RHU Keystroke Touchscreen Benchmark

Mohamad El-Abed, Mostafa Dafer, Christophe Rosenberger

RHU Keystroke Touchscreen Benchmark

Mohamad El-Abed[a], Mostafa Dafer[b], and Christophe Rosenberger[c,d,e]
[a] *Rafik Hariri University*
Meshref, Lebanon
[b] *Dalhousie University*
Halifax, Nova Scotia, Canada
[c] *Université de Caen Basse-Normandie, UMR 6072 GREYC, F-14032 Caen, France*
[d] *ENSICAEN, UMR 6072 GREYC, F-14050 Caen, France*
[e] *CNRS, UMR 6072 GREYC, F-14032 Caen, France*
alabedma@rhu.edu.lb
mostafadafer@dal.ca
christophe.rosenberger@ensicaen.fr

***Abstract*—Biometric systems are currently widely used in many applications to control and verify individual's identity. Keystroke dynamics modality has been shown as a promising solution that would be used in many applications such as e-payment and banking applications. However, such systems suffer from several performance limitations (such as cross-devices problem) that prevent their widespread of use in real applications. The objective of this paper is to provide researchers and developers with a public touchscreen-based benchmark collected using a mobile phone and a tablet (both portrait and landscape orientation each). Such a benchmark can be used to assess keystroke-based matching algorithms. Furthermore, It is mainly developed to measure the robustness of keystroke matching algorithms vis-à-vis cross-devices and orientation variations. An online visualizer for the database is also given to researchers allowing them to visualize the acquired keystroke signals.**

***Keywords*-Biometrics; Keystroke Dynamics; Benchmark; Performance Evaluation.**

I. Introduction

The use of biometric systems is widely spread on various applications that require individual's authentication [Jain et al., 2004]. They are used for controlling borders, managing access to specific places, *etc.*. Nowadays, many biometric modalities exist and are divided into three categories which are morphological (such as Face [Short et al., 2015]), behavioral (such as Keystroke [Giot et al., 2009b]) and biological (such as DNA [Hashiyada, 2004]). Each modality has its own advantages/disadvantages which leads their use to a specific target application and population [El-Abed et al., 2012].

Keystroke dynamics modality has been shown as a promising solution for the banking and e-commerce sectors for the following reasons:

- Invisibility: users are used to type a password to login into information systems.
- Low cost: no additional sensor is required for using keystroke modality, all input devices feature a keyboard.

However, in order to be used in real life applications, the performance evaluation of keystroke matching algorithms should be carefully addressed. As touch-screen keyboards scale according to the screen size and its resolution, touch-screen devices have introduced new challenges that keystroke dynamics researchers and developers need to solve. Physical keyboards had standard sizes and did not need to scale for different computers; but, as we aim to extend the usability of keystroke dynamics solutions for touch-screen devices, we need to address the variability of these dynamically scaled keyboards. Furthermore, the variation of the data collected from different devices (*i.e.*, keyboard size and/or orientation) would deteriorate the overall performance of the matching algorithm, mainly the false recognition rate (FRR). More generally speaking, data collected from a physical keyboard are different from those collected using a touchscreen keyboard, and those collected from different touchscreen keyboards/orientations would be also different, *etc.* so it is of utmost importance to still be able to verify the user in case the orientation is changed or even if the device is replaced with another of a different size. For all these reasons, it would be important to provide researchers and developers with a significant public benchmark to measure the robustness of their developed algorithms vis-à-vis cross-devices variation. Therefore, the objective of this paper is to provide researchers and developers with a public touchscreen-based benchmark collected using a mobile phone and a tablet (both portrait and landscape orientation each). To the best of our knowledge, it is the first benchmark collected using different devices (mobile and tablet) and orientations (portrait and landscape). An online visualizer for the database is also given to researchers allowing them to

visualize the acquired keystroke signals.

The outline of the paper is defined as follows. Section II presents an overview of the existing keystroke dynamics benchmarks. The presentation of the *RHU Touchscreen Keystroke* benchmark along with the keystroke visualizer is given in Section III. Section IV gives a conclusion and some perspectives of this work.

II. Literature Review

Once a biometric system has the potential to become reliable, different algorithms are developed to enhance the system as best as possible. In order to evaluate and compare biometric systems' algorithms, we need to compute their performance using a predefined protocol; mainly a public benchmark. Two types of benchmarks exist:

- **Real Benchmarks**
  These databases are collected by real contributors and thus are very time and energy consuming.
- **Synthetic Benchmarks**
  These databases are automatically generated using software to simulate real biometric data and thus are very quick and easy to generate. Very few biometric modalities are concerned [Cappelli et al., 2002].

As consumers have started depending more and more on phones and switching from laptops to tablets and hybrids, it was crucial to start new benchmarks specialized in touchscreen devices, especially that touchscreen devices feature portrait and landscape orientations and come in different screen sizes, thus affecting their software keyboards. However, collecting a database is not an easy task because it requires a lot of time, energy, contributors, and sometimes special material.

A lot of the newer benchmarks have respected constraints on the way of creating good behavioral biometrics databases (in terms of number of sessions, duration between each session, number of individuals and so on [Cherifi et al., 2009]).
In this section, we present the list of existing benchmarks from the state-of-the-art. Table I shows the dataset authors, year and availability. As shown in this table, few are the available benchmarks collected using a touchscreen phone. Furthermore, to the best of our knowledge, no touchscreen benchmark exists collected with different devices (touchscreen versus tablet) and orientations (landscape versus portrait). This is what we introduce in the next section which is a benchmark collected using different devices with different orientations. Such a benchmark would be used by researchers/developers during the design and development of usable touchscreen-based keystroke matching algorithms.

Table I
Existing benchmarks' authors & year, and availability. N/A*: We could not find an online link for the dataset.

Authors	**Availability**
Bleha et al. [1990]	N/A*
Joyce and Gupta [1990]	N/A*
Leggett et al. [1991]	N/A*
Lin [1997]	N/A*
Monrose and Rubin [1997]	N/A*
Obaidat and Sadoun [1997]	N/A*
Gunetti and Picardi [2005]	Previously found at: http://www.citefa.gov.ar/si6/k-profiler/dataset/
Rodrigues et al. [2006]	N/A*
Filho and Freire [2006]	http://www.biochaves.com/en/download.htm
Hocquet et al. [2007]	N/A*
Loy et al. [2007] (Queen Mary University)	http://www.eecs.qmul.ac.uk/~ccloy/downloads_keystroke100.html
Revett et al. [2007]	N/A*
Hosseinzadeh and Krishnan [2008]	N/A*
Giot et al. [2009a]	http://www.epaymentbiometrics.ensicaen.fr/index.php?view=article&id=19
Killourhy and Maxion [2009]	http://www.cs.cmu.edu/~keystroke/
Li et al. [2011] (Beihang University)	http://mpl.buaa.edu.cn/detail1.html
Giot et al. [2012]	http://www.epaymentbiometrics.ensicaen.fr/index.php?view=article&id=20
El-Abed et al. [2014]	http://www.coolestech.com/rhu-keystroke/
Tasia et al. [2014]	http://ty.ncue.edu.tw/N25/data.html
Vural et al. [2014]	http://clarkson.edu/citer/research/collections/index.html
Sun et al. [2016]	http://cubs.buffalo.edu/research/datasets

III. RHU Keystroke Touchscreen Benchmark

A. Overview of the Android KeyStroke Collecting Tool

It is an Android application that we developed in Java using Android Studio allowing the creation of keystroke dynamics benchmarks using Android devices. A screen capture of this application is presented in Figure 1. We developed this application in order to create our own keystroke dynamics database. The data are stored online by default using an API that we created which allows quick and easy extraction of specific information.

Figure 1. Screenshot of the Android KeyStroke Collecting Tool.

The main functionality of the application is as follows:

- Possibility to create new users. All users share the same password which is "rhu.university";
- Possibility to set the user age range, gender, and occupation which are saved to the database;
- Possibility for the user to train himself to type the password using a fake account;
- Possibility to continue acquisition from users in the form of login, so data can be acquired each time a user logs in normally.
- Capturing the typing information of a user. Extracted data is added to the database; it includes:
  - timing between a key pressure and a key release
  - timing between a key release and a key pressure
  - timing between two key pressures
  - timing between two key releases
  - total time to type the password
  - screen size and orientation
- As in most of static keystroke dynamics studies, typing correction is not allowed, when a user makes a mistake, he has to type the whole password again.
- For any keystroke capture, the extracted data features stored in the database are the timing differences between two events of these kinds: press/press, release/release, press/release, release/press, and total time. They are stored in the fields PP, RR, PR, RP, and TT respectively.

We have also used the same application on a tablet (to test keyboards of different sizes and orientations).

B. KeyStroke Benchmark

Using the KeyStroke Collecting Tool, we created Keystroke Benchmark which is collected using Nexus 5 touchscreen mobile phone and Samsung Galaxy Note 10.1 2014 tablet. The purpose of this benchmark is twofold, the data provided allows us to analyze the changes in typing patterns of the same person in different orientations of the same device using the same password, it also allows us to analyze the effect of device size on the keystroke dynamics patterns when typing in the same orientation and using the same password.

The database is freely accessible to the community and downloadable through the following link: www.coolestech.com/keystroke-web-visualizer
The population composing the database is represented in Figures 2, 3, and 4.

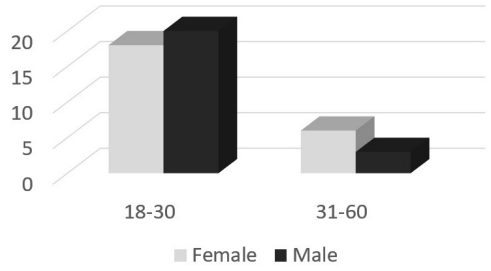


Figure 2. Population of RHU Keystroke Touchscreen Benchmark.

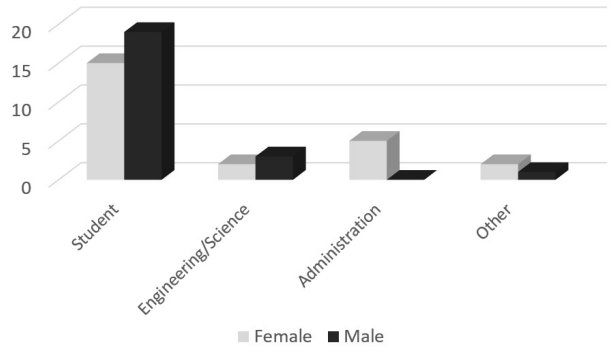


Figure 3. Distribution of Occupation Presented in RHU Keystroke Touchscreen Benchmark.

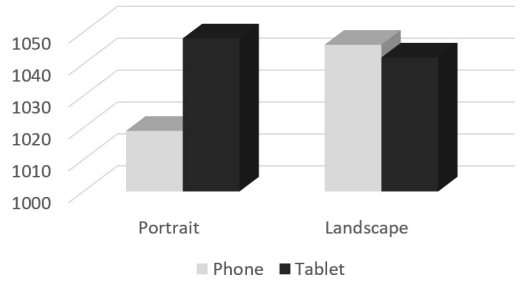


Figure 4. Trials Distribution on Different Devices and Orientations in RHU Keystroke Touchscreen Benchmark.

The hardware used to conduct this benchmark is as follows and presented in Figures 5 and 6.

HARDWARE

System chip	Qualcomm Snapdragon 800 MSM8974	Qualcomm Snapdragon 800 MSM8974
Processor	Quad-core, 2300 MHz, Krait 400	Quad-core, 2260 MHz, Krait 400
Graphics processor	Adreno 330	Adreno 330
System memory	3 GB RAM	2 GB RAM
Built-in storage	16 GB	32 GB
Storage expansion	up to 64 GB	

Figure 5. Samsung GALAXY Note 10.1 (2014 Edition) vs Google Nexus 5 Hardware PhoneArena [Retreived May 17, 2018]

DESIGN

Dimensions	9.57 x 6.75 x 0.31 inches (243.1 x 171.4 x 7.9 mm)	5.43 x 2.72 x 0.34 inches (137.84 x 69.17 x 8.59 mm)
Weight	19.29 oz (547 g) the average is 15.6 oz (445 g)	4.59 oz (130 g) the average is 5.6 oz (161 g)

DISPLAY

Physical size	10.1 inches	5.0 inches
Resolution	2560 x 1600 pixels	1080 x 1920 pixels
Pixel density	299 ppi	445 ppi
Technology	Super Clear LCD	IPS LCD
Screen-to-body ratio	70.97 %	70.89 %

Figure 6. Samsung GALAXY Note 10.1 (2014 Edition) vs Google Nexus 5 Design & Display PhoneArena [Retreived May 17, 2018]

Data representation:

- username: the username of the user
- occupation: the occupation of the user
- gender: stores the gender of the user
- age: stores the age range of the user (18-30 and 31-60)
- PP: timing between two key pressures (ms)
- PR: stores the timing between a key pressure and a key release (ms)
- RP: stores the timing between a key release and a key pressure (ms)
- RR: stores the timing between two key releases (ms)
- TT: the total time of typing the password (ms)
- Screen Orientation (1: portrait, 2: landscape)
- Screen Size (inch)

There are different sizes of touchscreen devices. Our idea is to make a data set of different keyboard sizes and screen orientations publicly available in order to be used as a reference database for testing keystroke dynamics algorithms and facilitate the comparison of previous and future keystroke dynamics authentication methods.

Forty seven (47) individuals have participated in the acquisition process by typing the password "rhu.university". In each session, an individual typed the password in all four configurations: phone/portrait, phone/landscape, tablet/portrait and tablet/landscape. At least 60 acquisitions were collected per session. In general, each user has 20 trials using each device orientation. In total, we have 4155 available acquisitions. All users were able to train themselves on the typing of the password on the keyboard as long as they wanted, because it is not their usual password and they do not have a pre-existing typing habit or pattern for it. In addition, touch smartphones' operating systems have different keyboards, and we wanted them to get used to the different keyboard sizes.

C. Keystroke Visualizer

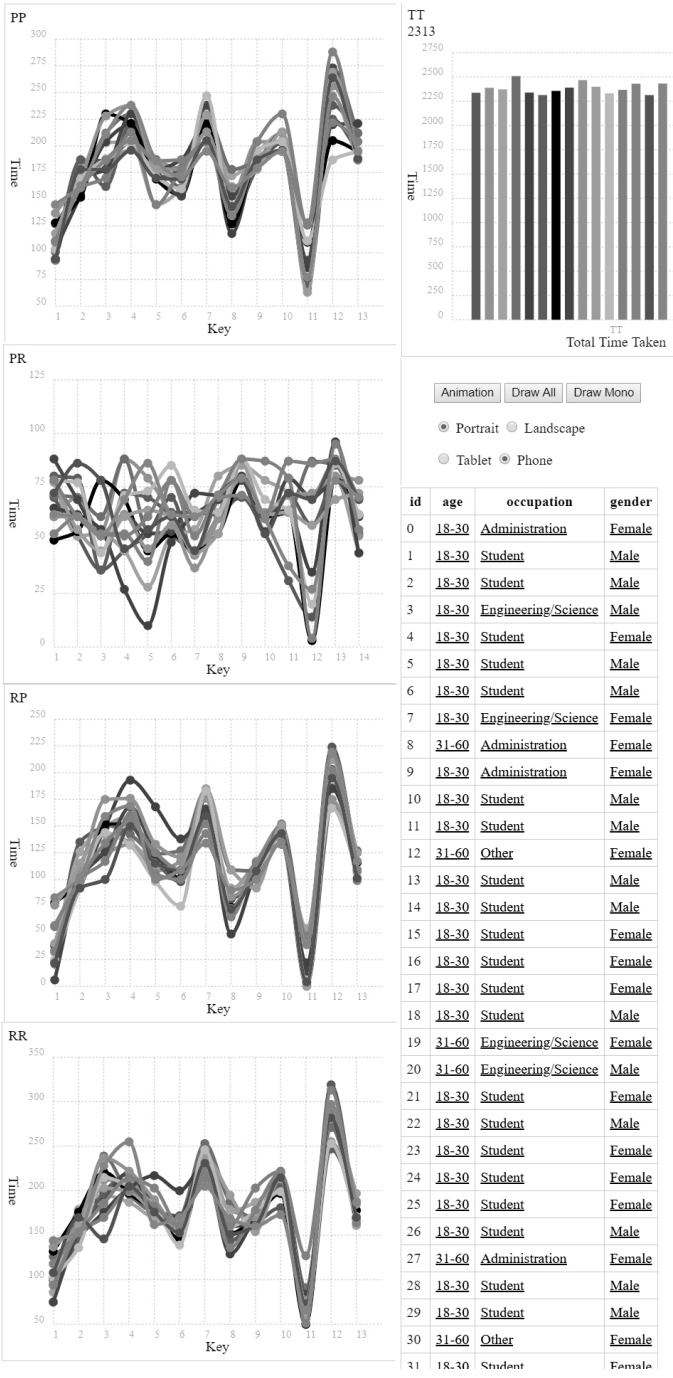
KeyStroke Visualizer, shown in Figure 7, is a web application that we developed allowing the visualization of a keystroke dynamics database created using the KeyStroke Collecting Tool.

We developed this application in order to visualize our own keystroke dynamics database. KeyStroke Visualizer simply connects to the database and automatically collects all information in order to visualize them on-demand.

The functionality of the application is as follows:

- A list of all users is automatically retrieved from the database when the application is started.
- The application provides each user's gender, age range and profession.
- By selecting a user, the application draws all PP, PR, RP, RR and TT (Total Time) graphs.
- Each trial is colored differently for easier differentiation between trials.

- Each point value in the chart can be viewed separately by hovering the cursor over it.
- Trials can be viewed in mono color (grey).
- Trials can be filtered to view the first 14 trials only for clearer view.
- Animation can be applied to the graphs for easier tracing of different trials.



id	age	occupation	gender
0	18-30	Administration	Female
1	18-30	Student	Male
2	18-30	Student	Male
3	18-30	Engineering/Science	Male
4	18-30	Student	Female
5	18-30	Student	Male
6	18-30	Student	Male
7	18-30	Engineering/Science	Female
8	31-60	Administration	Female
9	18-30	Administration	Female
10	18-30	Student	Male
11	18-30	Student	Male
12	31-60	Other	Female
13	18-30	Student	Male
14	18-30	Student	Male
15	18-30	Student	Female
16	18-30	Student	Female
17	18-30	Student	Female
18	18-30	Student	Male
19	31-60	Engineering/Science	Female
20	31-60	Engineering/Science	Male
21	18-30	Student	Female
22	18-30	Student	Male
23	18-30	Student	Female
24	18-30	Student	Female
25	18-30	Student	Female
26	18-30	Student	Male
27	31-60	Administration	Female
28	18-30	Student	Male
29	18-30	Student	Male
30	31-60	Other	Female
31	18-30	Student	Female

Figure 7. Screenshot of the Database Visualization Tool (KeyStroke Online Visualizer)

IV. Conclusion and Perspectives

We have presented in this paper a public benchmark named *RHU Keystroke Touchscreen* that would be used by researchers and developers to assess the performance of their developed keystroke authentication system. It is mainly developed to measure the robustness of keystroke matching algorithms vis-à-vis cross-devices and orientation variations. To the best of our knowledge, it is the first benchmark collected using different devices (mobile and tablet) and orientations (portrait and landscape).

We have used this benchmark to make a statistical analysis using the Kruskal-Wallis test (KW) [Higgins, 2003] to check the variation of the collected keystroke features from the different devices (mobile and tablet) and orientations (portrait and landscape). The results from this study recommend removing the *Press-Release* feature when dealing with touchscreen-based keystroke matching algorithms. The experimental results show significant differences between keystroke features collected using different devices and orientations. Such differences should be taken into consideration when designing touchscreen-based keystroke authentication systems.

As for the perspectives of this work, we aim to use the presented benchmark to develop a novel keystroke matching algorithm that would be robust against cross-devices and orientations problem. It would be also important to compare existing keystroke matching algorithms (mainly those based on data mining approaches) to measure their robustness on the presented variations. Extending the presented benchmark to other devices would be also important to assess the overall performance of keystroke matching algorithms. As a behavior modality, it would be also important to develop a quality index to assess the quality of the acquired keystroke signals to improve the overall performance of keystroke-based matching algorithms.

References

S. Bleha, C. Slivinsky, and B. Hussien. Computer-access security systems using keystroke dynamics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12: 1217 – 1222, 1990.

R. Cappelli, D. Maio, and D. Maltoni. Synthetic fingerprint-database generation. In *International Conference on Pattern Recognition (ICPR)*, pages 744–747, 2002.

F. Cherifi, B. Hemery, R. Giot, M. Pasquet, and C. Rosenberger. Performance evaluation of behavioral biometric systems. In *Behavioral Biometrics for Human Identification: Intelligent Applications*, pages 57–74, 2009.

M. El-Abed, C. Charrier, and C. Rosenberger. *New Trends and Developments in Biometrics*, chapter Evaluation of Biometric Systems, pages 149–169. InTech, 2012.

M. El-Abed, M. Dafer, and R. El Khayat. RHU keystroke: A mobile-based benchmark for keystroke dynamics systems. In *IEEE International Carnahan Conference on Security Technology (ICCST)*, pages 1–4, 2014.

J. R. M. Filho and E. O. Freire. On the equalization of keystroke timing histograms. 27:1440–1446, 2006.

R. Giot, M. El-Abed, and C. Rosenberger. Greyc keystroke: A benchmark for keystroke dynamics biometric systems. In *IEEE Third International Conference on Biometrics : Theory, Applications and Systems (BTAS)*, pages 1–6, 2009a.

R. Giot, M. El-Abed, and C. Rosenberger. Keystroke dynamics with low constraints SVM based passphrase enrollment. In *IEEE Third International Conference on Biometrics : Theory, Applications and Systems (BTAS)*, pages 425–430, 2009b.

R. Giot, M. El-Abed, and C. Rosenberger. Web-based benchmark for keystroke dynamics biometric systems: A statistical analysis. In *Proceedings of the IIHMSP 2012 conference*, pages 11 – 15, 2012.

D. Gunetti and C. Picardi. Keystroke analysis of free text. *ACM Transactions on Information and System Security*, 8:312–347, 2005.

M. Hashiyada. Development of biometric dna ink for authentication security. *Tohoku Journal of Experimental Medicine*, pages 109–117, 2004.

J. J. Higgins. An introduction to modern nonparametric statistics. *The American Statistician*, 2003.

S. Hocquet, J.Y. Ramel, and H. Cardot. User classification for keystroke dynamics authentication. In *International Conference on Biometrics (ICB)*, pages 531–539, 2007.

D. Hosseinzadeh and S. Krishnan. Gaussian mixture modeling of keystroke patterns for biometric applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 38(6):816–826, 2008.

A. K. Jain, S. Pankanti, S. Prabhakar, L. Hong, and A. Ross. Biometrics: A grand challenge. *International Conference on Pattern Recognition (ICPR)*, 2:935–942, 2004.

R. Joyce and G. Gupta. Identity authentication based on keystroke latencies. *Commun. ACM*, 33(2), 1990.

K. S. Killourhy and R. A. Maxion. Comparing anomaly-detection algorithms for keystroke dynamics. In *International Conference on Dependable Systems and Networks*, pages 125 – 134, 2009.

J. Leggett, G. Williams, and M. Usnick. Dynamic identity verification via keystroke characteristics. *International Journal of Man-Machine Studies*, 35:859–870, 1991.

Y. Li, B. Zhang, Y. Gao, S. Zhao, and J. Liu. Study on the beihang keystroke dynamics database. In *International Joint Conference on Biometrics (IJCB)*, pages 1 – 5, 2011.

D. T. Lin. Computer-access authentication with neural network based keystroke identity verification. In *IEEE Int. Conf. Neural Netw*, pages 174–178, 1997.

C. Loy, W. K. Lai, and C. P. Lim. Keystroke patterns classification using the artmap-fd neural network. In *IEEE Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 61 – 64, 2007.

F. Monrose and A. Rubin. Authentication via keystroke dynamics. In *Proceedings of the 4th ACM conference on Computer and communications security*, pages 48–56. ACM, 1997.

S. Obaidat and B. Sadoun. Verification of computer users using keystroke dynamics. *IEEE Trans Syst Man Cybern B Cybern*, 27(2), 1997.

PhoneArena. Samsung galaxy note 10.1 (2014 edition) vs google nexus 5. https://www.phonearena.com/phones/compare/Samsung-GALAXY-Note-10.1-2014-Edition,Google-Nexus-5/phones/8139,8148, Retreived May 17, 2018.

K. Revett, S.T. de Magalhaes, and H.M.D. Santos. On the use of rough sets for user authentication via keystroke dynamics. *Lecture Notes in Computer Science*, 4874:145–159, 2007.

R. Rodrigues, G. Yared, C. do N. Costa, J. Yabu-Uti, F. Violaro, and L. Ling. Biometric access control through numerical keyboards based on keystroke dynamics. In *International Conference on Biometrics (ICB)*, volume 3832, pages 640–646, 2006.

N. Short, S. Hu, P. Gurram, and K. Gurton. Exploiting polarization-state information for cross-spectrum face recognition. In *IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1 – 6, 2015.

Y. Sun, H. Ceker, and S. Upadhyaya. Shared keystroke dataset for continuous authentication. In *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6, Dec 2016. doi: 10.1109/WIFS.2016.7823894.

C. J. Tasia, T. Y. Chang, P. C. Cheng, and J. H. Lin. Two novel biometric features in keystroke dynamics authentication systems for touch screen devices. *Sec. and Commun. Netw.*, 7(4):750–758, 2014.

E. Vural, J. Huang, D. Hou, and S. Schuckers. Shared research dataset to support development of keystroke authentication. In *International Joint Conference on Biometrics (IJCB)*, pages 1 – 8, 2014.