentication system with smartbands. We implemented our solution on Empatica E4 Smartband [12], Samsung Gear S and Gear S2 [13] which are equipped with a photoplethysmography (PPG) sensor.

Let’s think about a scenario, where John has to make money transactions through a mobile banking smartphone application. He logins to the system via two-factor authentication. After he successfully makes the transaction, he forgets to logout. Then, an attacker withdraws all the cash from John’s account by using his session. If John has a continuous biometric authentication system, he would not have this problem. Such a system can be also used in services like city bikes and electric scooter rental services which have become popular in recent years. The literature on continuous authentication with physiological signals gathered from smartbands is limited. The effectiveness of HRV features derived from PPG sensors of smartwatches and smartbands are still unknown for continuous authentication [4]. Our proposed system enhances the state of the art as follows:

- Previous works employed mean heart rate per minute which requires longer recordings, and achieve a response time about five minutes [14]. The proposed system uses more sophisticated HRV features derived from PPG sensors of smartwatches and smartbands which are still unknown for continuous authentication [4]. Our proposed system enhances the state of the art as follows:
  - Previous works employed mean heart rate per minute which requires longer recordings, and achieve a response time about five minutes [14]. The proposed system uses more sophisticated HRV features derived from PPG sensors of smartwatches and smartbands which are still unknown for continuous authentication [4].
  - Most of the previous works with smartbands only focused on the continuous authentication while using a laptop, a smartphone, an ATM and used lifestyle related metrics. Our solution is not activity dependent. For example, a user may use this system during any activity such as live streaming, on social media, cycling, presentation or working in an office.
  - Our solution is comfortable, unobtrusive, seamless and works without interrupting the ordinary pattern of any activity.

In Section II we provide background information on the smartwatch framework and heart rate variability. In Section III the related work on continuous authentication and the comparison with our work in terms of its novelty are presented. In Section IV we describe the proposed system for continuous authentication with smartwatches using heart rate variability. In Section V we explain the conducted experiment for the proposed system. In Section VI we provide the results of our system. In Section VII we present the conclusion and future works of our study.

II. BACKGROUND

A. Smartband Framework

Smartbands (some people call them wristbands or smartwatches) are comfortable devices that can be attached to the wrist or arm. The devices used in this study are shown in Figure 1. Recently, the battery life of smartbands is increased to days. For example, the battery life of Empatica E4 is two days used all sensors are used [12]. Extended battery life of smartbands enables monitoring physiological signals gathered from individuals for long periods. Most of the wristbands are equipped with PPG which is an optically obtained signal that can be used to detect blood volume changes in the microvascular bed of tissue [15]. From the PPG signal, the time between the beats (RR intervals) can be computed. Most of the modern smartbands provide RR intervals thanks to their APIs. For example, Samsung smartwatches use the Tizen framework which has Human Activity Monitor API to gather the RR intervals. The sampling frequency of the PPG sensor can vary from 20Hz to 100Hz (64Hz for E4, 100Hz for Samsung Gear Series) in many smartbands. Cubic spline interpolation is used to detect the beats more accurately and most of the devices correct the heartbeats by using an accelerometer sensor for detecting the movements of the subject. This functionality is available in Tizen and Empatica API [12]. These devices are also equipped with Bluetooth and recently NFC chips which enable them to connect to smartphones, edge and cloud services. These short-range network interfaces can be used to check if the user is in the close proximity of the computer they are using.

![Fig. 1. Empatica E4, Samsung Gear S2 and Samsung Gear S are shown in the order.](image)
B. Heart Rate Variability

The variability of RR intervals is called HRV [16]. It is a very important feature for recognition of certain psychological, physiological and personal properties of an individual [16]. In the literature, Kubios [17] is a popular HRV feature extraction tool to compute the HRV. Non-linear, time domain and frequency domain features can define the variability of the heart [17]. The calculation of frequency domain and time domain features of HRV is computationally effective, thanks to Fast Fourier Transform (FFT) \( O(n \log n) \) for frequency domain and \( O(n) \) for time-domain features.

III. RELATED WORK

The behavioral and physiological biometrics from the wearable devices via sensors have become popular in individual recognition and authentication models. Some models focus on biometrics such as face [25], voice [26], fingerprints [27], [28], electroencephalography (EEG) [29], [30], [31], electrocardiograms (ECG) [32], [33], [28] and phonecardiography (PCG). [34]. The continuous authentication field is a fast-growing field, however, the literature on a system that is not dependent on a certain type of task is limited. Most of the previous work using physiological signals has been done on laboratory-grade equipment. Some of these sensors are not available in unobtrusive devices such as smartphones, smartwatches, and smart bands. On the other hand, fingerprints and face-based authentication systems can be easily deceived. Authentication with voice has privacy issues, which requires continuous recording of the environment.

As a method for recognizing individuals, Elkader et al. [18] presented a sensor-based motion biometric model that is suitable for 20 sedentary and non-sedentary activities (Vacuuming, Sweeping, Walking Downstairs, Walking Upstairs, Dusting, Iron Cloth, Folding Cloth, Washing Hands, Brushing Teeth, Washing Dishes, Washing Vegetables, Dicing, Peeling Vegetables, Grating, Stirring, Watching TV, Using PC, Talking on Phone, Texting on Phone, Writing with Pen). They used different combinations of 3 sensors (acceleration, gyroscope and magnetometer sensors) on 6 different body positions (dominant wrist, dominant upper arm, non-dominant wrist, chest, thigh, and ankle positions). They concluded that features extracted from the combination of six sensors reach the best classification accuracy in overall (98.3%). These activities are gathered in a laboratory environment with manual segmentation of the signals.

Another approach for implicit identification and authentication based on activity information, WearAI, Zeng et al. [19], proposed a biometric model that utilizes accelerometer and gyroscope sensors from five body locations such as left wrist (Shimmer 6DoF IMU), right ankle (Shimmer 6DoF IMU), center right hip/torso (Samsung Galaxy S4 i9500), left thigh/front pocket (Samsung Galaxy Nexus i9250), right upper arm (Samsung Galaxy Nexus i9250)). They achieved 97% accuracy with less than 1% false-positive rate. However, in both methods, placing many sensors on the body can be disturbing for the user in daily life usage.

Acar et al. [35] used smartwatches with key stroke dynamics for continuous authentication. Musale et al. [23] proposed a continuous authentication system based on Motorola 360 Sport by using accelerometer and gyroscope features. Vhaduri et al. [20] proposed continuous user authentication scheme that uses 44 features extracted from various biometrics (calorie burn, metabolic equivalent of task (MET), heart rate and step count) using Fitbit Charge HR device and they achieved average accuracy of 87.37% with Quadratic SVM classifier in one-to-many approach and average accuracy of 93% with Quadratic SVM classifier in one-to-one approach. In their revised scheme [14], they adopted more features (65) with different feature selection approaches and 93% (sedentary) and 90% (non-sedentary) with equal error rates of 5% is obtained. However, the Fitbit framework only provides only one sample each minute and access to the raw data is not possible. A system for continuous authentication with physiological signals should be low-cost and unobtrusive, and should not be dependent on certain activity for the sake of universality. It could be inferred from the comparison of the proposed system with the related work in Table I in terms of device and device position, features, unobtrusiveness, environment, and dependency to the activity type, our system outperforms other studies when feature engineering complexity, activity independence and unobtrusiveness are taken into consideration.

![Image](image.png)

Fig. 2. Our proposed solution continuously authenticate the user by processing the RR intervals coming from the smartband.

IV. PROPOSED SYSTEM

In this section, we explain our continuous authentication scheme. The overall multi-factor authentication diagram where a user initiate his/her session with a password, or fingerprint, is shown in Figure 2. In Figure
### TABLE I
PREVIOUS WORKS ON CONTINUOUS AUTHENTICATION.

| Article            | Year | Device(s)                          | Device Position(s)                  | Features                                                                 | Unobtrusive | Environment | Dependency to Activity Type |
|--------------------|------|------------------------------------|-------------------------------------|--------------------------------------------------------------------------|-------------|-------------|-----------------------------|
| S. A. Elkader et al. [18] | 2018 | A combination of acceleration, gyroscope and magnetometer sensors | Dominant wrist, dominant upper arm, non-dominant wrist, thigh, chest, ankle | Accelerometers, gyroscope and magnetometers features extracted from six sensor-based body locations | No | Laboratory | Kitchen area                  |
| Y. Zeng et al. [19] | 2017 | Shimmer 6DoF IMU, Samsung Galaxy S4 i9500, Samsung Galaxy Nexus i9500 | Wrist, ankle, hip/torso, thigh/front pocket, upper arm | Statistical features extracted from accelerometer sensor | Yes | Laboratory | Yes |
| S. Vhaduri et al. [20] | 2017 | Fitbit Charge HR device | Wrist | Statistical features extracted from step counts, heart rate, calorie burn, metabolic equivalent of task information | Yes | Daily Life | Yes |
| S. Vhaduri et al. [14] | 2019 | Fitbit Charge HR device | Wrist | Statistical features extracted from step counts, heart rate, calorie burn, metabolic equivalent of task information | Yes | Daily Life | Yes |
| Matsuyama et al. [21] | 2015 | Near-infrared Spectroscopy (NIRS) | Fore Head | Low-frequency brainwaves | No | Laboratory | Yes |
| Sarkar et al. [22] | 2016 | Front facing camera of mobile device Heartbeat detection kit as a wearable bracelet | In front of face | Deep Face Features | Yes | Daily Life | Yes |
| Ramli et al. [23] | 2016 | Fitbit Charge HR device | Wrist | ECG Wavelet Features | No | Laboratory | Yes |
| Ntantogian et al. [24] | 2015 | Camera, Motorola 360 Sport | In front of a person | Gait and gesture features | Yes | Laboratory | Yes |
| Musale et al. [25] | 2019 | Motorola 2nd Gen (smartwatch), Motorola G4 plus (smartphone) | Wrist | Statistical and human-action-related features from accelerometer and gyroscope sensor | Yes | Daily Life | Yes |
| Peng et al. [26] | 2017 | Google Glass | Head | Touch gestures (single-tap, swipe forward, swipe backward, swipe down, two-finger, swipe forward, and two-finger swipe backward) and voice commands | No | Laboratory | Yes |
| Our Work | 2019 | Samsung Gear S, Samsung Gear S2, Empatica E4 | Wrist | Heart Rate Variability features derived from PPG sensor | Yes | Real Life | No |

3. We show the data collection application, preprocessing for artifact detection, feature extraction, feature selection and classification units of our system. In Figure 4, we describe our preprocessing for artifact detection module.

A. Data Collection Application

We developed a data collection application in Tizen 3.0 framework [36] for Samsung Gear S1 and S2. The application collects inter-beat intervals and 3D accelerometer data and stores them as downloadable comma-separated values (CSV files). Empatica E4 has a cloud-based data collection application. The physiological signals can be downloaded as a CSV file. The gathered RR intervals from two different participants are shown in Figure 5.

B. Preprocessing and Artifact Removal

We implemented our preprocessing module in MATLAB [37] (see Figure 4). First, we loaded the CSV file provided from the smartbands. The signal is segmented...
into non-overlapping time windows of 120 seconds. According to the HRV guidelines, 2 minutes is the minimum window length for calculation of short-term HRV features [16]. Since response time is important for a continuous authentication system, we selected the minimum possible duration. Therefore, the minimum required duration of physiological data for authentication is 2 minutes. The artifacts in the RR intervals are detected by checking the difference between the consecutive points. We labeled the points exceeding more than 20% of the local average as artifacts, and the other points as the validated RR intervals, this threshold is selected from the previous works [38]. The points labeled as artifacts are deleted. After the removal, we implemented two different techniques. The first one is to interpolate the missing data points using a cubic spline interpolation algorithm which is commonly used [17]. The second technique is to apply minimum consecutive time and sample constraints on the remaining data to be regarded as meaningful. For example, if the minimum sample constraint is set to 5, we do not count three consecutive samples followed by a missing data point because the sequence is too short to be evaluated. In this study, we applied the former technique because it achieved better results [9].

C. Feature Extraction

We extracted time and frequency domain heart rate variability features from the segmented time windows. We used Marcus Volmer’s toolbox [39] which is implemented in MATLAB. We selected the features which are commonly used in the previous works related to heart rate variability [9], [38] and [40]. In order to compute the frequency domain features, the RR intervals are interpolated using 4Hz cubic spline interpolation, because RR intervals are unevenly sampled. We applied FFT to the interpolated windows. The computed features are shown in Table II. The total number of extracted features are 11 for each window.

D. Dimensionality Reduction

It is known that dimensionality reduction leads to better performance for the machine learning systems since it removes the unrelated features with the desired prediction [41]. We used Principal Component Analysis (PCA) based dimensionality reduction which is available in Weka toolkit [42]. PCA is a very powerful tool when applied with machine learning (ML) classifiers. It converts the set of vectors to uncorrelated variables. We explored the effect of different selection of covered variance and PCA variables. It is known that the covered variance affects the classification performance, therefore different values are evaluated (0.8, 0.85, 0.9 and 0.95) we reported the best achieving results, when the variance is selected as 0.9.

E. Handling Class Imbalance

Since some of the windows are deleted due to improper placement of the devices or heavy movements.
TABLE II
HEART RATE VARIABILITY FEATURES AND THEIR DEFINITIONS.

| HRV Feature     | Description                                                                                   |
|-----------------|-----------------------------------------------------------------------------------------------|
| Mean RR         | Mean value of the RR intervals                                                               |
| STD RR          | Standard deviation of the RR intervals                                                       |
| RMSSD           | Root mean square of successive difference of the RR intervals                                |
| pNN50           | Percentage of the number of successive RR intervals varying more than 50ms from the previous interval |
| HRV triangular index | Total number of RR intervals divided by the height of the histogram of all RR intervals measured on a scale with bins of 1/128 s |
| TINN            | Triangular interpolation of RR interval histogram                                             |
| SDSD            | Related standard deviation of successive RR interval differences                             |
| LF              | Power in low-frequency band (0.04-0.15 Hz)                                                    |
| HF              | Power in high-frequency band (0.15-0.4 Hz)                                                     |
| LF/HF           | Ratio of LF-to-HF                                                                            |
| VLF             | Power in very low-frequency band (0.00-0.04 Hz)                                               |

There is a class imbalance between participants. We applied the majority class subsampling to equalize the number of windows for each participant. This method is the most commonly used in the literature [43]. The majority of class instances are subsampled to 35. After the majority subsampling, the number of the imposter samples is 945 and the genuine samples are 35 for each authentication models.

Fig. 5. The RR intervals gathered from two participants with Samsung Gear S. The difference can be seen by looking at the raw signals.

Fig. 6. The MLP model used in the proposed system.

F. Machine Learning

We used k-Nearest Neighbour (kNN), Random Forest (RF), Multi-Layer Perceptron (MLP) and Linear Discrimenent Analysis (LDA) classifiers which are available in Weka Machine Learning software [42]. We investigated the different set of parameters for the classifiers and we fine tuned the parameters for the classifiers. The best performing feature set is the following; N selected as 3 for kNN, number of trees is selected as 100 for random forest and the hidden layer is selected as 1 and hidden unit as 5 for MLP, the MLP architecture is shown in Figure 6. We created a binary authentication model for each user. The selected user’s label is set to 1 and others as 0.

We applied 10 fold stratified cross-validation (the distribution of class labels are equal in each fold) for evaluating our system and fine tuning the parameters where 90% of the dataset is used for training and the rest is used for testing by changing the folds. We report the fine-tuned classification results.

G. Evaluation Metrics

In order to present the results of our authentication system, we provide the performance metrics used in the literature [14], [44], [18], [19], [8]. In authentication systems, there are two types of error which are False Acceptance Rate (FAR) and False Rejection Rate (FRR). These errors are affected from the selection of threshold which can be between 0 and 1 for the ML classifiers. A smaller selection of threshold will lead a low FAR but high FRR. The points of equilibrium is important for such system. This equilibrium is called Equal Error Rate (EER). The calculation of EER is implemented in Weka Machine Learning Software. The definitions are provided below:

- False Acceptance Rate (FAR): It is the ratio of false acceptance divided by the total attempts.
Fig. 7. Data collection procedure for the physiological continuous authentication system.

- False Rejection Rate (FRR): It is the ratio of denied legitimate attempts to the total number of attempts.
- Equal Error Rate (EER): This is the metric where FRR and FAR are equal.

V. DATA COLLECTION

In this section, we describe the data collection in real life and the ethics procedure. We collected physiological data from 28 people in controlled real-life settings, during a summer school for teachers. All of the participants are healthy teachers who have no prior medical condition. Before the data collection, subjects received and filled a consent. The gender of participants are 16 male and 12 female, the ages are between 25 and 35. The data collection procedure is shown in Figure 7. The duration of the total data collection is 110 minutes. The dataset has a baseline (20 minutes), lecture (40 minutes), free-time (10 minutes), examination (20 minutes) and recovery session (20 minutes). We did not use the free-time session which might create a bias on the results. The reason that we had these different scenarios is to create a daily life sequence. During the free-time participants were allowed to take a break from the lecture. We applied our implementation of Trier Social Stress Test [44] (TSST) which is frequently used for inducing stress. We selected questions from the mathematics Olympics (which is very hard for the normal population). We told the subjects that this is a test for measuring their intelligence, and we said that a moderate person achieves at least a 75% score. Every subject participated in every session and they did not know the objective of the study. The physiological data is gathered with different brands of commercial smartwatches (8 Empatica E4, 3 Samsung Gear S and 17 Samsung Gear S2).

Ethics

The procedure of the methodology used in this study is approved by the Institutional Review Board for Research with Human Subjects of Boğaziçi University with the approval number 2018/16. Prior to the data acquisition, each participant received a consent form which explains the experimental procedure and its benefits and implications to both the society and the subject. The procedure was also explained vocally to the subject. The data collection procedure and all of the interventions in this research fully meet the 1964 Declaration of Helsinki [45]. The data is stored anonymously.

VI. EXPERIMENTAL RESULTS AND DISCUSSION

We examined the results in two different subsections. In the first one, we presented and evaluated the authentication results of different devices and the whole system performance. In the second part, by applying a signal data quality filter, we improved the performance of the system.

A. Effect of Device Type on the Biometric Authentication Performance

We tested all 28 subjects and calculated EER results in Table III. These results are calculated by one vs. all tests for all subjects. EER results for four different classifiers are presented in Table VI. We also added the device type and average data quality columns to this table. Data quality presents the non-interpolated percentage of the data after removal of artifacts. As an example, if the average data quality is 70%, the remaining 30% of the data is interpolated. Data quality along with the device type affects the EER results significantly (see Figure 8). We achieved the best performance with Gear S as 98.48% and 3.96% EER. The selection of classifier has also important effect on the EER results. For example, Empatica E4 achieves 19.43% EER with kNN and 6.77% with RF classifier. The best achieving classifier is selected as RF in terms of EER. Design of the watch strap as show in Figure 1, PPG sensor quality, built-in processing algorithms of devices might be the factors for the difference in EER results.

B. Effect of Data Quality Constraint Filter

In daily life, seamless wrist-worn devices can get noisy signals, which drops the quality of the derived features. It is not possible to collect high-quality data all the time during a day because of various reasons such as high activity level and improper use of smartwatches. After observing that the data quality has a major effect on the authentication performance, we applied a data quality constraint on our data. Suppose that the data quality of a device is 50%. This means that the other half of the data is obtained by synthetic cubic interpolation data. Therefore, we expect that when the data is compared with other participants’ data, it could not be discriminated because it lost most of the unique characteristics of the PPG data. In Figure 8, we evaluated the effect of a quality threshold on EER. For this purpose, the evaluation has been done concerning different data quality thresholds for all devices. As we investigate the EER results of different device types, Samsung Gear S
TABLE III
THE EER METRICS OF EACH PARTICIPANTS WITH DIFFERENT CLASSIFIERS.

| Participant/Classifier | kNN  | RF   | MLP  | LDA  | Device | Average Quality |
|------------------------|------|------|------|------|--------|----------------|
| P1                     | 20.00 | 1.67 | 3.30 | 5.00 | E4     | 96.00         |
| P2                     | 22.84 | 5.00 | 10.00| 8.73 | E4     | 94.00         |
| P3                     | 5.68  | 3.23 | 5.65 | 9.95 | E4     | 96.00         |
| P4                     | 31.30 | 12.52| 15.52| 15.52| E4     | 93.00         |
| P5                     | 30.17 | 9.45 | 27.27| 27.27| E4     | 91.50         |
| P6                     | 23.96 | 8.30 | 15.87| 19.05| E4     | 97.00         |
| P7                     | 17.96 | 11.03| 22.22| 25.04| E4     | 92.00         |
| P8                     | 3.58  | 2.99 | 4.48 | 5.97 | E4     | 93.00         |
| P9                     | 4.00  | 5.33 | 5.33 | 6.67 | Gear S | 87.00         |
| P10                    | 1.33  | 2.67 | 4.00 | 5.33 | Gear S | 95.00         |
| P11                    | 9.20  | 3.39 | 5.08 | 9.32 | Gear S | 80.00         |
| P12                    | 37.47 | 18.73| 23.26| 21.13| Gear S | 85.00         |
| P13                    | 30.15 | 10.41| 21.13| 22.54| Gear S | 80.00         |
| P14                    | 13.52 | 3.80 | 8.45 | 9.86 | Gear S2| 62.00         |
| P15                    | 31.35 | 4.69 | 12.50| 17.19| Gear S | 68.00         |
| P16                    | 18.43 | 6.78 | 20.59| 20.59| Gear S | 54.00         |
| P17                    | 36.25 | 18.23| 27.78| 29.17| Gear S | 64.00         |
| P18                    | 43.34 | 24.55| 43.75| 40.63| Gear S | 79.00         |
| P19                    | 21.88 | 2.82 | 8.45 | 11.27| Gear S2| 69.00         |
| P20                    | 34.50 | 11.43| 18.75| 20.31| Gear S | 66.00         |
| P21                    | 39.30 | 19.55| 21.13| 22.54| Gear S | 66.00         |
| P22                    | 25.60 | 13.41| 18.75| 18.75| Gear S | 68.00         |
| P23                    | 31.67 | 14.08| 23.94| 21.13| Gear S | 62.00         |
| P24                    | 33.10 | 14.10| 23.44| 31.25| Gear S2| 53.00         |
| P25                    | 20.81 | 7.04 | 15.49| 19.72| Gear S | 65.00         |
| P26                    | 33.22 | 16.45| 24.23| 23.44| Gear S | 69.00         |
| P27                    | 37.48 | 20.54| 28.17| 29.58| Gear S | 76.00         |
| P28                    | 36.63 | 19.18| 23.44| 26.56| Gear S | 64.00         |

TABLE IV
AUTHENTICATION PERFORMANCE RESULTS, EER VALUE OF EMATICA E4, SAMSUNG GEAR S AND SAMSUNG GEAR S2 IS PRESENTED.

| Device      | kNN  | RF   | MLP  | LDA  | Average Quality |
|-------------|------|------|------|------|----------------|
| Empatica E4 | 19.43%| 6.77%| 13.03%| 14.56%| 94.06%         |
| Gear S      | 4.84% | 3.96%| 4.80% | 7.10%| 87.33%         |
| Gear S2     | 30.86%| 13.28%| 21.36%| 22.68%| 67.64%         |
| All Devices | 24.37%| 10.08%| 16.98%| 18.69%| 77.30%         |

VII. CONCLUSION

We proposed a scalable, unobtrusive and seamless continuous authentication system with commercial grade smartwatches and smartbands. We collected physiological data from 28 participants and demonstrated the EER measures for each of the participants in a real-life scenario. We proposed state-of-the-art preprocessing for signals coming from real-life data with artifacts due to the physical construction of the smartwatches. We achieved promising results by using our system (4.4% EER with Empatica E4). We showed the effect of different smartwatches. The selection of the classifier for the proposed system is very important. We applied feature based signal processing along with machine learning classifiers (kNN, RF, MLP and LDA). Even though, Gear S2 is a newer model of Gear S, due to the its leather strap, the signals coming from heart rate monitoring unit contained higher amount of artifacts, therefore it effects the overall quality of the RR intervals and the authentication system’s performance. For the authentication systems based on PPG sensors, sport straps can be a better choice, as shown in Figure 1. We showed that HRV can be used for continuous authentication without interrupting the activity of the user. We applied a signal removal procedure by using the overall RR interval quality measure, a higher quality leads to better performance after 80% quality threshold. The performance of the scheme vary between individuals. This conclusion is aligned with the literature [14], [8].

The minimum required amount of recording to apply our system for authentication is 2 minutes, once that is satisfied, authentication can be validated in seconds thanks to the sliding window approach. Our system can be implemented on any wrist-worn device which can provide RR intervals without a need for raw PPG.

Fig. 8. The change of EER metric with respect to minimum quality thresholds. Best achieved EER is 3.96%.
The proposed methodology can be used with various applications requiring continuous authentication. As a future work, we plan to apply our system completely in the wild settings with more participants and longer physiological recordings and show the performance of the framework. All of the evaluations are done in the same context, therefore in different types of contexts, the system might achieve better performance.

ACKNOWLEDGMENT

This work is supported by the Turkish Directorate of Strategy and Budget under the TAM Project number DPT2007K120610.

REFERENCES

[1] A. C. Baktir, C. Tunca, A. Ozgovde, G. Salur, and C. Ersoy, “Sdn-based multi-tier computing and communication architecture for pervasive healthcare,” IEEE Access, vol. 6, pp. 56,765–56,781, 2018.

[2] C. Ntantogian, S. Malliaros, and C. Xenakis, “Gaithashing: A two-factor authentication scheme based on gait features,” Computers & Security, vol. 52, 04 2015.

[3] Statista Research Department. (2016) Internet of things (iot) connected devices installed base worldwide from 2015 to 2025 (in billions). [Online]. Available: https://www.statista.com/statistics/411264/iot-number-of-connected-devices-worldwide/

[4] S. W. Shah and S. S. Kanhere, “Recent trends in user authentication – a survey,” IEEE Access, vol. 7, pp. 112,505–112,519, 2019.

[5] A. Ross, K. Nandakumar, and A. K. Jain, “Introduction to multibiometrics,” in Handbook of biometrics. Springer, 2008, pp. 271–292.

[6] N. Akhtar, H. Gite, G. Rabban, and K. Kale, “Heart rate variability for biometric authentication using time-domain features,” in International Symposium on Security and Communication. Springer, 2015, pp. 168–175.

[7] N. Akhtar, J. F. Mahdi, and G. R. Manza, “Microcontroller based data acquisition system for heart rate variability (hrv) measurement,” International Journal of Applied Science and Engineering Research, vol. 1, no. 4, pp. 576–583, 2012.

[8] D. Ramli, M. Hooi, and K. Chee, “Development of heartbeat detection kit for biometric authentication system,” vol. 96, 12 2016, pp. 305–314.

[9] Y. S. Can, N. Chalabianloo, D. Eikiz, and C. Ersoy, “Continuous stress detection using wearable sensors in real life: Algorithmic programming contest case study,” Sensors, vol. 19, no. 8, p. 1849, Apr. 2019.

[10] K. Phua, J. Chen, T. H. Dat, and L. Shue, “Heart sound as a biometric,” Pattern Recognition, vol. 41, no. 3, pp. 906–919, 2008.

[11] S. Pirbhulal, H. Zhang, S. Mukhopadhyay, C. Li, Y. Wang, G. Li, W. Wu, and Y. T. Zhang. “An efficient biometric-based algorithm using heart rate variability for securing body sensor networks.” Sensors, vol. 15, no. 7, pp. 15,067–15,089, 2015.

[12] Samsung, 2018, accessed at December 2019. [Online]. Available: https://www.samsung.com/

[13] S. Vhaduri and C. Poellabauer, “Multi-modal biometric-based implicit authentication of wearable device users,” IEEE Transactions on Information Forensics and Security, vol. 14, no. 12, pp. 3116–3125, Dec. 2019.

[14] Y. Sun and N. Thakor, “Photoplethysmography revisited: From contact to noncontact, from point to imaging,” IEEE Transactions on Biomedical Engineering, vol. 63, no. 3, pp. 463–477, Mar. 2016.

[15] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleieger, A. Malliani, A. J. Moss, and P. J. Schwartz, “Heart rate variability: Standards of measurement, physiological interpretation, and clinical use,” European Heart Journal, vol. 17, no. 3, pp. 354–381, 03, 1996.

[16] “Kubios user guide,” 2019. [Online]. Available: https://www.kubios.com/downloads/Kubios_HRV_Users_Guide.pdf

[17] S. A. Elked, M. Barlow, and E. Lakshika, “Wearable sensors for recognizing individuals undertaking daily activities,” in Proceedings of the 2018 ACM International Symposium on Wearable Computers - ISWC. ACM Press, 2018.

[18] Y. Zeng, A. Pande, J. Zhu, and P. Mohapatra, “WearA: Wearable device implicit authentication based on activity information,” in 2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM). IEEE, Jun. 2017.

[19] S. Vhaduri and C. Poellabauer, “Wearable device user authenti-
cation using physiological and behavioral metrics,” Oct. 2017.

[20] Y. Matsuyama, M. Shozawa, and R. Yokote, “Brain signal’s low-frequency fits the continuous authentication,” vol. 164, 03 2015.

[21] S. Sarkar, V. M. Patel, and R. Chellappa, “Deep feature-based face detection on mobile devices,” in 2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), Feb 2016, pp. 1–8.

[22] P. Musale, D. Baek, N. Werellagama, S. S. Woo, and B. J. Choi, “You walk, we authenticate: Lightweight seamless authentication based on gait in wearable IoT systems,” vol. 7, 2019, pp. 37,883–37,895.

[23] G. Peng, G. Zhou, D. T. Nguyen, X. Qi, Q. Yang, and S. Wang, “Continuous authentication with touch behavioral biometrics and voice on wearable glasses,” IEEE Transactions on Human-Machine Systems, vol. 47, no. 3, pp. 404–416, June 2017.

[24] M. Ghayourmi, “A review of multimodal biometric systems: Fusion methods and their applications,” in 2015 IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS). IEEE, Jun. 2015.

[25] R. Brunelli and D. Falavigna, “Person identification using multiple cues,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 10, pp. 955–966, Oct. 1995.

[26] E. Camlikaya, A. Kholmatov, and B. Yanikoglu, “Multi-biometric templates using fingerprint and voice,” in Biometric Technology for Human Identification V, B. V. Kumar, S. Prabhakar, and A. A. Ross, Eds. SPIE, Mar. 2008.

[27] M. Hammad, Y. Liu, and K. Wang, “Multimodal biometric authentication systems using convolution neural network based on different level fusion of ECG and fingerprint,” IEEE Access, vol. 7, pp. 26,527–26,542, 2019.

[28] M. D. Bugdol and A. W. Mitas, “Multimodal biometric system combining ECG and sound signals,” Pattern Recognition Letters, vol. 38, pp. 107–112, Mar. 2014.

[29] Q. Gui, Z. Jin, and W. Xu, “Exploring EEG-based biometrics for user identification and authentication,” in 2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). IEEE, Dec. 2014.

[30] M. Rangoussi, N. Alexandris, A. Evangelou, and M. Poulos, “Person identification from the EEG using nonlinear signal classification,” Methods of Information in Medicine, vol. 41, no. 01, pp. 64–75, 2002.

[31] C. Zhang, Y.-M. TIAN, and H.-W. WANG, “Review of ECG signal identification research,” DEStech Transactions on Computer Science and Engineering, no. aice-ncs, Mar. 2017.

[32] H.-S. Choi, B. Lee, and S. Yoon, “Biometric authentication using noisy electrocardiograms acquired by mobile sensors,” IEEE Access, vol. 4, pp. 1266–1273, 2016.

[33] A. Sarkar, A. L. Abbott, and Z. Doererrz, “Biometric authenti-
cation using photoplethysmography signals,” in 2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS). IEEE, Sep. 2016.

[34] A. Acsar, H. Aksu, A. S. Uluagac, and K. Akkaya, “Waca: Wearable-assisted continuous authentication,” in 2018 IEEE Security and Privacy Workshops (SPW), May 2018, pp. 264–269.
[36] Tizen, *Tizen API*, 2015, accessed at December 2018. [Online]. Available: [https://developer.tizen.org/development/api-references/web-application](https://developer.tizen.org/development/api-references/web-application)

[37] MATLAB, version 7.10.0 (R2010a). Natick, Massachusetts: The MathWorks Inc., 2010.

[38] B. Cinaz, B. Armarich, R. Marca, and G. Tröster, “Monitoring of mental workload levels during an everyday life office-work scenario,” *Personal Ubiquitous Comput.*, vol. 17, no. 2, pp. 229–239, Feb. 2013.

[39] M. Vollmer, *MarcusVollmer/HRV Toolbox*, 2017, accessed at December 2019. [Online]. Available: [https://www.github.com/MarcusVollmer/HRV](https://www.github.com/MarcusVollmer/HRV)

[40] M. Gjoreski, M. Luštrek, M. Gams, and H. Gjoreski, “Monitoring stress with a wrist device using context,” *Journal of biomedical informatics*, vol. 73, pp. 159–170, 2017.

[41] D. Barber, *Bayesian Reasoning and Machine Learning*. Cambridge University Press, 2012.

[42] F. Eibe, M. Hall, and I. Witten, “The weka workbench. online appendix for” data mining: Practical machine learning tools and techniques,” *Morgan Kaufmann*, 2016.

[43] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, “Handling imbalanced datasets: A review,” *GECTS International Transactions on Computer Science and Engineering*, vol. 30, pp. 25–36, 11 2005.

[44] C. Kirschbaum, K.-M. Pirke, and D. H. Hellhammer, “The ‘triier social stress test’—a tool for investigating psychobiological stress responses in a laboratory setting,” *Neuropsychobiology*, vol. 28, no. 1-2, pp. 76–81, 1993.

[45] W. M. Association et al., “World medical association declaration of helsinki. ethical principles for medical research involving human subjects.” *Bulletin of the World Health Organization*, vol. 79, no. 4, p. 373, 2001.

Yağmur Ceren Dardağan is a senior student in Computer Engineering Department of Bogazici University, Turkey. Her research interests include physiological signal processing, machine learning and pervasive health applications.

Cem Ersoy received the Ph.D. degree from Polytechnic University, New York, NY, USA, in 1992. He was an R&D Engineer in NETAS A. S. from 1984 to 1986. He is a Professor of computer engineering at Bogazici University, Turkey. He is also the Vice Director of the Telecommunications and Informatics Technologies Research Center, TETAM. His research interests include wireless/cellular/ad-hoc sensor networks, activity recognition, and ambient intelligence for pervasive health applications, green 5G and beyond networks, mobile cloud/edge/fog computing, software-defined networking, and infrastructure-less communications for disaster management. He was the Chairman of the IEEE Communications Society Turkish Chapter during the last eight years. He is a member of the IFIP.

Deniz Ekiz received the MS degree from Computer Engineering Department, Bogazici University, Turkey in 2019. He is a Ph.D. candidate in the Computer Engineering Department of Bogazici University, Turkey. His research is focused on the health-related applications of wearable technology.

Yekta Said Can studied Computer Engineering in Faculty of Engineering, Bogazici University, Istanbul, Turkey where he obtained his B.Sc Degree in 2012. He obtained his M.Sc degree at the same department, in 2014 while working as a researcher at TUBITAK BILGEM for two years. He is pursuing a PhD degree right now in Computer Engineering at Bogazici University. He also works as a teaching assistant in the same department. His research interest includes watermarking, speech and speaker recognition, physiological signal processing and machine learning.