Multi-ID: A Novel Dataset for Advancing Understanding of Cross-Task and Cross-Device Input Dynamics

Meghna Chaudhary, Saandeep Aathreya, Parush Gera, Shaun Canavan, Tempestt Neal
Department of Computer Science and Engineering, University of South Florida, Tampa, FL USA
tjneal@usf.edu

July 22, 2024

Abstract

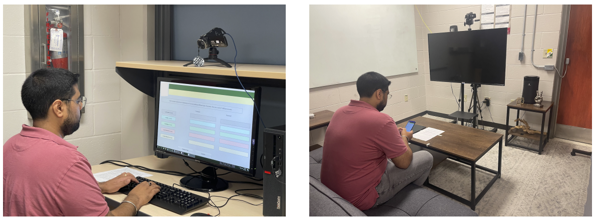
This paper presents the Multi-ID (Multiple Inputs and Devices) Dataset, a collection of keystroke, mouse, and touch dynamics, obtained from over 30 participants engaged in diverse tasks on both desktop computers and mobile phones. Multi-ID serves as a valuable asset for advancing collaborative, multi-device scenarios, laying the groundwork for the development of systems that can adapt to diverse user preferences and device usage patterns. One application of this dataset lies in security-focused applications spanning multiple devices. In particular, in-depth analyses of keystrokes, touch gestures, and mouse dynamics across different devices, as facilitated by Multi-ID, could inform the development of robust cross-device user authentication systems. This is particularly crucial in collaborative settings where participants frequently transition between devices. However, the challenge of cross-device authentication is substantial due to intra-person variation in input dynamics across devices, even when inputting the same phrase or gesture. To underscore this challenge and showcase the utility of Multi-ID, we conducted user authentication experiments utilizing three classifiers and various train–test scenarios involving fixed and free-form keystroke data. These experiments encompass within- and across-task and device scenarios, revealing notably low equal error rates, reflecting the very challenging yet practical scenario of cross-device user authentication.

1 Introduction

This paper presents the Multi-ID (Multiple Inputs and Devices) Dataset, an on-going collection of keystroke, mouse, and touch dynamics, collected thus far from over 30 research participants from both a desktop computer and a mobile phone as they performed various tasks. The Multi-ID dataset holds immense potential for enhancing collaborative scenarios. For example, collaborative systems that integrate both mobile devices and desktop computers could leverage insights from the Multi-ID dataset to inform the development of adaptive interfaces to ensure seamless transitions and consistent user experiences across devices [26]. Further, Multi-ID could play a crucial role in shaping the design of collaborative applications in specialized areas like virtual collaboration, where participants interact through various devices [20, 24], among other applications.

Another significant application of the Multi-ID dataset is its potential as a valuable resource for improving security mechanisms. Specifically, Multi-ID could contribute to the creation of intelligent authentication mechanisms that prioritize both security and user convenience, particularly in settings where users frequently switch between devices, by aiding understanding of the unique input patterns associated with each user and device combination.

However, user authentication across devices is challenging, particularly considering the dynamics captured in the Multi-ID dataset. For instance, previous studies have investigated the influence of various devices on keystroke-based authentication systems. Acien et al. [1] employed stacked Long Short-Term Memory networks to authenticate 168,000 subjects based on 136 million keystrokes collected from both physical and touchscreen keyboards. Their findings revealed a significant increase (21.4%) in equal error rates (EERs) — the point at which the false acceptance rate (FAR) equals the false rejection rate (FRR), providing a balanced assessment of the system's performance — when training on the physical keyboard dataset and testing on the touchscreen keyboard dataset, compared to training and testing on the same physical keyboard dataset (2.2%). Similarly, Alsuhibany and Almuqbil [2] investigated the persistence of typing patterns across different devices, observing an increase in EERs from 0% (training and testing on the same device) to 6% (training and testing on different devices). Further, several factors are known to impact the performance of biometric systems using input dynamics, like keystroke, touch, and mouse dynamics, for authentication. For instance, keystroke time intervals are influenced by users' ages [33, 13] and by the typing task at hand [21], while touch gestures are known to differ for the same user across different mobile applications [12]. On the other hand, input dynamics are common identifiers for user authentication on mobile devices, laptops, and workstations



(a) Desktop setup

(b) Phone setup

Figure 1: Data collection setup.

[18, 16, 8, 19, 11, 9, 35, 22, 7, 33]. Consequently, the design of cross-device and multi-input-based authentication systems is challenging yet warrants increased attention.

2 The Multi-ID Dataset

The Multi-ID Dataset includes input dynamics collected from research participants aged 6 and older, with 3 individuals aged 17 and under, 18 individuals aged 18 to 29, 4 individuals aged 30 to 49, and 7 individuals aged 50 or older. At the time of this writing, the dataset includes data from 32 participants (13 female and 19 male; 12 Asian, 11 White, 4 Middle Eastern or North African, 3 Black or African American, and 2 Hispanic, Latino, or of Spanish origin). Additionally, there are 21 iOS users and 11 Android users.

Before participating in the study, individuals undergo an initial virtual meeting where a researcher explains the study's purpose and details, and participants read and sign a consent form. Once participants agree to take part in this study, they schedule a time to come into the lab. During their first session, participants complete a demographic form. For children under 18 years old, demographic information is provided by their accompanying guardian.

Participants attend three separate sessions in the lab, each involving tasks performed on a Lenovo ThinkCentre M710 workstation and a OnePlus Nord N10 5G smartphone. In each session, various input dynamics are collected transparently. This includes fixed (participants type a provided phrase) and free-form (participants enter a phrase of their choice) keystroke dynamics, directed (participants navigate using the mouse based on provided tasks) and free-form (participants use the mouse as they please) mouse dynamics, as well as touch dynamics. The data collection setup is illustrated in Figure 1, and screenshots for Tasks 2, 3, and 4 can be found in Figure 2, as described below:

Task 1: Essay Compose a provided essay (approximately 500 words for adults and 150 words for children) on the desktop computer. Children engage with a shorter, child-friendly story instead of a lengthy essay for better participation and to minimize fatigue. Collected desktop modalities include fixed keystroke and free-form mouse dynamics.

Task 2: Password Entry Enter three passwords five times each on both the desktop and smartphone (`GmxPV3L`, `Nv5PHS!8kP8`, and `jxK&5sDpwfE+U` for adults, and `schoolRocks`, `g@me&play`, and `GmxPV3L` for children). These passwords are designed to vary in difficulty, incorporating special characters, numbers, and a mix of lower and uppercase letters. Desktop modalities include fixed keystroke and free-form mouse dynamics, while phone modalities include fixed keystroke and free-form touch dynamics.

Task 3: Recipe Search Search the web for a recipe of the participant's choice on the desktop. Collected desktop modalities encompass free-form keystroke and both directed and free-form mouse dynamics. Participants can opt to use a provided website for a guided search (directed mouse) or freely explore the internet (free-form mouse).

Task 4: Mock Credentials Generate mock username and password combinations as if setting up personal email, utility, banking, work-hosted email, and school-hosted email accounts on both the workstation and smartphone. Mock accounts for children include a streaming service (e.g., YouTube), a chat account for communication with friends, a banking or money-saving account, a gaming account, and a school account. Collected desktop modalities involve free-form keystroke and mouse dynamics, while phone modalities encompass free-form keystroke and touch dynamics.

Task 5: Text Messaging Engage in a brief conversation with a member of the research team via phone, simulating text messaging. The exchange is tailored to the participant's age group, addressing their study experience, covering topics such as:

1. Identifying the easiest and hardest tasks during the study.
2. Inquiring about their typical authentication methods on their own device.
3. Exploring reasons for not using any authentication method, including a password.
4. Understanding the participant's emotions when their chosen authentication method fails.
5. Gauging their opinion on continuous authentication approaches, such as those based on typing patterns.

Note that this task captures both input dynamics and the chat itself, and both components are incorporated into the Multi-ID Dataset. Phone modalities collected encompass free-form keystroke and touch dynamics.

To enhance participation across all three sessions, we offer flexible scheduling options for participants, allowing them to coordinate sessions based on their availability. Instead of scheduling sessions on set days, we send reminder emails after each session, prompting participants to schedule the next one. As a result, the time between sessions varies, averaging 13 days. The average duration of each session is 22, 19, and 19 minutes, respectively, contributing to approximately 32 hours of data thus far. This data collection has received Human Subjects approval from the University of South Florida's Institutional Review Board (STUDY002291). Participants receive compensation in the form of a $35 e-gift card per completed session.

To our knowledge, the Multi-ID Dataset is only one of two (in addition to the BB-MAS dataset [5]) to consist of three different input dynamics, and the only to include fixed and free-form keystroke tasks, directed and free-form mouse tasks, and touch gestures, each collected across three sessions per subject. Table 1 provides a summary of comparable datasets that include keystroke, touch, and mouse data. While these datasets include various input dynamics, only one [5] includes all three input dynamics, while the others do not cover across device or task scenarios, have limited demographic details, or consider a single session or multiple sessions within the same day.

3 Authentication Experiments

We explored four train/test scenarios to evaluate authentication performance using the Multi-ID dataset across sessions, devices, and tasks to demonstrate its use for assessing multi-input user authentication, including:

1. Training and testing using data collected on a single device within a single task.
2. Training and testing using data collected on a single device from different tasks.
3. Training and testing using data collected on different devices within a single task.
4. Training and testing using data collected on different devices from different tasks.

In each of these scenarios, the training data consisted of a single session, while the testing data consisted of a future session (i.e., session 1 as training data and session 2 as testing data, session 2 and 3 as training and testing data, respectively, and session 1 and 3 as training and testing data, respectively).

3.1 Methodology

We extracted keystroke dynamic features from data collected within the Multi-ID Dataset Tasks 2 and 4; these tasks include desktop and phone data that reflect identical activities across both devices. We extracted features common in other keystroke dynamic-based approaches [4], including

1. *Press-to-Press*: Time between pressing the first key and pressing the next key.
2. *Press-to-Release*: Time between pressing the first key and releasing the same key.
3. *Release-to-Press*: Time between releasing the first key and pressing the next key.
4. *Release-to-Release*: Time between releasing the first key and releasing the next key.
5. *Press-to-Release*: Time between pressing the first key and releasing the next key.

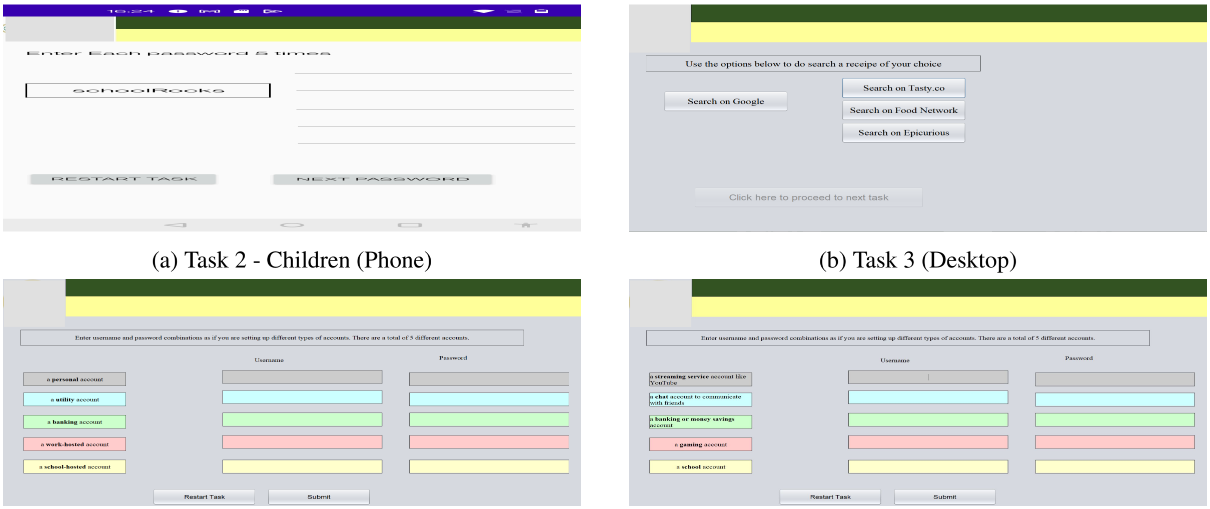
We trained three different machine learning classifiers using these features, including a support vector machine (SVM), decision tree (DT), and a logistic regression (LR) classifier per participant using a one-versus-rest approach to conduct user authentication experiments. Each classifier was tuned using a grid search (*SVM*: kernel: [linear, rbf, poly], C: [1, 10], gamma: [auto]; *DT*: max_depth: [1, 2, 3, 4], criterion: [gini, entropy]; and *LR*: penalty: [l1, l2, elasticnet, None], solver: [liblinear, newton-cholesky], C: [1, 10]).

3.2 Results

Figure 3 plots the range of EERs across all devices and tasks. In this figure, the x-axis indicates the training and testing set, where X/Y indicates the training (X) data and testing (Y) data, D indicates data extracted from the desktop, P indicates data extracted from the phone, and T2 and T4 indicate Task 2 and Task 4, respectively. Here, we briefly summarize overall findings to elucidate open challenges that this dataset might help address.

Training with Desktop/Fixed The first subplot shows the range of EERs achieved when training on data gathered during Task 2 on the desktop (D-T2). When testing with D-T2, illustrating a same device, same task train/test scenario, we find that the EER distribution exhibits lower variance with a slightly lower median compared to other testing scenarios. On the other hand, the largest median EER value is observed in the different device, same task scenario (i.e., D-T2 vs P-T2), although the variance in EER values is smallest. Notably, the largest range of EER values is observed in the most challenging scenario of different device, different task (i.e., D-T2 vs P-T4).

Training with Phone/Fixed When training authentication systems on data collected during Task 2 on the phone, we found all EER distributions tightly clustered. This could suggest that the matching scores between enrolled and query features are generally the same, despite the test set. However, because most of these EERs range around 0.50, these results also indicate that our particular experimental setup leads to consistently producing false accept and false reject rates that show 50% of users are either falsely authenticated or falsely rejected. This observation contrasts other training scenar-



Enter Each password 5 times

schoolRocks

Use the options below to do search a receipe of your choice

Enter username and password combinations as if you are setting up different types of accounts. There are a total of 5 different accounts.

	Username	Password
a **personal** account		
a **utility** account		
a **banking** account		
a **work-hosted** account		
a **school-hosted** account		

Enter username and password combinations as if you are setting up different types of accounts. There are a total of 5 different accounts.

	Username	Password
a **streaming service** account like YouTube		
a **chat** account to communicate with friends		
a **banking or money savings** account		
a **gaming** account		
a **school** account		

(a) Task 2 - Children (Phone)

(b) Task 3 (Desktop)

(c) Task 4 - Adults (Desktop)

(d) Task 4 - Children (Workstation)

Figure 2: Screenshots of input screens used in our data collection procedures for Tasks 2, 3, and 4.

Table 1: Comparable publicly available datasets of input dynamics. The *Across Devices* column indicates if the dataset contains keystrokes (note that touch and mouse are specific to certain devices, while keystroke data can be collected from multiple devices) collected from full physical keyboards (such as desktop and laptop keyboards) and touch-based keyboards (such as those on mobile phones and tablets). The *Demographics* column indicates if the dataset includes ethnicity, gender, and age of its participants. An asterick indicates datasets with additional modalities beyond keystroke, touch, and mouse dynamics.

Dataset	Modalities	Subjects	No. of Sessions	Across Devices	Demographics
BB-MAS* [6]	Fixed keystroke, free-form keystroke, touch, free-form mouse	117	2	Yes	Yes
HMOG* [30]	Free-form keystrokes, touch	100	8 (same day)	No	No
BrainRun* [25]	Touch	2,218	1	No	Yes
Touchalytics [14]	Touch	41	1 (optional follow-up session)	No	No
LTU Touch [28]	Touch	190	2	No	No
CU Mobile I & II* [27]	Fixed keystroke, free-form keystroke, touch	88	8 (same day)	No	No
UB [31]	Fixed keystroke, free-form keystroke, fixed and free form mouse	157	3	No	Yes
CMU Keystroke [17]	Fixed Keystrokes	51	8	No	Yes
GREYC Keystroke [15]	Fixed Keystrokes	113	5	No	Yes
Soft Biometrics Database [32]	Fixed Keystrokes	120	1	No	Yes
Multi-K [34]	Free-form Keystrokes	86	1	No	No
Clarkson II [23]	Free-form Keystrokes, Free-form mouse	113	Longitudinal (2.5 Years)	No	No
Aalto [10]	Fixed Keystrokes	168,000	1	No	Yes
Balabit Mouse Challenge [29]	Free-form Mouse	10	1	No	No
DFL [3]	Free-form Mouse	21	1	No	No
Multi-ID	**Fixed keystroke, free-form keystroke, directed mouse, free-form mouse, touch**	**30+**	**3**	**Yes**	**Yes**

ios, where some EERs are much lower (P-T4/P-T4) or have a greater range (D-T2/P-T4).

Training with Desktop/Free-form It is interesting to find contrasting ranges of EERs associated with training an authentication system using free-form keystroke dynamics collected from the desktop, particularly when compared to training with fixed keystroke dynamics from a desktop (e.g., D-T2/P-T4 vs D-T4/P-T4). Such differences highlight variations in typing patterns depending on how the typist is instructed to provide input. However, the amount of variance in EERs is relatively consistent with the Desktop/Fixed training scenario, where, besides X/P-T4, the range of EER values is small. Thus, we suspect that while the task impacts one's typing dynamics, the task used for testing also might lead to some expected range of accuracy assuming the training device and task is the same.

Training with Phone/Free-form When using data collected during Task 4 from the phone, we found the lowest EER median for the same device, same task scenario (i.e., P-T4/P-T4), in addition to the same device, different task scenario (i.e., P-T4/P-T2). There is a sharp increase in EERs (from approximately 0.20-0.23 to 0.40-0.50) when changing the test set to data collected from a desktop. We note that this observation was not as prominent when training on data collected from the desktop and testing on data collected from the phone.

Overall, this small study demonstrates the impact of training on the same versus different devices and the implications of performing authentication tasks on the same versus different tasks. Notably, when training and testing occur on the same device and task, the EER distribution exhibits lower variance with a slightly lower median. However, the challenge arises when dealing with different devices and tasks, leading to increased EER values and wider ranges. This underscores the importance of understanding the dynamics of authentication across various scenarios. The introduction of the Multi-ID Dataset emerges as a promising solution to address these and other challenges. For instance, promising avenues of research facilitated by the Multi-ID Dataset might include

Device-Agnostic Authentication Training authentication models using data from one device and testing on data collected from another device, investigating the extent to which models can generalize across different devices, identifying potential challenges and opportunities for device-agnostic authentication.

Task-Dependent Authentication Examining authentication performance across different tasks within the same device and across devices, assessing the impact of task-specific behaviors on authentication accuracy and exploring whether certain tasks contribute more to authentication challenges.

Keystroke Dynamics Exploration Analyzing fixed and free-form keystroke dynamics separately, understanding how the choice of input (provided phrase vs. user's choice) influences authentication accuracy.

Mouse and Touch Dynamics Investigation Investigating correlations between mouse and touch dynamics.

Multi-Modal Fusion Approaches Develop multi-modal authentication models by fusing data from keystrokes, mouse movements, and touch dynamics, evaluating the effectiveness of combining multiple input modalities in enhancing overall authentication accuracy and resiliency to device and task variations.

By conducting these experiments, researchers can gain deeper insights into the complexities of multi-input, cross-device, and cross-task user authentication, paving the way for more robust and adaptable authentication systems in real-world settings. Further, findings from these experiments can significantly enhance collaborative systems or environments by informing the development of user authentication protocols that seamlessly accommodate diverse devices and tasks, fostering a user-friendly and adaptable collaborative ecosystem. Incorporating insights from multi-input analyses can lead to authentication systems that intelligently adapt to users' behaviors across devices and tasks, promoting a secure and efficient collaborative environment where users can seamlessly interact and collaborate without unnecessary authentication hindrances.

4 Conclusion

This article presents the initial version of the Multi-ID Dataset, which includes keystroke, touch, and mouse dynamics from both a workstation and a mobile phone, encompassing free-form and fixed text input and web browsing. Initial experiments demonstrate the potential of this dataset to not only expose real-world security challenges in various within- and across-tasks and device scenarios, but to also help facilitate the resolution of such challenges by, for example, the development of device-agnostic or task-dependent authentication models.

Importantly, the Multi-ID Dataset presents a wealth of opportunities for exploration and advancement across various domains. Beyond its implications for user authentication and collaborative systems, researchers can leverage this dataset for human-computer interaction studies, offering insights into user preferences and interactions and aiding in the design of more intuitive and user-friendly interfaces. Additionally, the dataset's cross-device and cross-task nature makes it valuable for understanding user behaviors in diverse contexts, potentially informing the development of personalized computing experiences. Its applications could also extend to fields such as accessibility, where insights into user interactions can contribute to designing inclusive technologies. Overall, the Multi-ID Dataset serves as a versatile tool for multidisciplinary research, encompassing human-computer interaction, personalized computing, the CSCW community, accessibility, and beyond.

5 Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 2039373. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

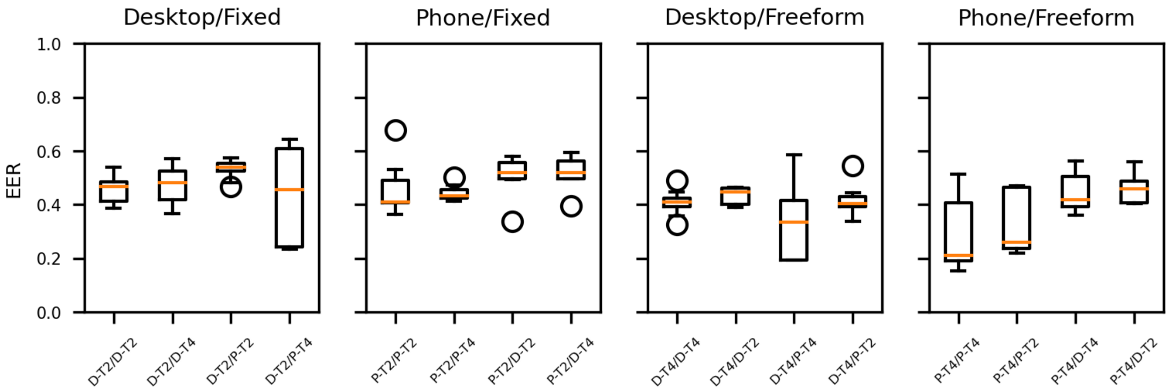


Figure 3: Range of EERs across all devices and tasks. The x-axis indicates the training and testing set, where X/Y indicates the training (X) data and testing (Y) data, D indicates data extracted from the desktop, P indicates data extracted from the phone, and T2 and T4 indicate Task 2 and Task 4, respectively. Each plot's title provides the training data.

References

[1] Alejandro Acien, Aythami Morales, John V. Monaco, Ruben Vera-Rodriguez, and Julian Fierrez. Typenet: Deep learning keystroke biometrics. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 4(1):57–70, 2022.

[2] Suliman A Alsuhibany and Afnan S Almuqbil. Impact of using different-sized touch keyboards on free-text keystroke dynamics authentication in the arabic language. *Scientific Reports*, 12(1):15866, 2022.

[3] Margit Antal and Lehel Denes-Fazakas. User verification based on mouse dynamics: a comparison of public data sets. In *2019 IEEE 13th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, pages 143–148, 2019.

[4] Salil P Banerjee and Damon L Woodard. Biometric authentication and identification using keystroke dynamics: A survey. *Journal of Pattern Recognition Research*, 7(1):116–139, 2012.

[5] Amith K. Belman and Vir V. Phoha. Discriminative power of typing features on desktops, tablets, and phones for user identification. *ACM Trans. Priv. Secur.*, 23(1), feb 2020.

[6] Amith K Belman, Li Wang, SS Iyengar, Pawel Sniatala, Robert Wright, Robert Dora, Jacob Baldwin, Zhanpeng Jin, and Vir V Phoha. Insights from bb-mas–a large dataset for typing, gait and swipes of the same person on desktop, tablet and phone. *arXiv preprint arXiv:1912.02736*, 2019.

[7] Lucia Cascone, Michele Nappi, Fabio Narducci, and Chiara Pero. Touch keystroke dynamics for demographic classification. *Pattern Recognition Letters*, 158:63–70, 2022.

[8] Robert Cockell and Basel Halak. On the design and analysis of a biometric authentication system using keystroke dynamics. *Cryptography*, 4(2), 2020.

[9] Ingo Deutschmann, Peder Nordström, and Linus Nilsson. Continuous authentication using behavioral biometrics. *IT Professional*, 15(4):12–15, 2013.

[10] Vivek Dhakal, Anna Maria Feit, Per Ola Kristensson, and Antti Oulasvirta. Observations on typing from 136 million keystrokes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–12, New York, NY, USA, 2018. Association for Computing Machinery.

[11] El-Sayed M. El-Kenawy, Seyedali Mirjalili, Abdelaziz A. Abdelhamid, Abdelhameed Ibrahim, Nima Khodadadi, and Marwa M. Eid. Metaheuristic optimization and keystroke dynamics for authentication of smartphone users. *Mathematics*, 10(16), 2022.

[12] Tao Feng, Jun Yang, Zhixian Yan, Emmanuel Munguia Tapia, and Weidong Shi. Tips: Context-aware implicit user identification using touch screen in uncontrolled environments. In *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications*, HotMobile '14, New York, NY, USA, 2014. Association for Computing Machinery.

[13] Gianni Fenu, Mirko Marras, and Ludovico Boratto. A multi-biometric system for continuous student authentication in e-learning platforms.

Pattern Recognition Letters, 113:83–92, 2018. Integrating Biometrics and Forensics.

[14] Mario Frank, Ralf Biedert, Eugene Ma, Ivan Martinovic, and Dawn Song. Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. *IEEE Transactions on Information Forensics and Security*, 8(1):136–148, 2013.

[15] Romain Giot, Mohamad El-Abed, and Christophe Rosenberger. Greyc keystroke: A benchmark for keystroke dynamics biometric systems. In *2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, pages 1–6, 2009.

[16] Pilsung Kang and Sungzoon Cho. Keystroke dynamics-based user authentication using long and free text strings from various input devices. *Information Sciences*, 308:72–93, 2015.

[17] Kevin S. Killourhy and Roy A. Maxion. Comparing anomaly-detection algorithms for keystroke dynamics. In *2009 IEEE/IFIP International Conference on Dependable Systems Networks*, pages 125–134, 2009.

[18] Junhong Kim and Pilsung Kang. Freely typed keystroke dynamics-based user authentication for mobile devices based on heterogeneous features. *Pattern Recognition*, 108:107556, 2020.

[19] Sowndarya Krishnamoorthy, Luis Rueda, Sherif Saad, and Haytham Elmiligi. Identification of user behavioral biometrics for authentication using keystroke dynamics and machine learning. In *Proceedings of the 2018 2nd International Conference on Biometric Engineering and Applications*, ICBEA '18, page 50–57, New York, NY, USA, 2018. Association for Computing Machinery.

[20] Stefan Marks and David White. Multi-device collaboration in virtual environments. In *Proceedings of the 2020 4th International Conference on Virtual and Augmented Reality Simulations*, ICVARS '20, page 35–38, New York, NY, USA, 2020. Association for Computing Machinery.

[21] John V. Monaco and Charles C. Tappert. The partially observable hidden markov model and its application to keystroke dynamics. *Pattern Recognition*, 76:449–462, 2018.

[22] Soumik Mondal and Patrick Bours. A study on continuous authentication using a combination of keystroke and mouse biometrics. *Neurocomputing*, 230:1–22, 2017.

[23] Christopher Murphy, Jiaju Huang, Daqing Hou, and Stephanie Schuckers. Shared dataset on natural human-computer interaction to support continuous authentication research. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*, pages 525–530, 2017.

[24] Patrick Aggergaard Olin, Ahmad Mohammad Issa, Tiare Feuchtner, and Kaj Grønbæk. Designing for heterogeneous cross-device collaboration and social interaction in virtual reality. In *Proceedings of the 32nd Australian Conference on Human-Computer Interaction*, OzCHI '20, page 112–127, New York, NY, USA, 2021. Association for Computing Machinery.

[25] Michail D. Papamichail, Kyriakos C. Chatzidimitriou, Thomas Karanikiotis, Napoleon-Christos I. Oikonomou, Andreas L. Symeonidis, and Sashi K. Saripalle. Brainrun: A behavioral biometrics dataset towards continuous implicit authentication. *Data*, 4(2), 2019.

[26] Seonwook Park, Christoph Gebhardt, Roman Rädle, Anna Maria Feit, Hana Vrzakova, Niraj Ramesh Dayama, Hui-Shyong Yeo, Clemens N. Klokmose, Aaron Quigley, Antti Oulasvirta, and Otmar Hilliges. Adam: Adapting multi-user interfaces for collaborative environments in real-time. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–14, New York, NY, USA, 2018. Association for Computing Machinery.

[27] Aratrika Ray-Dowling, Ahmed Anu Wahab, Daqing Hou, and Stephanie Schuckers. Multimodality mobile datasets for behavioral biometrics research: Data/toolset paper. In *Proceedings of the Thirteenth ACM Conference on Data and Application Security and Privacy*, CODASPY '23, page 73–78, New York, NY, USA, 2023. Association for Computing Machinery.

[28] Abdul Serwadda, Vir V. Phoha, and Zibo Wang. Which verifiers work?: A benchmark evaluation of touch-based authentication algorithms. In *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–8, 2013.

[29] Balabit Mouse Challenge Data Set. How to: Use sentiment analysis and opinion mining. https://github.com/balabit/Mouse-Dynamics-Challenge, 2016.

[30] Zdeňka Sitová, Jaroslav Šeděnka, Qing Yang, Ge Peng, Gang Zhou, Paolo Gasti, and Kiran S. Balagani. Hmog: New behavioral biometric features for continuous authentication of smartphone users. *IEEE Transactions on Information Forensics and Security*, 11(5):877–892, 2016.

[31] Yan Sun, Hayreddin Ceker, and Shambhu Upadhyaya. Shared keystroke dataset for continuous authentication. In *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6, 2016.

[32] Syed Zulkarnain Syed Idrus, Estelle Cherrier, Christophe Rosenberger, and Patrick Bours. Soft biometrics database: A benchmark for keystroke dynamics biometric systems. In *2013 International Conference of the BIOSIG Special Interest Group (BIOSIG)*, pages 1–8, 2013.

[33] Syed Zulkarnain Syed Idrus, Estelle Cherrier, Christophe Rosenberger, and Patrick Bours. Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords. *Computers Security*, 45:147–155, 2014.

[34] Ahmed Anu Wahab, Daqing Hou, Mahesh Banavar, Stephanie Schuckers, Kenneth Eaton, Jacob Baldwin, and Robert Wright. Shared multi-keyboard and bilingual datasets to support keystroke dynamics research. In *Proceedings of the Twelfth ACM Conference on Data and Application Security and Privacy*, CODASPY '22, page 236–241, New York, NY, USA, 2022. Association for Computing Machinery.

[35] Lu Xiaofeng, Zhang Shengfei, and Yi Shengwei. Continuous authentication by free-text keystroke based on cnn plus rnn. *Procedia Computer Science*, 147:314–318, 2019. 2018 International Conference on Identification, Information and Knowledge in the Internet of Things.