IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 March 2024, accepted 19 April 2024, date of publication 29 April 2024, date of current version 6 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3395128

RESEARCH ARTICLE

Wearable Wisdom: A Bi-Modal Behavioral Biometric Scheme for Smartwatch User Authentication

ATTAULLAH BURIRO[1], ZAHID AKHTAR[2], FRANCESCO RICCI[3], AND FLAMINIA L. LUCCIO[1]

[1]Department of Environmental Sciences, Informatics and Statistics, University of Venezia Ca' Foscari, 30123 Venice, Italy
[2]Department of Network and Computer Security, State University of New York Polytechnic Institute, Utica, NY 13502, USA
[3]Faculty of Engineering, Free University of Bolzano-Bozen, 39100 Bolzano, Italy

Corresponding author: Attaullah Buriro (attaullah.buriro@unive.it)

This work was supported in part by Projects "SEcurity and RIghts In the CyberSpace —SERICS" under the National Recovery and Resilience Plan (NRRP) funded by European Union (EU)—NextGenerationEU under Grant PE00000014-CUP H73C2200089001, in part by the "Interconnected Nord-Est Innovation Ecoscheme—iNEST" under NRRP funded by European Union—NextGenerationEU under Grant ECS00000043-CUP H43C22000540006, and in part by PRIN/PNRR "Automatic Modelling and ∀erification of Dedicated sEcUrity deviceS—AM∀DEUS" under NRRP funded by European Union—NextGenerationEU under Grant P2022EPPHM-CUP H53D23008130001.

ABSTRACT Multi-modal biometric schemes, which leverage more than one sensor, offer significant advantages over uni-modal schemes in terms of accuracy and robustness against attacks. In this paper, we present a novel behavioral multi-biometric user authentication scheme that employs data fusion for biometric-based authentication. The proposed scheme extracts two types of data (Electromyography (EMG) and movements) from a single user's clapping action type, via a worn smartwatch. During the user enrollment phase, the scheme creates a digital identity of the user based on the arm's movement and EMG signatures generated through the entire period of clapping action. In the verification phase, the scheme authenticates users based on the detected combined EMG and movement signature. The experimental analysis of the proposed scheme employing a Deep Neural Network classifier demonstrates its efficacy; our classifier achieves a True Accept Rate of 94.37%, a False Accept Rate of just 0.11%, and an accuracy of 97.13%. Furthermore, it is experimentally demonstrated that augmenting the training dataset via Generative Adversarial Networks leads to even better performance; with improved TAR (up to ≈96%) and accuracy (up to 97.94%), while decreasing FAR (from 0.11% to 0.082%). These results showcase the effectiveness and robustness of the proposed bi-modal behavioral biometric scheme for smartwatch user authentication.

INDEX TERMS Smartphone authentication, behavioral biometrics, sensors.

I. INTRODUCTION

Smartwatches have now become popular wearable devices. Besides their more traditional use (e.g., showing time, handling text messages and phone calls), smartwatches are now widely exploited for performing sensitive operations, such as, touchless payments [1], obtaining medical alerts [2], fitness tracking [3], tracking heart rate and/or detecting heart rhythm irregularities, accessing cars [4], and opening a garage door, just to name a few. Given the increasing use of smartwatches for critical applications, it becomes of utmost importance to protect the private data stored in or accessed from smartwatches against any unauthorized access.

Authentication, which is the verification of the identity of claimants, is a fundamental component of any cybersecurity solution. Authentication aims at preventing unauthorized access and maintaining the required level of confidentiality and integrity of private data. User authentication could be performed by leveraging the user-specific secret (e.g., PIN, password, etc.). However, these approaches are

The associate editor coordinating the review of this manuscript and approving it for publication was Julien Le Kernec.

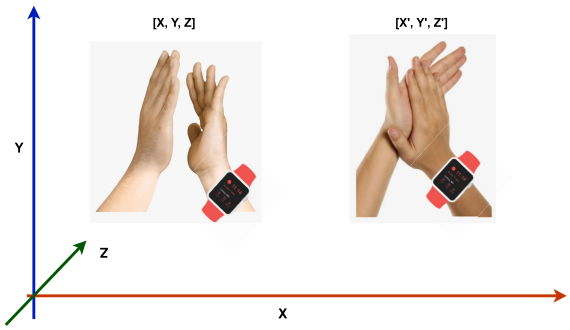


FIGURE 1. Clapping in 3d space.

not considered to be optimally suited for smartwatch usage because of their well-known security and usability limitations [5]. Usability is limited by the imposed requirement of active user cooperation, and the difficulty of entering or sketching a secret pattern on the small touch screen. For instance, security limitations such as susceptibility to shoulder surfing attacks and insufficiently robust authentication mechanisms further impede their effectiveness Alternately, biometric-based authentication establishes the user's identity through biological modalities, such as the user's face, fingerprint, or retina. These solutions already exist on recent smartphones (e.g., Apple Face ID [6], and fingerprint sensors in iPhones [7]). However, their deployment on smartwatches is still limited,[1] as the use of physical biometrics also pose usability and/or privacy concerns. For example, iris recognition requires cumbersome iris entry hence requiring extra user effort. Some studies have shown that users even raise privacy concerns on using these technologies [8]. Consequently, behavioral biometric-based authentication solutions are garnering considerable attention in both academia [9], [10], and industry [11]. For example, it is worth noting that Facebook is currently researching EMG wearables in its virtual reality lab to find if such a device could be used to precisely interact with computing devices [11]. According to them, this wrist-based interaction could be used as an alternative mean of intentional input with a maximum level of portability. To this end, we expect EMG-based smartwatches to arrive soon in the market.

Unlike physical biometrics, such as those based on face, fingerprint, and iris recognition, behavioral-biometrics-based authentication schemes assume that each person has a distinctive manner of performing actions such as typing [12] or swiping [13] on a keyboard, or holding and moving a smartwatch [9], [10], [14], [15], [16], [17], [18], [19]. These activities can be analyzed in order to create a user profile that can be used for authentication purposes on smartwatches.

Towards the design of higher usability and user-friendly user authentication methods on smartwatches, in this article, we propose EMG-assisted clap as a behavioral-biometric modality. The proposed scheme authenticates users by exploiting two invisible behavioral biometric signatures, namely, EMG and sensor-based clap signature (see Figure 1), to overcome the limitations of existing knowledge-based and physiological user authentication approaches. Our presented scheme exploits built-in smartwatch hardware, i.e., 3-dimensional sensors to register the arm movements generated during the performed clapping action, and the EMG sensors to collect electromyographic signals. A user is identified based on the differences between the currently collected EMG-assisted clap signature and all the signatures collected during the enrollment phase. Our proposed scheme is very convenient, as it does not ask users to memorize and type any secret, which increases its usability. Moreover, replicating the clapping action in a manner that accurately matches both the movement and EMG data signatures of the target user proves to be challenging. This level of alignment significantly bolsters the security and robustness of the scheme, as it necessitates a sophisticated understanding and mimicking of the user's unique behavioral patterns across multiple modalities. We have implemented and evaluated our proposed scheme with multiple alternative state-of-the-art classifiers, namely, K-Nearest Neighbor (KNN), Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Deep Neural Network (DNN) on our created chimerical dataset. DNN outperforms the other classifiers and attains the highest True Accept Rate (TAR) of 94.37% at just 0.11% False Accept Rate (FAR) and accuracy of$\approx$ 97.13%. These results were obtained by using a particular chimerical data set, which was obtained by combining each user (data) from the clap data set with another user (data) in the EMG data set. However, there was the risk that the specific matching of users in the two data sets to build the chimerical data may have determined the observed result. To remove this possible bias, we have performed additional experiments using 50 chimerical datasets: each one is created by randomly pairing users from the clap and EMG datasets. The outcome was a consistent performance observed across all these datasets, which confirms the robustness of the classifier to variations in the construction of the chimerical dataset, and suggests that good performance can also be obtained by using non chimerical data sets. Moreover, we further explored the potential of DNN by training it on an augmented dataset, generated using Generative Adversarial Network (GAN). Remarkably, this approach led to even better performance, with the DNN improving TAR (up to $\approx$ 96%) and accuracy (up to $\approx$ 98%) while decreasing FAR (from 0.11% to 0.082%).

The main contributions of this paper are listed below:

- The proposal of a novel user-friendly and secure behavioral biometric-based authentication scheme for smartwatches. Our scheme is the first to combine EMG and arm-movement for user authentication. The scheme uses the differences between the EMG-assisted sensory readings generated using the performed clapping action during the enrollment phase, and the same signature collected during the verification phase. We note that

[1] https://www.retaildive.com/ex/mobilecommercedaily/samsung-paypal-tap-fingerprint-technology-in-smartwatch-response-to-apple

users only need to perform a faint clapping, with or without sound. The main goal is collecting movements, not the sound.
- Experimental evaluation of the proposed scheme's efficacy and robustness on a created chimerical dataset of EMG and clapping action using advanced machine learning classifier, e.g., DNN. Experimental evaluation of the reliability and validity of DNN on cross-user chimerical datasets: we generated several chimerical datasets by randomly combining clap and EMG users. The results showed that our classifier achieves stable and consistent performance across different user combinations. This indicates that our classifier is not biased by a specific selection of the chimerical data set.
- The proposal and evaluation of a GAN scheme for data augmentation to attain elevated precision.

Paper Organization: The rest of the paper is organized as follows: Section II surveys the relevant research studies published on smartwatch unlocking. Section III presents a high-level explanation of our approach. Section IV presents the evaluation strategy, and analysis and discusses the obtained results. Section V concludes this work with a summary of the findings and of future work.

II. RELATED WORK

There exists a wide and diverse range of multimodal biometric methods. For instance, the progress of multimodal biometric schemes in authentication can be seen in [20] and [21]. In this section, we survey relevant user authentication approaches that are similar to our work. We have chosen to focus on the studies exploring behavioral biometric approaches, particularly those leveraging arm movements and EMG signals, as these are the two key modalities utilized in our proposed bi-modal scheme for smartwatch user authentication. The rationale behind this decision is that our work is specifically aimed at advancing the state-of-the-art in behavioral biometric-based authentication for wearable devices. Prior studies have demonstrated the uniqueness and stability of arm movements and EMG signals as promising traits for user verification on wearable platforms, offering advantages in terms of seamless integration and user acceptance. By surveying the relevant literature in these two specific areas, we can clearly position our novel contribution of combining these behavioral biometric traits in a unified scheme for smartwatch user authentication, effectively highlighting how our work advances the current state-of-the-art in this domain. Furthermore, we have opted not to incorporate advanced techniques such as Convolutional Neural Networks (CNN) and Transformer-based classifiers due to their requirement of large training set. Specifically, during the bootstrapping phase where a small-sized dataset is utilized, these sophisticated models are deemed impractical. Therefore, in the context of our research scheme, characterized by limited data availability, the utilization of CNN and Transformer-based classifiers is not considered feasible.

A. ARM-MOVEMENT-BASED USER AUTHENTICATION

Existing smartwatches are equipped with multiple sensors that could potentially be used to detect arm rotations, arm movements, heart rate, blood oxygen level, skin temperature and conductance. This information can be exploited to develop unobtrusive user authentication schemes. Specifically, relevant to this work, there exist a few arm-movement based behavioral biometric authentication schemes that exploit the device built-in 3 dimensional sensors for authentication purposes [9], [10], [14], [15].

Yang et al. [22] and Lewis et al. [9] were the first ones to propose a motion-based authentication solution for smartwatches. Yang et al. [22] introduced MotionAuth which was implemented on Android and its effectiveness was tested on the collected data of 30 users. Authors tested four simple and natural gestures, e.g., drawing a circle using a smartwatch worn arm, and attained an Equal Error Rate (EER) of 2.6% using Dynamic Time Warping (DTW), and a histogram-based techniques for classification. Similarly, Lewis et al. [9] employ free-form arm movement as a behavioral modality and reported an 84.6% accuracy using DTW as a classifier, over the collected dataset of 5 users.

Moreover, we consider the works presented in [14], [15], [16], [23], [24], and [25] to be very close to our approach. Authors of these studies have proposed motion-based finger-snapping [14], in-air-writing [15], Hand-punch [16], motion-based clapping gesture [23], [24], and motion-based handwriting gesture [25], for user authentication on smartwatches. We summarise these studies and compare them with our approach in Table 1.

B. EMG-BASED USER AUTHENTICATION

Few studies (e.g., [26], [27], [28]) advocated using EMG as a behavioral biometric for user authentication on smartphones. However, to the best of our knowledge, none of them has utilized it either for smartwatch-based solutions, or has proposed a motion-based clapping action that uses both EMG and smartwatch's built-in sensors for user authentication on smartwatches.

We consider the study presented in [28] close to our work. This study presents an EMG-assisted pattern-based unlocking mechanism for smartphones. The scheme collects the EMG signals while the pattern lock is being drawn on the touchscreen by the user. In this way, even if someone knows the pattern, he would not be able to access the smartphone because either none or a different EMG signal would be acquired. The authors recruited 10 users and recorded 20 trials from each user. Later, they extracted time-domain features from the recorded data and trained their chosen 1-class classifiers, namely Support Vector Machine (SVM) and Local Outlier Factor (LoF). The authors reported the results in the form of Half Total Error Rate (HTER), i.e., the average of FAR and False Reject Rate (FRR). They achieved an HTER ranging from 7.25% to 23.50%, for different users. Another relevant study [29] uses the EMG signals acquired

from Myo armbands to authenticate smartphone users. The authors conducted several experiments on their collected dataset of 53 users and obtained an average TAR of 91.81%, at a FAR of 7.43%, by using a Convolutional Siamese Network classifier.

While prior work has explored the use of arm movement and EMG signals individually for user authentication, the combination of these two behavioral biometric traits in a multimodal scheme remains an underexplored area, particularly for smartwatch applications.

Our authentication scheme differs from prior motion-based schemes in the following ways: firstly, while prior work has explored the use of arm movement and EMG signals individually for user authentication, the combination of these two behavioral biometric traits in a multimodal scheme remains an underexplored area, particularly for smartwatch applications. Thus our approach utilizes a novel bi-modal scheme that extracts two types of data (arm movements and EMG signatures) from a single behavior, which is clapping. Secondly, previous studies (as seen in Table 1) have not utilized a Deep Neural Network as classifier for tackling this task. Thirdly, our approach employs data augmentation through the use of a Generative Adversarial Network. Fourthly, our approach is robust against adversarial examples, i.e., Fast Gradient Sign Method (FGSM) attacks. Finally, our approach surpasses the state of the art in terms of True Accept Rate (TAR), False Accept Rate (FAR), and accuracy. It attains better results compared to any state-of-the-art arm-movement based approach listed in Table 1.

III. PROPOSED APPROACH: EMG-ASSISTED CLAP BIOMETRICS

Our clapping-based approach leverages the users' specific way of clapping to authenticate them. In particular, our proposed scheme collects EMG and arm movements signals generated while a user performs the clapping action (for 3.5 sec). The proposed approach then extracts statistical features from: a) 3-dimensional sensors (accelerometer and gyroscope) for capturing the arm movements, and b) 8-channel EMG stream for capturing the muscle activity to register the clap modality. Since the scheme exploits users' familiarity with the clapping gesture and uses a second type of sensor that invisibly collects a complementary type of data, the scheme can achieve high usability and user acceptance. Additionally, it is secure because it requires a high level of similarity between the two types of data generated by the clapping behavior, which is hard to achieve by an imitator to access the user's smartwatch.

Figure 2 illustrates the proposed approach. During users' enrollment, the smartwatch collects the sensory readings from the accelerometer, gyroscope, and EMG sensors, and extracts statistical features (explained in Section in III-B) from the required clapping actions. Then, these extracted features from the accelerometer and gyroscope are fused to make feature vectors, one for each clapping action. Next, it applies GANs to generate more samples (explained in Section in III-C), and these expanded feature vectors are finally saved in the database for both training the classifier and future comparisons during verification. During user verification, the user is required to clap once. The user's smartwatch collects the corresponding sensory readings, extracts the same features, and compares this query feature vector with the feature vectors stored in the database during enrollment. If the recent feature vector is found similar to one previously stored, the access is granted, otherwise, the user is considered an impostor.

A. DATASETS

A chimerical dataset is a synthetic dataset that is created by combining data from different modalities or sources, such as EMG and clapping, or face and eye [51]. A chimerical dataset can be useful for evaluating the performance of multimodal biometric schemes, especially when there is no real dataset available that contains the desired modalities. Thus, we utilize two distinct publicly available datasets of EMG and clapping coming from different persons. In particular, since none of the available smartwatches have EMG sensors at the moment, we rely on a publicly available dataset [30] that used the Myo Armband[2] for EMG. Conversely, for movements, we relied on "WISDM Smartphone and Smartwatch Activity and Biometric Dataset" [31].

We acknowledge that the EMG data used in our experiments comes from a different type of behavior than clapping. However, we argue that this does not invalidate our experiments, as we are interested in evaluating the performance of our scheme on fusing two types of data (EMG and movements) from a clap action for user authentication. We believe that our scheme can generalize to real EMG data collected from clapping, as the EMG signals are consistent and distinctive for each user.

Dealing with the nonexistence of bi-sensor datasets including EMG and clapping biometric traits, we assembled a bi-modal *chimerical* dataset (i.e., a dataset in which every user has an EMG and clapping trait) by combining the EMG and clapping data of pairs of users of publicly available EMG [30] and clapping [31] datasets. Namely, because the two adopted datasets do not contain the same users, we uniquely and randomly matched the EMG modality and clapping modality of pairs of distinct users of respective datasets, obtaining a chimerical database of 50 users, with 128 genuine samples per user. Moreover, to assess the robustness and validity of our classifier, we generate multiple chimerical datasets created by randomly matching users in the two available data sets. In practice, we combined the features of a randomly chosen user of the clap dataset with randomly selected user in the EMG dataset, and repeated this process 50 times to generate 50 chimerical datasets. Then, we tested our classifier on each of them. In fact, we aimed to ensure that the classifier's performance is stable and reliable across the alternative data combinations

[2]https://wearabletech.io/myo-bracelet/

TABLE 1. Comparison of our scheme with the related work. Our comparison is limited to the work that involved sensory readings, the mode, the classifiers (i.e., Dynamic Time Warping (DTW), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), Deep Neural Network (DNN), Recurrent Neural Networks (RNNs)), the number of users, and the obtained results.

Paper	Input Method	Mode	Sensors	Classifiers	Users	Best Results
[9]	Free-form arm-movement	Uni-modal	Acceleromter Gyroscope	DTW	5	Accuracy = 84.6%
[10]	password-based arm-movement	Uni-modal	Acceleromter Gyroscope	RNNs	310	EER = 6.09%
[14]	Finger-snaping-based movements	Uni-modal	Acceleromter Gyroscope	MLP	11	TAR = 82.34% FAR = 34.12%
[15]	In-air-finger-based movements	Uni-modal	Acceleromter Gyroscope	DTW MLP	11	TAR = 80.52% FAR = 21.65%
[16]	Hand-punch movement	Uni-modal	Accelerometer	SVM	20	Accuracy = 95.45%
[22]	Arm-movement: 1. Circle 2. Arm-rotation 3. Hand-up 4. Hand-down	Uni-modal	Acceleromter Gyroscope	DTW Historgram	30	EER = 2.6%
[23]	Motion-based activities (18)	Uni-modal	Accelerometer Gyroscope	KNN DT RF	50	Accuracy = 96.65% (clapping)
[24]	Hand-clapping-based movements	Uni-modal	Accelerometer Gyroscope	KNN MLP RF	50	TAR = 93.3% FAR = 0.22% Accuracy = 96.54%
[25]	Hand-writing-based movements	Uni-modal	Acceleromter Gyroscope	MLP SVM	21	EER = 6.56%
This work	**EMG-assisted-claping movements**	**Bi-modal**	**Accelerometer Gyroscope**	KNN MLP SVM **DNN**	50	**TAR = 95.95% FAR = 0.082% Accuracy = 97.94%**

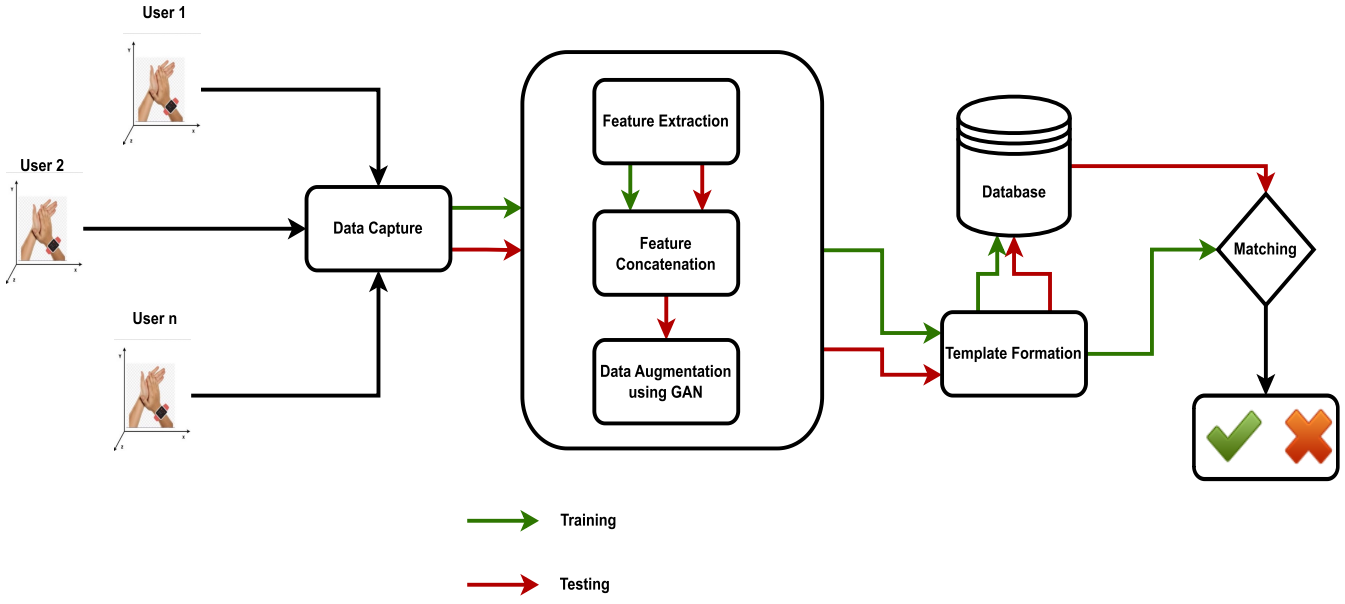


FIGURE 2. Block diagram of the proposed EMG-assisted clap biometrics approach.

(not only for a specific chimeric combination). Please note that assembling chimerical datasets is a common practice in the multimodal biometric schemes works when no real databases are obtainable [32], [34], [35], [51].

1) WISDM SMARTPHONE AND SMARTWATCH ACTIVITY AND BIOMETRIC DATASET

The "WISDM Smartphone and Smartwatch Activity and Biometrics Dataset" [31] contains sensory values collected

from accelerometer and gyroscope sensors of both smartphones and smartwatches from 51 users while they were performing 18 diverse daily living activities. Each activity was performed for 3 minutes, so each user contributed 54 minutes of data. We have used only the data related to the clapping activity recorded using smartwatch from the first 50 users in this study.

2) MULTI-CHANNEL ELECTROMYOGRAPHY SIGNAL ACQUISITION OF FOREARM [30]

This publicly available dataset consists of EMG samples collected from 50 users using 8-channels Myo Armband for 10 different arm-movements (i.e., wrist in neutral, pronation, supination, wrist extension, wrist flexion, ulnar deviation, radial deviation, fine pinch, power grip, and hand open). These movements are related to the position, rotation, and angle of the wrist and forearm, as well as the type of grip and finger spread of the hand. Since these factors can affect the way a person claps, we assume that clapping-based movements will also be different from user to user, because each user may have a unique style, preference, and habit of clapping, influenced by their personality, culture, mood, and context. Moreover, clapping is a complex and dynamic behavior that involves coordination of multiple muscles and joints, which can vary in strength, flexibility, and sensitivity among users. Therefore, we expect that clapping-based movements will have sufficient variability and distinctiveness to be used as a biometric modality for user authentication.

The data was collected at a sample rate of 200 samples/sec from each of its eight channels also known as skin surface electrodes. We chose this dataset because this is the more realistic data to emulate the smartwatch sensors. We are hopeful that the soon-to-launch Facebook smartwatch will have 8-channels as well.

B. FEATURES EXTRACTION

Our smartwatch user authentication solution uses built-in 3-dimensional accelerometer and gyroscope sensors, as well as 8 EMG channels, to extract two types of data from a single action type - clapping. Accelerometer and gyroscope sensors are 3-dimensional: *X*, *Y*, and *Z*. Myo Armband was used to prepare EMG dataset [30] that contains eight electrodes, often termed as channels. Each electrode senses the muscle activities and produces electrical energy. In this way, the EMG data is recorded under each of the channels in millivolts. Consequently, the final dataset encompasses an 8-channel stream (octuple) for each timestamp. We extracted 5 statistical features from the accelerometer and the gyroscope sensors, namely, Min, Max, Mean, Variance, and Standard Deviation from each dimension from a captured sample (of 3.5 sec). Additionally, we wanted to understand how the streams are related to each other. As such, we also measured correlation, absolute difference, and cosine similarity between the dimensions, e.g., as motion sensors are 3-dimensional we have *XY*, *XZ*, and *YZ* correlations. In total, we obtained 60-features long feature vector for the clapping modality. For the EMG modality, we compute the same features, but here the number of dimensions is 8 (unlike accelerometer and gyroscope which are 3-dimensional), thereby we have a larger feature vector of size 140-features. For creating a chimeric user profile, we concatenated the two feature vectors for each observation per user and form a 200-features-long feature vector for the bi-sensor scheme. Technically speaking, for a chimeric user (e.g., user-1), we concatenate feature vectors from each observation of both modalities of user-1 available in both datasets and so on. As we extracted 128 observations for each user from each of the dataset. Hence the total number of observations in chimeric dataset is 6400 samples.

C. DATA AUGMENTATION USING GAN

Data augmentation is a technique commonly used to increase the size of a dataset by transforming existing data into new and unique samples. The reasons behind applying augmentation are (i) improvement of model performance, (ii) reduce the need for more data, and (iii) increase model's robustness. We adapted Generative Adversarial Network (GAN) for data augmentation.

The GAN, introduced by Goodfellow et al. [36], is an unsupervised method for creating synthetic data. GAN uses two sub-models: a generator and a discriminator. The discriminator is first trained on real data, and the generator produces new samples that the discriminator classifies as either ''real'' or ''not real.'' The feedback from the discriminator allows the generator to iteratively improve the quality of the generated samples. This approach has been applied in a variety of fields such as image generation [37], video generation [38], wireless communications (for spectrum sensing [39] and signal spoofing [40]), and voice and language processing [41], [42]. However, generating one-dimensional (1D) tabular data, particularly in the area of user authentication [43], [44], [45], has received limited attention.

Our implementation of a Generative Adversarial Network (GAN) used the Python programming language and the Keras library.[3] In this work, the architecture of the generator *G* and discriminator *D* was used for GAN-based synthetic EMG-assisted clapping sample generation.

In this work, the developed generator and discriminator networks are small neural networks. The generator has four hidden layers with 128, 16, 16, and 8 units respectively, while the discriminator has five hidden layers with 288, 352, 512, 64, and 32 units respectively. We experimented with different numbers of hidden layers and units in the generator to arrive at this configuration, which showed excellent results in generating synthetic samples, as can be seen in Figure 3, that illustrates the comparison of some ''real'' and ''non-real'' features for a randomly chosen user across all 5000 epochs.

The generator's first layer takes in a fixed-size (100-dimensional) noise vector, and its final output is a 200-dimensional vector that resembles the 200-dimensional vector of ''real'' samples. These 200-dimensional vectors

[3]https://keras.io/

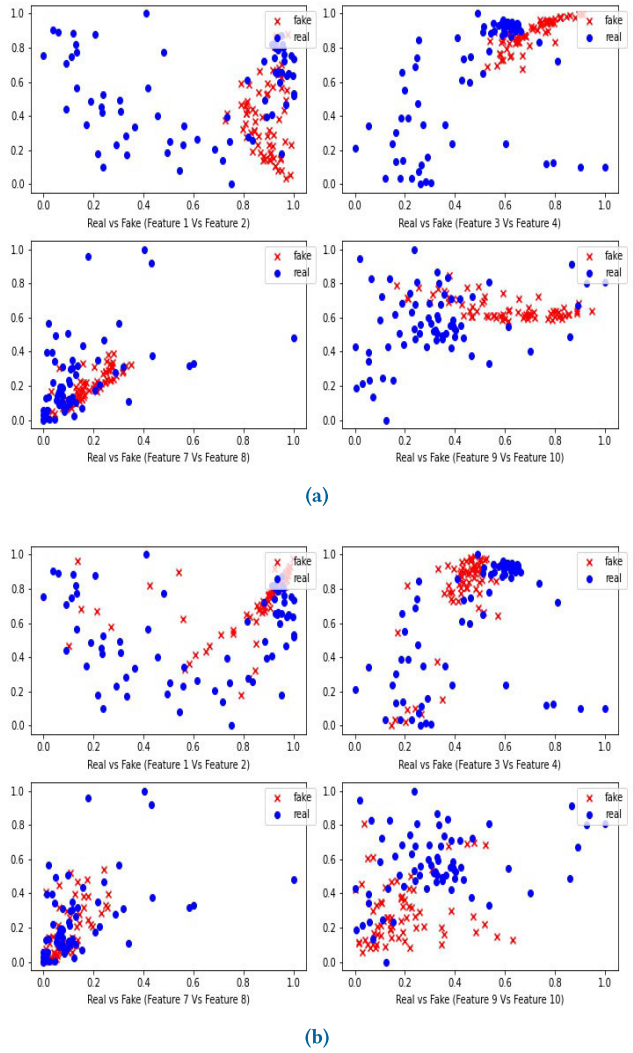


FIGURE 3. Comparison of "real" (in blue) and "non-real" (in red) features after: (a) 1000, (b) 5000 epochs.

from both "real" and "non-real" samples serve as input to the discriminator, whose output is binary, indicating whether the sample is "real" or "non-real".

After 5000 epochs of generating "non-real" samples and saving the generated data every 1000 epochs, we compared some original and generated features in Figure 3; in this figure, we show this comparison only for 1000 (see Figure 3a) and 5000 (see Figure 3b). The synthetic "non-real" (in red) samples were very similar to the "real" (in blue) samples after 5000 epochs as depicted in the figure.

t-Distributed Stochastic Neighbor Embedding (t-SNE) [50], is a dimensionality reduction technique commonly used to visualize high-dimensional data in a lower-dimensional space, typically two. t-SNE is a nonlinear dimensionality reduction technique, meaning it can capture complex relationships and structures in the data that linear methods, e.g., PCA would not be able to capture effectively. It is particularly useful in preserving both local and global structures in the data. Local structures refer to the relationships between neighboring points, while global structures involve relationships across the entire dataset.

In Figure 4, we present the t-SNE visualization of the original and generated samples for some randomly chosen users, aiming to illustrate the similarities in two dataset. The t-SNE plots depicted in Figure 4, provide a comprehensive snapshot of the distribution of samples in the high-dimensional space, effectively reduced to two dimensions for visualization purposes. The original and synthetic samples are color-coded for clarity, with blue representing the original data and red depicting the synthetic data (generated after 5000 epochs). It is evident from the figure that the clusters corresponding to original and synthetic samples significantly overlap which means our GAN was successful in learning well the distribution of original samples and hence managed to generate high-quality synthetic samples.

To evaluate the quality of these synthetic samples, we added them to the original train set and observed an increase in accuracy.

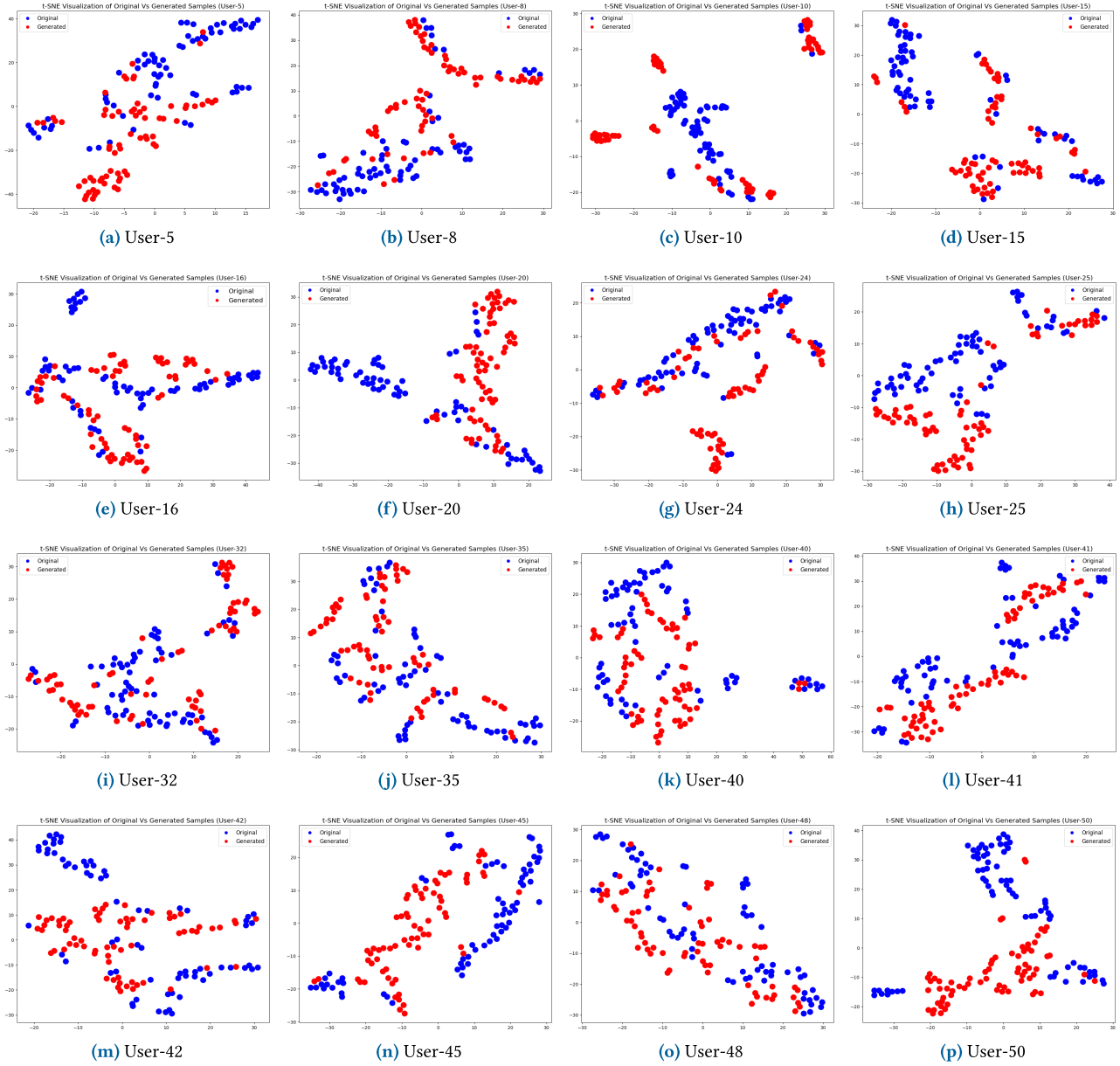
IV. EXPERIMENTAL EVALUATION

In this section, we describe the experimental evaluation of the proposed bimodal smartwatch user authentication scheme.

A. MACHINE LEARNING CLASSIFIERS

The proposed scheme is an attempt of solving the problem of user authentication in client-server architecture such as banking, remote access, etc. In this scenario, the chosen ML classifier is trained on training samples of multiple users. We considered three state-of-the-art multi-class classifiers (i.e., KNN, MLP, and SVM) for obtaining baseline results as well as for comparison with related studies. The rationale behind selection of these diverse classifiers is to provide a comprehensive evaluation of our proposed scheme. By including state-of-the-art baseline classifiers such as KNN, MLP, and SVM, we were able to benchmark our performance against well-established models. Furthermore, we developed an optimized Deep Neural Networks (DNNs) based classification/verification scheme specifically to leverage the complex, bimodal nature of the arm movement and EMG data for user authentication. To this aim, we searched the best number of layers, the best-required units in each layer, and the learning rate. It is worth noting that we performed a grid search to determine the best hyperparameters of adopted classifiers among the listed parameters in Table 2. We leveraged Scikit-learn[4] a python library for hyperparameter optimization of baseline classifiers and Keras-tuner[5] for DNN network. It is worth mentioning that we exploited training and validation data splits for finding best hyperparameters. While we recognize the potential of advanced CNN and Transformer-based classifiers in achieving state-of-the-art performance, their efficacy is heavily dependent on the large-scale, diverse datasets for training. In our case, the

[4]https://scikit-learn.org/stable/
[5]https://keras.io/keras_tuner/



(a) User-5 (b) User-8 (c) User-10 (d) User-15

(e) User-16 (f) User-20 (g) User-24 (h) User-25

(i) User-32 (j) User-35 (k) User-40 (l) User-41

(m) User-42 (n) User-45 (o) User-48 (p) User-50

FIGURE 4. Comparison of t-SNE plots for different users for "real" (in blue) and "non-real" (in red) samples.

dataset size (70 samples/user) did not meet the requirements for leveraging the full capabilities of these models.

B. EXPERIMENTAL PROTOCOL

For each user, as we previously said, we obtained 128 observations. We took the first 70 observations (emulating 20 Hz for clapping and 200 Hz for EMG for 3.5 sec duration) for training the classifiers. The next 30 observations were selected for validation, i.e., performing parameter optimization. The final 28 (from index 101 – 128) samples were used for the final testing of the classifiers. It is worth noting that this test set remains unseen by the classifier during the training or parameter optimization and data augmentation processes.

C. EXPERIMENTAL RESULTS

For authentication, we summarize our results in terms of True Accept Rate (TAR), False Reject Rate (FRR), False Accept Rate (FAR), True Reject Rate (TRR), and Accuracy. For adversarial robustness analysis, we report the results in terms of accuracy. TAR, FRR, FAR, and TRR, are defined as the fraction of legit attempts correctly classified as legit, the legit attempts misclassified as adversarial, the adversarial attempts misclassified as legit, and the adversarial attempts classified as adversarial, respectively. The accuracy is the ratio of correct classifications and the total attempts. Since the FRR and TRR can also be computed as $1 - TAR$ and $1 - FAR$, respectively, we only show TAR and FAR to avoid redundancy.

TABLE 2. Parameter optimization of all chosen classifiers. The parameters "α" and "C" are regularization parameters for MLP and SVM, respectively, where as "Tanh" (hyperbolic tangent) is an activation function used in MLP.

Classifiers	Parameters	Range	Best	Best validation Accuracy (%)
KNN	*# of Neighbors*	*1 to 50 (step-size=1)*	1	80.12
MLP	"*Size of hidden layer*"	*8 to 64 (stepsize=8)*	64	95.60
	"*activation*"	*"tanh", "relu"*	*tanh*	
	"*solver*"	*"sgd", "adam"*	*adam*	
	"α"	"0.0001", "0.001", "0.01"	0.0001	
	"*learning_rate*"	*"constant", "adaptive"*	*constant*	
SVM	"*C*"	*0.1 to 2.0 (stepsize=2)*	0.1	96.88
	"γ"	"1", "0.1", "0.01," 0.001"	1	
	"*Kernel*	*"linear", "rbf", "poly", "sigmoid"*	*linear*	
DNN	"*num_layers*"	*2 to 10 (stepsize=1)*	5	93.46
	"*num_units*"	*32 to 512 (stepsize=32)*	288, 352, 512, 64, 32	
	"*learning_rate*"	"0.01", "0.001", "0.0001"	0.0001	

1) USER AUTHENTICATION RESULTS

We summarize user authentication scheme performance results in Figures 5 and 6.

Figure 5 illustrates the results of the performed analysis on individual modalities: we show the performance of our chosen classifiers when trained on original data points for clap (see Figure 5a) and EMG (see Figure 5c) modalities, and when trained on augmented data points for clap (see Figure 5b) and for EMG (see Figure 5d) modalities. Overall, the highest accuracy achieved by DNN classifier in unimodal settings is 90.4% (for clap) and 89.56% (for EMG) on original data points. This accuracy is further improved, when the the augmented data points are used, up to 93.69% (for clap) and 93.84% (for EMG), respectively.

The authentication results of our bimodal approach are reported in Figure 6. In Figure 6a, we show the TAR, FAR, and accuracy of the chosen baseline classifiers in their default settings. It is evident that the MLP classifier yields the highest 86.07%, 0.28%, and 92.89%, TAR, FAR, and accuracy, respectively. This figure does not include the results of DNN because, unlike these classifiers, it does not have any baseline implementation (researchers have to design this architecture depending on their datasets, applications, and so on).

In Figure 6b, we summarize the results of our tuned classifiers trained on the original training set (which contains 70 samples of each user). We can see that performance of KNN and MLP remains similar to what they achieved in the default settings. But the performance of the SVM classifier is improved noticeably, e.g., TAR goes from 84.47% to 87.81% leading to a nearly 2% increase in accuracy. DNN classifier with 5 hidden layers (as discussed earlier) outperformed all the baseline classifiers and obtains the highest TAR (94.37%), lowest FAR (just 0.11%), and highest accuracy (97.13%).

It is worth repeating that there is a potential risk that the unique chimerical dataset, constructed and employed in our experiments earlier, might have biased the results. To eliminate this risk, we have conducted additional experiments with 50 distinct datasets (obtained by varying the matching of users in the two available data sets). The uniform performance across these diverse datasets proved the resilience of the classifier to alterations in the composition of the chimerical datasets. We have applied the identical training/validation/testing split that was previously used to each dataset. We selected DNN to conduct this experiment as it is the most effective classifier. The results are stable across the 50 datasets. We achieve an average TAR of 92.905% (with a standard deviation of 1.07), an average FAR of 0.145% (with a standard deviation of 0.0219), and an average Accuracy of 96.381% (with a standard deviation of 0.546). These results show that the performance of the classifier is robust and is not influenced by the specific way users in the two datasets are combined to form the chimerical dataset.

In Figure 6c, we summarize the results of tuned classifiers when they were trained on the augmented dataset. The purpose of data augmentation was to answer a basic question: *is it possible to improve the classifier performance on an increased number of samples without asking users to provide more of them*, and the answer is yes. Namely, we could exploit GAN to generate synthetic "non-real" samples and by training on this augmented training set, we can obtain enhanced performance. It is notable that, in this study, the generation of the sample was performed using training samples (70 only) for each user. Normally, GAN requires a large number of samples for good quality generation. We can notice that the performance of baseline-tuned classifiers did not change much as they were found unable to extract meaningful insights from the augmented training set. However, the

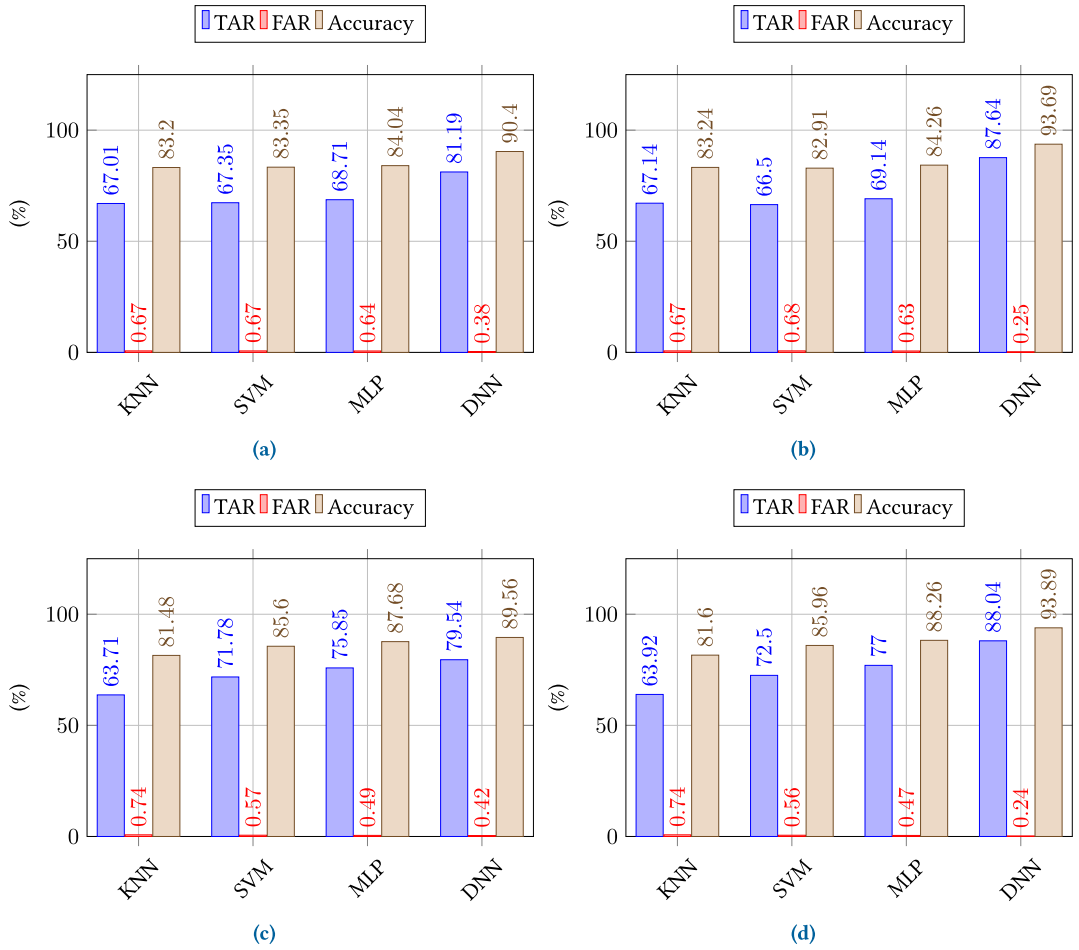


FIGURE 5. Results of chosen classifiers in unimodal settings: Clap on original and augmented data (a) and b), EMG on original and augmented data(c) and (d).

performance of DNN improved noticeably. On the augmented dataset, the TAR and accuracy are further increased and FAR is lowered. We can observe that the DNN achieved a TAR of 95.95%, FAR of just 0.082%, and accuracy of 97.94%.

The most common way of determining a Deep Neural Network (DNN) overfitting is by monitoring the model's learning curves and comparing the accuracy and losses of training and validation sets. If the accuracy of the training set continues to increase with more iterations of training and the accuracy of the validation set decreases, then this is a clear indication of overfitting. Similarly, we can analyze the gap between training and validation loss. If training loss continues to decrease while the validation loss increases, then this also suggests overfitting. Due to space limitations, we show the learning curves of our DNN network in Figure 7 for bimodal settings only. In both curves, we can see that the trend of accuracy and loss remains the same throughout the training process which proves that our classifier is not overfitting.

2) ROBUSTNESS ANALYSIS AGAINST ADVERSARIAL EXAMPLES

Though DNNs-based schemes now can achieve remarkable performances, it has been empirically shown that DNNs are vulnerable to adversarial examples [46], [47], [48]. Adversarial examples are carefully crafted versions of genuine samples which are intentionally perturbed using adversarial noise with the aim of confusing DNNs, leading to the mis-classification of a given input.

In this paper, we are simulating a specific scenario: we assume that authorized users conduct Fast Gradient Sign Method (FGSM) attacks on their own devices (perform ''penetration testing'') - a white-box attack to ensure that they have control over the accuracy of their enabled authentication mechanism. By doing these attacks, users will be able to verify the strength of their authentication mechanism and become sure of its security against potential hackers. Further, evaluation of the pre-trained neural network-based schemes under adversarial examples is considered extremely important because it helps to assess robustness, uncover

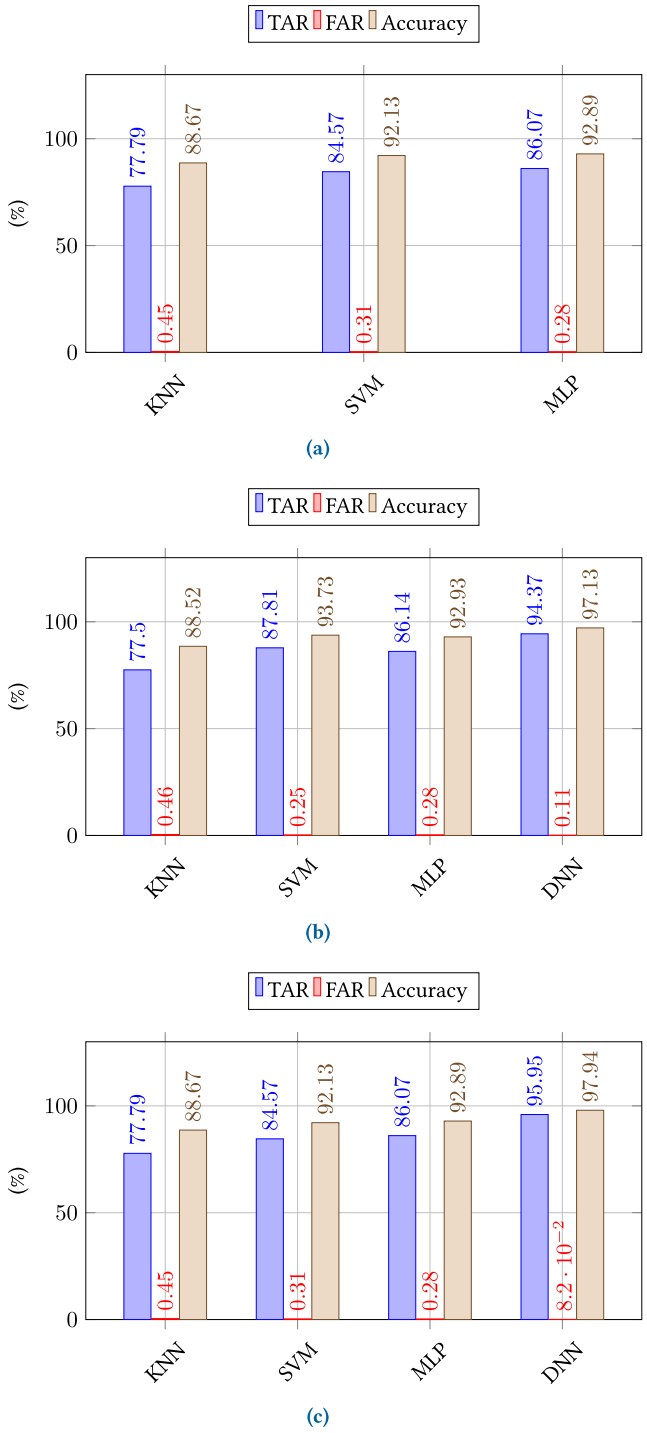


FIGURE 6. Baseline classification results: (a) in default settings, (b) in optimized settings, and (c) in optimized settings on augmented train set.

vulnerabilities, and ensure that the model's performance is not overly dependent on the distribution of the training data. Thus, we also evaluated the robustness of the proposed user authentication scheme against adversarial examples.

There exist several methods to generate adversarial examples [47]. In this study, we adopted FGSM.[6] FGSM is a

[6]https://pytorch.org/tutorials/beginner/fgsm_tutorial.html

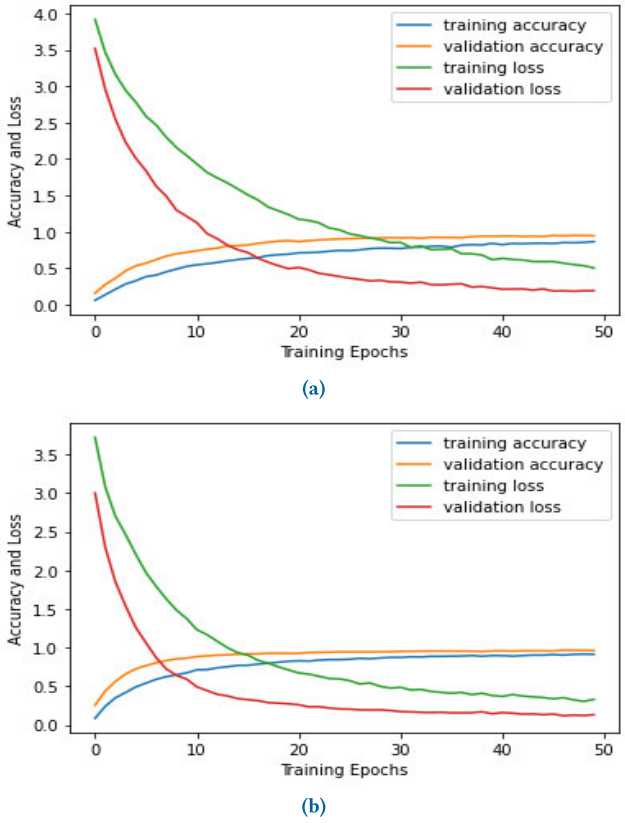


FIGURE 7. Learning curves of the DNN classifier (in bimodal settings): (a) On original training set (b) on the augmented training set.

white-box attack, i.e., the attacker has complete knowledge of the model's architecture and parameters. FGSM is performed by adding a small perturbation to the input data, calculated by taking the sign of the gradient of the model's loss function to the input data. The objective of the attack is causing the model to misclassify the perturbed input.

There are several ways of interpreting the results of an FGSM attack, including accuracy, confidence scores, adversarial example visualizations, and feature importance. However, accuracy is the most commonly used metric. If the accuracy of the classifier on the original data split is significantly higher compared to FGSM-generated data, it means the classifier is likely to be vulnerable to adversarial attacks.

We applied the FGSM attacks on our pre-trained models and the results are shown in Figure 8. The model trained on the augmented training set shows higher resistance to FGSM attack, as the accuracy only drops by an acceptable small margin compared to without any attack. We generated FGSM on a range of ϵ values (i.e., from 0.001 to 0.01) as these values are better suitable for FGSM generation on tabular data. However, the values chosen for image data range from 0.01 to 1 [49]. Epsilon ϵ is a hyperparameter that controls the strength of the attack. Technically speaking, It determines how much perturbation is added to the original input data to create an adversarial example. We believe that our proposed solution offers robust security against adversarial attacks/examples and provides a secure platform

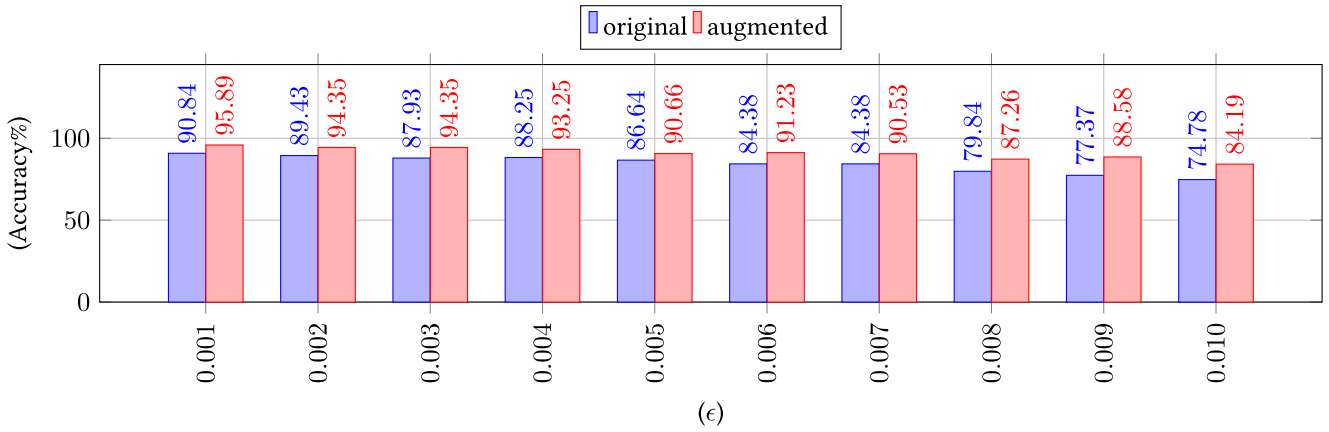


FIGURE 8. FGSM attack on pre-trained model on the original and augmented training sets on different ϵ values.

for user authentication in the smartwatch user authentication domain.

V. CONCLUSION AND FUTURE WORK

This paper presents a novel bimodal behavioral biometric-based user authentication scheme for smartwatches, which has high usability and efficacy. The proposed scheme extracts two types of data (clapping and EMG signal) from clapping action, to register the user and authenticate them later. We consider our scheme user-friendly because it exploits the users' familiarity with the clapping action and collects the electromyographic signals without any inconvenience. Moreover, the addition of an invisible layer (created using the 8-channel EMG sensory readings) enhances the security of the scheme as it is very difficult to imitate both the clapping and the EMG signal of another user.

We have evaluated and compared our proposed scheme with four different state-of-the-art machine learning classifiers (including KNN, MLP, and SVM as baseline classifiers) and Deep Neural Networks (DNN). Our results show that DNN outperforms its counterparts. By using DNN, we have obtained the highest accuracy of 97.13% (trained on the original training set) and 97.94% (trained on an augmented training set). It is noteworthy that our DNN classifier exhibited stable and robust performance across all the selected chimeric datasets. We have empirically shown that by increasing the number of training samples using GANs (i.e., augmented training set), the accuracy of such schemes increase as reported by other studies as well [44], [45]. The proposed scheme could be used not only as a standalone method to authenticate users on smartwatches but also to complement any existing method, such as face recognition on laptops and fingerprint recognition in automotive applications.

The proposed scheme offers a seamless integration of behavioral biometrics, granting users greater flexibility. Our future plans involve transforming the scheme into an adaptive scheme, that could keep collecting the EMG-assisted clap data passively to ensure reaching a significant number of samples and re-training to attain the higher accuracy. Considering this, we intend to revisit the possibility of incorporating advanced CNN and Transformer-based models into our methodology once we have access to a larger and more diverse dataset. Additionally, we also intend to evaluate the proposed scheme in different scenarios as some studies show that behavioral patterns may change depending on the context or environment. Also, we will prototype a proof-of-the-concept application based on our findings, and perform usability and security testing. Finally, we are also interested to carry out the performance evaluation in terms of power, memory, and timings (i.e., sample acquisition time and decision time, etc.).

ACKNOWLEDGMENT

The authors would like to thank all the volunteers for their participation in the experiments and groupmates for their valuable and insightful thoughts.

REFERENCES

[1] (2022). *Garmin Pay Contactless Payments*. [Online]. Available: https://explore.garmin.com/en-U.S./garmin-pay/
[2] (2022). *Amie Clark, Apple Watch Medical Alert*. [Online]. Available: https://www.theseniorlist.com/medical-alert-watch/apple/
[3] (2022). *Garmin Pay, Fitness Tracking Products*. [Online]. Available: https://www.garmin.com/en-U.S./c/sports-fitness/activity-fitness-trackers/
[4] (2022). *BMW To Use IPhone's U1 Chip for Digital Car Keys*. [Online]. Available: https://www.macworld.com/article/676033/bmw-to-use-iphones-u1-chip-for-digital-car-keys.html/
[5] W. Meng, D. S. Wong, S. Furnell, and J. Zhou, "Surveying the development of biometric user authentication on mobile phones," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1268–1293, 3rd Quart., 2015.
[6] (2022). *Apple, About Face ID Advanced Technology*. [Online]. Available: https://support.apple.com/en-us/HT208108/
[7] (2022). *Apple, Use Touch ID on IPhone and IPad*. [Online]. Available: https://support.apple.com/en-us/HT201371/
[8] (2020). *State Of Password and Authentication Security Behaviors Report*. Accessed: Apr. 13, 2023. [Online]. Available: https://pages.yubico.com/2020-password-and-authentication-report
[9] A. Lewis, Y. Li, and M. Xie, "Real time motion-based authentication for smartwatch," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Oct. 2016, pp. 380–381.
[10] C. X. Lu, B. Du, X. Kan, H. Wen, A. Markham, and N. Trigoni, "VeriNet: User verification on smartwatches via behavior biometrics," in *Proc. 1st ACM Workshop Mobile Crowdsensing Syst. Appl.*, Nov. 2017, pp. 68–73.
[11] (2022). *Facebook Reportedly Developing a Smartwatch and Researching EMG Sensors*. [Online]. Available: https://www.myhealthyapple.com/facebook-smart-watch-emg-sensors//

[12] T. Nguyen and N. Memon, "Tap-based user authentication for smartwatches," *Comput. Secur.*, vol. 78, pp. 174–186, Sep. 2018.
[13] T. V. Nguyen, N. Sae-Bae, and N. Memon, "DRAW-A-PIN: Authentication using finger-drawn PIN on touch devices," *Comput. Secur.*, vol. 66, pp. 115–128, May 2017.
[14] A. Buriro, B. Crispo, M. Eskandri, S. Gupta, A. Mahboob, and R. Van Acker, "Snapauth: A gesture-based unobtrusive smartwatch user authentication scheme," in *Proc. Int. Workshop Emerg. Technol. Authorization Authentication*, 2018, pp. 30–37.
[15] A. Buriro, R. Van Acker, B. Crispo, and A. Mahboob, "AirSign: A gesture-based smartwatch user authentication," in *Proc. Int. Carnahan Conf. Secur. Technol. (ICCST)*, Oct. 2018, pp. 1–5.
[16] G. C. Liang, X. Y. Xu, and J. D. Yu, "User-authentication on wearable devices based on punch gesture biometrics," in *Proc. ITM Web Conf.*, vol. 11, 2017, p. 1003.
[17] K. A. Rahman, N. Alam, J. Musarrat, A. Madarapu, and M. S. Hossain, "Smartwatch dynamics: A novel modality and solution to attacks on cyber-behavioral biometrics for continuous verification?" in *Proc. IEEE Int. Symp. Netw., Comput. Commun.*, 2020, pp. 1–5.
[18] J. Lee, S. Park, and E.-K. Lee, "Behavior-based authentication using user biological data to IoT device having touchscreen," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2021, pp. 750–752.
[19] G. Stragapede, R. Vera-Rodriguez, R. Tolosana, A. Morales, A. Acien, and G. Le Lan, "Mobile behavioral biometrics for passive authentications," *Pattern Recognition Letters*, vol. 157, pp. 33–41, May 2022.
[20] S. B. Abdullahi, Z. A. Bature, P. Chophuk, and A. Muhammad, "Sequence-wise multimodal biometric fingerprint and finger-vein recognition network (STMFPFV-Net)," *Intell. Syst. Appl.*, vol. 19, Sep. 2023, Art. no. 200256.
[21] S. B. Abdullahi, C. Khunpanuk, Z. A. Bature, H. Chiroma, N. Pakkaranang, A. B. Abubakar, and A. H. Ibrahim, "Biometric information recognition using artificial intelligence algorithms: A performance comparison," *IEEE Access*, vol. 10, pp. 49167–49183, 2022.
[22] J. Yang, Y. Li, and M. Xie, "MotionAuth: Motion-based authentication for wrist Worn smart devices," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, Mar. 2015, pp. 550–555.
[23] G. M. Weiss, K. Yoneda, and T. Hayajneh, "Smartphone and smartwatch-based biometrics using activities of daily living," *IEEE Access*, vol. 7, pp. 133190–133202, 2019.
[24] A. Buriro and F. Ricci, "ClapAuth: A gesture-based user-friendly authentication scheme to access a secure infrastructure, emerging technologies for authorization and authentication," in *Proc. 5th Int. Workshop*, Sep. 30, 2022, pp. 15–30.
[25] I. Griswold-Steiner, R. Matovu, and A. Serwadda, "Handwriting watcher: A mechanism for smartwatch-driven handwriting authentication," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 216–224.
[26] H. Tamura, T. Gotoh, D. Okumura, H. Tanaka, and K. Tanno, "A study on the S-EMG pattern recognition using neural network," *Int. J. Innov. Comput., Inf. Control*, vol. 5, no. 12, pp. 4877–4884, 2009.
[27] H. Yamaba, A. Kurogi, S.-I. Kubota, T. Katayama, M. Park, and N. Okazaki, "Evaluation of feature values of surface electromyograms for user authentication on mobile devices," *Artif. Life Robot.*, vol. 22, no. 1, pp. 108–112, Mar. 2017.
[28] Q. Li, P. Dong, and J. Zheng, "Enhancing the security of pattern unlock with surface EMG-based biometrics," *Appl. Sci.*, vol. 10, no. 2, p. 541, Jan. 2020.
[29] B. Fan, X. Su, J. Niu, and P. Hui, "EmgAuth: Unlocking smartphones with EMG signals," *IEEE Trans. Mobile Comput.*, vol. 22, no. 9, pp. 5248–5261, Sep. 2022.
[30] (2022). *IVN Joel Ramrez Ngeles and Marco Aceves-Fernandez, Multi-Channel Electromyography Signal Acquisition of Forearm*. [Online]. Available: https://data.mendeley.com/datasets/p77jn92bzg/1/
[31] G. M. Weiss, "WISDM smartphone and smartwatch activity and biometrics dataset," *UCI Mach. Learn. Repository*, vol. 7, pp. 133190–133202, Jan. 2019.
[32] N. Poh and S. Bengio, "Can chimeric persons be used in multimodal biometric authentication experiments?" in *Proc. Int. Workshop Machine Learn. Multimodal Interact.*, 2005, pp. 87–100.
[33] P. Lopes Silva, E. Luz, G. Moreira, L. Moraes, and D. Menotti, "ChimericalDataset creation protocol based on doddington zoo: A biometric application with face, eye, and ECG," *Sensors*, vol. 19, no. 13, p. 2968, Jul. 2019.
[34] Z. Akhtar, A. Buriro, B. Crispo, and T. H. Falk, "Multimodal smartphone user authentication using touchstroke, phone-movement and face patterns," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2017, pp. 1368–1372.
[35] J. L. Wayman, "A path forward for multi-biometrics," in *IEEE Int. Conf. Acoust. Speed Signal Process. Proc.*, Aug. 2006, pp. 1–15.
[36] I. J. Goodfellow et al., "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, Jun. 2014, pp. 2672–2680.
[37] J. Yang, A. Kannan, D. Batra, and D. Parikh, "LR-GAN: Layered recursive generative adversarial networks for image generation," 2017, *arXiv:1703.01560*.
[38] C. Vondrick, H. Pirsiavash, and A. Torralba, "Generating videos with scene dynamics," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 613–621.
[39] K. Davaslioglu and Y. E. Sagduyu, "Generative adversarial learning for spectrum sensing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.
[40] Y. Shi, K. Davaslioglu, and Y. E. Sagduyu, "Generative adversarial network in the air: Deep adversarial learning for wireless signal spoofing," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 1, pp. 294–303, Mar. 2021.
[41] J. Li, W. Monroe, T. Shi, S. Jean, A. Ritter, and D. Jurafsky, "Adversarial learning for neural dialogue generation," 2017, *arXiv:1701.06547*.
[42] L. Yu, W. Zhang, J. Wang, and Y. Yu, "Seqgan: Sequence generative adversarial nets with policy gradient," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 1–17.
[43] J. V. Monaco, M. L. Ali, and C. C. Tappert, "Spoofing key-press latencies with a generative keystroke dynamics model," in *Proc. IEEE 7th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Sep. 2015, pp. 1–8.
[44] T. Piplani, N. Merill, and J. Chuang, "Faking it, making it: Fooling and improving brain-based authentication with generative adversarial networks," in *Proc. IEEE 9th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Oct. 2018, pp. 1–7.
[45] A. Buriro, F. Ricci, and B. Crispo, "SWIPEGAN: Swiping data augmentation using generative adversarial networks for smartphone user authentication," in *Proc. 3rd ACM Workshop Wireless Secur. Mach. Learn.*, Jun. 2021, pp. 85–90.
[46] J. Zhang and C. Li, "Adversarial examples: Opportunities and challenges," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 7, pp. 2578–2593, Jul. 2020.
[47] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, and D. Mukhopadhyay, "Adversarial attacks and defences: A survey," 2018, *arXiv:1810.00069*.
[48] J. Monteiro, I. Albuquerque, Z. Akhtar, and T. H. Falk, "Generalizable adversarial examples detection based on bi-model decision mismatch," in *Proc. IEEE Int. Conf. Syst. Man Cybern. (SMC)*, Bari, Bari, Italy, Oct. 2019, pp. 2839–2844.
[49] A. Musa, K. Vishi, and B. Rexha, "Attack analysis of face recognition authentication systems using fast gradient sign method," *Appl. Artif. Intell.*, vol. 35, no. 15, pp. 1346–1360, Dec. 2021.
[50] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal Machine Learning Research*, vol. 9, no. 11, pp. 1–19, 2008.
[51] P. L. Silva, E. Luz, G. Moreira, L. Moraes, and D. Menotti, "Chimerical dataset creation protocol based on Doddington zoo," *Sensors*, vol. 19, no. 13, p. 2968, Aug. 2019.

ATTAULLAH BURIRO received the Ph.D. degree in information and communication technology (security and privacy) from the University of Trento, Italy, in February 2017. He is currently an Assistant Professor with the University of Venezia Ca' Foscari, Venice, Italy. Prior to this, he was an Assistant Professor, from September 2020 to August 2023; and a Postdoctoral Researcher with the Free University of Bolzano-Bozen, Bolzano, from September 2019 to August 2020, and the University of Trento, Italy, from March 2017 to August 2019. His research interests include biometrics, access control, the Internet of Things (IoT), computer vision, machine learning, artificial intelligence, and data mining. He has developed several secure, user-friendly, implicit behavioral biometric-based authentication solutions for smartwatches, smartphones, and critical infrastructures.

ZAHID AKHTAR received the Ph.D. degree in electronic and computer engineering from the University of Cagliari, Italy. He is currently an Assistant Professor with the Department of Network and Computer Security, State University of New York (SUNY) Polytechnic Institute, USA. Prior to that, he was a Research Assistant Professor with the University of Memphis, USA, and a Postdoctoral Fellow with the INRS-EMT, University of Quebec, Canada, the University of Udine, Italy, Bahcesehir University, Turkey, and the University of Cagliari, Italy. His research interests include computer vision and machine learning with applications to cybersecurity, biometrics, affect recognition, image and video processing, and audiovisual multimedia quality assessment.

FRANCESCO RICCI is currently a retired Professor with the Faculty of Engineering, Free University of Bozen-Bolzano. He was a Senior Researcher and the Technical Director of the E-commerce and Tourism Research Laboratory (eCTRL), ITC-IRST, Trento, Italy, from 2000 to 2006. From 1998 to 2000, he was a Scheme Architect with the Research and Technology Department (Process and Reuse Technologies), Sodalia S.p.A. His research interests include recommender schemes, user modeling, machine learning, and ICT applications to travel and tourism. He is the author of more than 200 refereed publications. According to Google Scholar, his H-index is 64 and has around 25,000 citations. He is a Co-Editor of the *Recommender Schemes Handbook* (Springer 2022).

FLAMINIA L. LUCCIO is currently an Associate Professor with the University of Venezia Ca' Foscari, Venice, Italy. She was visiting for research purposes, Carleton University of Ottawa, Ottawa University, University of Montreal, Canada, École Normale Supérieure de Lyon, and Université Bordeaux 1, France. Her research interests include the development and analysis of algorithms for distributed schemes and on computer security, with a particular emphasis on cryptography, security APIs, trusted hardware, and on usable security. She has published more than 100 papers in international journals and conferences. She has been a Scientific Coordinator of the Ca' Foscari Research Group of different projects, including the Amadeus Prin 2022 PNRR Project and the Serics SCAI PNRR Mission 4 Project. She has been in the PC of many international conferences.

Open Access funding provided by 'Università Ca' Foscari Venezia' within the CRUI CARE Agreement