Data Article

A chimerical dataset combining physiological and behavioral biometric traits for reliable user authentication on smart devices and ecosystems

Sandeep Gupta [a, *], Attaullah Buriro [a], Bruno Crispo [a, b]

[a] Department of Information Engineering & Computer Science (DISI), University of Trento, Italy
[b] Department of Computer Science (DISI), DistriNET, KULeuven, Belgium

ARTICLE INFO

ABSTRACT

We present a chimerical dataset that combines both physiological and behavioral biometric traits, for reliable user authentication on smart devices and ecosystems [1]. The data are composed of statistical features computed from swipe-gesture, voice-prints, and face-images. The swipe and voice-prints data presented hereinafter are collected using a customized Android application - DriverAuth, however, the face data is obtained from the MOBIO Dataset [2].

We collected 10,320 swipe and voice-prints samples from 86 users worldwide by collaborating with a professional crowd-sourcing platform and formed a chimerical dataset adjunct to the publicly available MOBIO dataset with our collected dataset. The dataset consists of various statistical features computed from the raw data for all three traits, i.e., swipe, voice-print, and face.

---

* Corresponding author.
E-mail addresses: sandeep.gupta@unitn.it (S. Gupta), attaullah.buriro@unitn.it (A. Buriro), bruno.crispo@unitn.it (B. Crispo).

Specifications Table

Subject area	*Information Security*
More specific subject area	*User authentication, Physiological and Behavioral Biometrics, smart devices and ecosystems*
Type of data	*Text files*
How data was acquired	*We developed "DriverAuth" - an Android app, to collect swipe and voice-prints. A crowdsourcing company - UBERTESTERS was hired for the data collection and they recruited the testers to perform the experiment on our prototype application.*
	Face data is obtained from the MOBIO Dataset [2]
Data format	*CSV*
Experimental factors	*A chimerical dataset combining three distinct traits, i.e., swipe, voice, and face. In the experiment, swipe and voice data are collected from 86 participants and face data is obtained from the MOBIO Dataset [2].*
Experimental features	*In total 393 statistical features are extracted from 3 biometric traits, i.e. swipe (33), voice (104) and face (256)*
Data source location	*DISI, University of Trento, Italy*
Data accessibility	*Dataset is uploaded with this article.*
Related research article	*DriverAuth: A Risk-based Multi-modal Biometric-based Driver Authentication Scheme for Ride-sharing Platforms (Ref. COSE_1458) [1] https://doi.org/10.1016/j.cose.2019.01.007*

Value of the Data

- Data can be used by scientists, researchers or mobile devices manufacturing companies, in order to build a multimodal user authentication scheme.
- Swipe gesture, voice-prints, and face-images can be used for authentication purposes on smart devices and ecosystems, in either unimodal or multimodal settings. They have shown to be a reliable alternative to traditional authentication mechanisms, as they are considered secure and useable.
- The experiment was performed with participants from diverse background and we collected the participants' age, location, and handedness, etc. Among 86 participants, 56 were males, 29 were females and 1 undisclosed with 77 right-handed and 9 left-handed. The majority of the participants were Asian (28) and European (52) continents. From the age perspective, 60 were between 20 and 30, 17 were between 30 and 40, and 3 were 40 above.
- Swipe and voice-prints were collected using Ubertesters[1] — a crowdsourcing platform for testing purposes. Ubertesters recruited approximately 150 participants worldwide for this experiment. However, we approved only 86 testers out of 150 participants based on the availability of sensors, completeness of experiment, and the quality of collected data.
- Face data of 86 users (56 males and 30 females) was obtained from the MOBIO Dataset [2].

[1] https://ubertesters.com

1. Data

The dataset, enclosed with this paper, is organized in four CSV files, i.e., swipe features (*SwipeFeatures.csv*), voice features (*VoiceFeatures.csv*), face features (*FaceFeatures.csv*), and Features vs. Weight (*FeaturesVs.Weight.csv*). Each data file, namely, *SwipeFeatures.csv, VoiceFeatures.csv, and FaceFeatures.csv* contain 10,320 rows, i.e., 10,320 observations of 86 users with 120 observations per user.

- *SwipeFeatures.csv* contains $33 \times 10,320$ observations. The columns contain the following 33 features, extracted from swipe-gesture:

No.	Swipe Features			
1–4	Duration (1)	Average event size (2)	Event size down (3)	Pressure down (4)
5–8	Start X (5)	Start Y (6)	End X (7)	End Y (8)
9–12	Velocity X Min (9)	Velocity X Max (10)	Velocity X Average (11)	Velocity X STD (12)
13–16	Velocity X VAR (13)	Velocity Y Min (14)	Velocity Y Max (15)	Velocity Y Average (16)
17–20	Velocity Y STD (17)	Velocity Y VAR (18)	Acceleration X MIN (19)	Acceleration X Max (20)
21–24	Acceleration X AVG (21)	Acceleration X STD (22)	Acceleration X VAR (23)	Acceleration Y MIN (24)
25–28	Acceleration Y Max (25)	Acceleration Y AVG (26)	Acceleration Y STD (27)	Acceleration Y VAR (28)
29–32	Pressure Min (29)	Pressure Max (30)	Pressure AVG (31)	Pressure STD (32)
33	Pressure VAR (33)	—	—	—

- *VoiceFeatures.csv* contains 104 × 10,320 observations. Columns contain 104 statistical features, namely, Mean, Standard Deviation, Kurtosis, and Skewness computed from a 2-D Mel Frequency Cepstral Coefficients (MFCC) vector of filtered voice signals. In total, 8 statistical features (each of size 1 × 13) are generated from each left and the right voice channel. Finally, these 8 vectors of size 1 × 13 are concatenated to form a single 1-D feature vector of dimension 1 × 104.
- *FaceFeatures.csv* contains 256 × 10,320 observations. Here, columns contain 256 features, computed using Binarized Statistical Image Features (BSIF) filter of size 3 × 3 with 8 bits word-length per image.
- *FeaturesVs.Weights.csv* contains Features vs. Weights for each modality using the ReliefF feature selection algorithm. Feature-wise weight is computed for each modality in unimodal settings and after their fusion in bimodal (*swipe + voice, swipe + face, voice + face*) and trimodal (*swipe + voice + face*) settings.

2. Experimental design, materials, and methods

We developed an Android customized prototype application, namely DriverAuth, that replicates the functioning of ride-booking apps (as shown in Fig. 1). DriverAuth app alerts for each new ride-assignment, which testers can accept with their voice command, to continue. The next screen shows the customer information and pick-up details, which testers can accept by swiping with their finger on the touchscreen of their smartphone. Finally, testers are prompted to take a selfie by the smartphone camera to conclude the new ride-assignment process.

The prototype application is built for Android OS (OS version *4.4.x* and above). It uses built-in hardware, i.e., touchscreen sensors and microphone, to acquire touch-based data generated as a result of swipe-gesture and recording of the user's voice.

The experiment was conducted in 4 sessions over a span of 3 days. Each user trained the application in 3 sessions with 30 training patterns per session for 15 minutes each. In the fourth session, user-tested the application for 30 times. In total 120 observations per user, comprising of 7740 (86 × 90) training samples and 2580 (86 × 30) testing samples, were collected. However, the data can be used in any ratio for the generation of training data for model training and testing, to test the trained classification model.

Our prototype applications use client-server architecture. The data generated, as a result of user's actions, i.e., swipe and voice command, was encrypted and packetized on the client device, i.e., smartphone, and was instantaneously transferred to the server, for further processing, i.e., verification of the user's identity [1].

(a) Speak to continue   (b) Swipe to accept   (c) Take the selfie

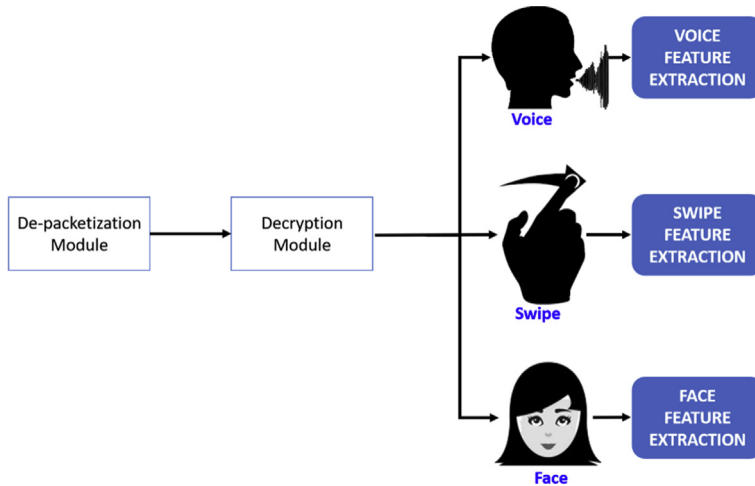**Fig. 1.** DriverAuth: A new ride-assignment process.

Fig. 2. Data de-packetization, decryption, and feature extraction to generate a classification model.

At the server end, the data is de-packetized and decrypted for features extraction (as shown in Fig. 2). Subsequently, the extracted features can be fused and ranked for generating an efficient classification model to predict between a legitimate user and an impostor.

Acknowledgments

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dib.2019.104924.

References

[1] Sandeep Gupta, Attaullah Buriro, Crispo Bruno, DriverAuth: a risk-based multi-modal biometric-based driver authentication scheme for ride-sharing platforms, Comput. Secur. 83 (June 2019) 122–139.
[2] Philip Tresadern, Chris McCool, Norman Poh, Pavel Matejka, Abdenour Hadid, Christophe Levy, Tim Cootes, Sebastien Marcel, Mobile Biometrics (Mobio): joint face and voice verification for a mobile platform, IEEE Pervasive Comput. (2012).