Confounding Factors in Keystroke Dynamics

Osvaldo Andrés Pérez-García(✉)

Advanced Technologies Application Center (CENATAV),
7ma A #21406 e/214 y 216, Rpto. Siboney, Playa, CP 12200 La Habana, Cuba
osvaldo.perez@cenatav.co.cu

Abstract. Authentication is the verification of the identity of a person to access a resource or perform an activity. Authentication based on keystroke dynamics biometrics validates a legitimate user, comparing his typing on keyboard with his stored template. An important group of factors influences the capture of the raw data generated by the user's typing. These Confounding Factors have been addressed in the literature from different approaches, and most of these studies agree that their influence affects the reliability of Keystroke Dynamics. In this research, a taxonomy of Confounding Factors is proposed, and several mitigation actions are discussed to face them.

Keywords: Keystroke dynamic · Authentication · Behavioral biometric · Confounding factors

1 Introduction

Authentication in computer security is the process of linking a user with their identity. The basis of authentication process is the knowledge that the system must have about authorized users, and the validation of this information. This validation is performed based on (i) What secret a user knows (e.g., password), (ii) What a user has (e.g., a token); (iii) Who is he/she (e.g., biometric traits), or (iv) Where he/she is (e.g., connected from a particular network)[1].

In 1975, Spillane [2] suggested the user authentication based on the user's typing behavior and Gaines, et al. in [3], were the first to study the possibilities of Keystroke Dynamics (KD) for this purpose. After that, KD has been the focus of multiple research works, some of them dedicated to the negative influence of several factors in experiment results and practical systems [4-6].

The present paper discusses KD and its relationship with those negative factors. A taxonomy to facilitate their study is proposed. The existence and poor understanding of these factors, their impact on KD Authentication Systems (KDAS) reliability, and the possibility of their exploitation by impostors, for example, to generate synthetic forgery attacks, were the main motivations to do this work.

According to this, and to accomplish our aim, it is necessary to know the bases of KD briefly. Thus, this paper was organized as follow: KD systems are described in the next section 2. In section 3, the proposed taxonomy, and an explication about abovementioned factors are discussed. At the same time, mitigation actions that may be applied in experimental and real scenarios are proposed. Finally, the conclusions of the study are presented.

A. Pardo and J. Kittler (Eds.): CIARP 2015, LNCS 9423, pp. 192–199, 2015.
DOI: 10.1007/978-3-319-25751-8_24

2 Keystroke Dynamics in Brief

KDAS compare user's templates with keystroke samples. In function of that, the user may type fixed text when a predefined text is required or free text when restrictions about content or length of text are not necessary. KDAS are used in a static manner, for example, at the beginning of a session, or in a dynamic (continuous) way while the user's session lasts.

Five recommended components to be used in KDAS are described in [7]. That is, (a) Data acquisition: Original data are captured and processed. (b) Features extraction: Raw data acquired are processed and users' profiles are conformed, which contain extracted user's traits. (c) Classification and matching: Selected criteria are applied to categorize data. (d) Decision: User's data are compared using selected algorithms (classification or matching). Finally, (e) Adaptation: Although an enrolled user have been correctly recognized by the system, is recommended that the user re-enroll to add new information and update stored user profile.

According to the literature, the features distinctive of the user's typing behavior, but not only, are: (1) the pressure that is exerted on each key; (2) the position adopted by hands when user types; (3) functional relationship between fingers and keys; (4) the sound generated by keystroke; (5) the vibration generated by keystroke; (6) the sequence used to perform an action; (7) quantity of errors committed when writing and methods to correct them; (8) the use of special keys; (9) keystroke speed (total typing); (10) time interval that remains on a pressed a key, and (11) time interval between pressing a key and then another.

The first five features require special devices to be gathered (cameras, pressure sensor, motion sensor or microphone). The rest only needs the keyboard. Most public research is focused on the use of time intervals, viz. 9, 10 and 11. Of these, the two latest features have attracted more interest. For example, in [7] was reported that the 90 percent of the studied works used features based on timing, and only 5 percent of them were focused on pressure. According to literature, all mentioned characteristics are present while users are typing (universality), which helps distinguish one user from another (uniqueness), and they permit to recognize a person for long periods of time (permanence) [1].

Whereas the user types, a sequence of two consecutive events take place: a) Key-Down, when he/she presses the key, and b) Key-Up, when the same key is released. While a keystroke occurs, the KDAS must capture the key's identification (or key's name), Key-Down's timestamp and Key-Up's timestamp. Knowing these data, it is possible to calculate how long it takes the user to press two keys (digraph), three keys (three graphs), or several consecutive keys (n-graph), and a posteriori, it is possible to extract features, and obtain the user's templates.

2.1 Keystroke Dynamics Classification

Classification methods for KD have been treated in the literature through different approximations, using classic statistical methods, machine learning, with emphasizing in neural networks, pattern recognition techniques, and hybrid approaches. In [1, 7-10]

surveys of these methods are presented. From analysis of these studies, two issues stick out, the variability in the design of experiments and the insufficient amount of public data sets, making it too complex the execution of performance analysis and comparison of algorithms [5]. This complexity increases due to the influence of a group of confounding factors (CF), which are ignored by most of researches when designing experiments, when creating the users' templates database, and worse, when a KDAS is applied in real scenarios. Section 3 is dedicated to explaining CF.

2.2 Evaluation Metrics

To evaluate the performance of a KDAS in experimental scenarios [1], four metrics are applied. (i) False *Rejection Rate* (FRR): percentage of genuine users rejected. (ii) *False Acceptance Rate* (FAR): percentage of impostors accepted. (iii) *Equal Error Rate* (EER): Point on the ROC curve where FRR is equal to FAR, and (iv) *Half-Total Error Rate (HTER),* which is defined as the mean of both error rates FAR and FRR[11], that is, HTER = (FAR + FRR) / 2.

For systems with very high security needs, it is desired to reduce the value of FAR to the minimum possible [11, 12]. In contrast, if the system needs standard security, but emphasizing in user's acceptation, similar and small values of FAR and FRR are preferred. In addition, selecting the adequate performance metric depend on how you design the authentication system. For example, if the intention is to develop a continuous authentication system, it is necessary to collect continuous or periodically the user's keystrokes, and, on this basis, both ANIA (Average Number of Impostor Actions) as ANGA (Average Number of Genuine Actions) could be most appropriated to evaluate the system performance than the aforementioned metrics [13].

3 Confounding Factors (CF), Background and Related Works

The CF in KDAS are internal or external elements that affect in some way the value of data captured at system's stages, causing error rates unexpected variability, and misinterpretation of outcomes.

Problems associated with the inability to repeat experiments, or variability in the results of applying the same classifier to different data sets, are described in the literature as probable effects of CF. For example, Maxion, in an excellent analysis [14], which explains different CF that according to his criteria had disturbed KD experiments conducted up to that time. Maxion's article has two antecedents. First, [15], where the dependence of the KDAS of the internal clock of the computers, and the participation of the operating systems in this dependence are demonstrated. And second, [4], where authors measured the effects of 5 factors over the KDAS's data capture mechanism, and demonstrated how one of them, the Operating System load (I\O load specifically), had affected that process. In that paper, the analyzed factors were included in Environmental class. In contrast, in our work, the same factors are classified into the Technical factors class because, in our opinion, this class does not influence the way of the user's typing, as Environmental factors do. See Figure 1.

Lee et al. [16], referred that keystroke-based authentication systems show much higher error rates than others behavioral biometric systems because of the influence of several factors, between them, types of keyboard, physical status of user, etc. On the other hand, in [17], authors identified six factors that, in their view, "might influence the error rates of keystroke-dynamics detectors". These are, the algorithm, training amount, feature set, updating, impostor practice and typist-too-typist variation; some of them included in Table 1 and grouped as part of the Experimental and System Design class.

3.1 Confounding Factors Taxonomy

According outcomes from review of current papers about KD, what stands out is the fact that the authors correctly identify most of the Confounding Factors, but assuming that those factors are inherent to the environment where the experiment has been applied or the behavior and characteristics of users. Without going to judge the validity of this approach, the premise in this paper is that by mixing the origin of the CF, it is more complicated to determine how to minimize its effects on experimental and real scenarios. Then, in the light of our knowledge, and according to the reviewed literature, there are at least four sources of CF (see Table 1):

1. **Technical Type:** For KDAS, it is very important the precision of raw data captured. As it was aforementioned in section 2, the majority of research works about KDAS have focused on features based on time intervals. The captures of these periods between keystrokes depend on the computer's hardware, Operating System, keyboard and application used in keystrokes acquisition. In [15] and [18], authors reported the incidence of these factors in their measurements, which showed significant variations of error rates, while Villani, et al. in [19] demonstrated the importance of using the same keyboard in experiments

   In order to identify in practice this kind of factors (see Figure 1), it is necessary to understand that they do not affect the way of the user types, but the data value captured by the sensor, which can be considered as the union of the keyboard, the computer hardware, the Operating System and the final application. Knowing how each factor affects the samples, it allows us, from the design phase, to consider how to mitigate their influence, for example, by removing the affected data samples, and include automated controls to detect inappropriate sensors.

   Another important characteristic of Technical Factors is their influence over all users involved in tests. That is, no matter who is the user, his keystroke timing measures will be affected by the combination of the computer's hardware, Operating System, keyboard and the application used in keystrokes acquisition.
2. **Environmental Type:** This class includes every elements belonging to the location where the experiment takes place or where the final application will run, and that they may affect the way the user types: the illumination, keyboard position in relation to the user, room temperature, possibility of interruptions (e.g., colleagues, phone calls), and others. Some authors refer to these CF, but no revised article demonstrates the existence of a relation between them and the error rate variability[20]. On the above, it is logical and obvious to think that the environment influences on the mood of the user, and this, in turn, affects his way of typing. See Figure 1.

Table 1. Confounding factors taxonomy

Confounding Factor				Description
Type	**Factor**			
Technical	Operating System (OS) [11, 15]			Data from keyboard events are recorded running system-calls.
	Keyboard [11, 19]			Keyboard type: shape, language, keys' positions (ZERTY, QWERTY…), technical characteristics (interface), defective device.
	OS load			IO load, multithread, etc.
	RAM and µprocessor [18]			Amount of RAM, type of µprocessor
	Level of data capture			For instance: driver or application level
	Computer			E.g., Laptop or Desktop
Environment	Lighting			Low light
	Keyboard position			Position of keyboard to user
	Interruptions			Colleges, phones, etc.
Behavioral [21]	Transients	Physical tiredness, emotions, stress, drowsiness		Impact on the way of user's typing
	Permanents	User experience		
		User age, soft biometric [22]		
Experimental and system Design	**Design**	Text	Type	Fix (structured) or Free (unstructured) text
			Length	How many characters
		Monitored	Continuous	While user types
			Periodic	Time windows
		Task		Copy from dictated or reading
	Database	Type		Public or private
		Size		How many samples per user?
	Detectors	Classifier and metrics		Statistical, Machine Learning…
		Feature Set		Dimension
	Metrics			Performance metrics to continuous or static authentication
	Subjects (users & impostors)	Subjects number		Class number
		Number of attempts and time between them.		Impact on user experience
		Training amount		More training, less error rate

3. **Behavioral Type:** This kind of factor includes mood, fatigue, and others. In [23] these factors are named Transient factors, but the user experience is not a temporal element, neither the user age, and both affect the way the user types. In [24], the relation between emotional states and the variations in the rhythm of fingering if proved, and in [21], was demonstrated how the user's emotional states affect for short time his/her typing rhythm.
4. **Experimental Type:** These factors are discussed in most of the reviewed papers, but very few of them deal with this topic in the necessary depth, despite their importance. All experiments designed and carried out with the required quality must ensure repeatability, reproducibility and validity. Under those principles, Maxion in [14] proposed to take into account a set of environmental factors (including into this denomination our Technical class) and behavioral factors in the design of experiments, due to the demonstrated influence of these factors on experimental outcomes. Before that, Killourhy and Maxion [5] showed the drawbacks to trying to compare the experimental results published until that moment, precisely due to the influence of several factors described in our taxonomy inside the current class.

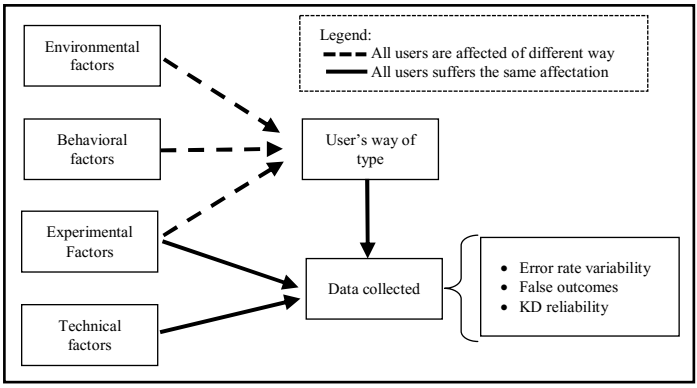

Fig. 1. Influence of Confounding Factors over users and data collection process

3.2 About Mitigations Actions

In general, the correct design of the experiments ensuring their repetition, reproduction and their validity should be the first step in any debate about how to mitigate the influence of CF in KDAS. And second, but not less important, it is essential to understand that CF are probable sources of error, and though their most significant impact is perceived in the assessment of the results, it is practically impossible to recognize their presence without knowledge about their existence.

On this basis, it is necessary to establish as the requirements for the execution of the experiments and developing of an application of KD-based authentication, the use of specific hardware and software (e.g., computer, keyboard, memory RAM, Operating System). In addition, it is essential to include automatic controls to ensure compliance with these requirements inclusive until level code.

For further analysis the following works are recommended: [5], where the authors, Killourhy and Maxion, presented some of the ways to mitigate the influence of Experimental CF; and [25], where Deng and Yu Zhong proposed the application of the Identity Vector (i-vector) method. This is a new approach taken from the voice biometric domain, and they demonstrated how much it could help to mitigate CF actions. The own Killourhy, in [6], proposed a much more comprehensive approach to face CF.

To mitigate the influence of technical factors, it must first understood that they act like a monolithic sensor, affecting each other, while causing changes in the readings of the keyboard event times. Therefore, it can be understood to be the same influence for all users, which is true, but it is not all the time, because it depends on all the variables involved. For example, the characteristics of the microprocessor and RAM are constants whose influence varies according to the load of operating system tasks. And so on, a similar analysis can be performed to all components of such factors.

Then, it is possible to adjust the operating system timer (for example, to 1 millisecond) at the enrollment stage, or use the same trick to the continuous authentication while the user types. To implement both ideas, it is recommended to study carefully its negative effects, and take into account the differences between laptops and desktop computers[26].

4 Conclusions

Keystroke dynamics is a cost effective mechanism for security and authentications systems. Its development has been marked by the negative effect of several factors, which cause false experimental outcome and error rate variability, wounding the KD reliability.

In this work, a taxonomy of Confounding Factors is proposed. This classification was obtained from the analysis of current and public papers, which are focused on KDAS particularities, and their relation with CF. In addition, this work aims to facilitate the study of CF sources, and to contribute with other researchers to improve the design of experiments and final applications for KDAS.

References

1. Zhong, Y., Deng, Y.: A survey on keystroke dynamics biometrics: approaches, advances, and evaluations. In: Zhong, Y., Deng, Y. (eds.) Recent Advances in User Authentication Using Keystroke Dynamics Biometrics. Science Gate Publishing (2015)
2. Spillane, R.: Keyboard apparatus for personal identification. IBM Technical Disclosure Bulletin **17**, 3346 (1975)
3. Gaines, R.S., Lisowski, W., Press, S.J., Shapiro, N.: Authentication by keystroke timing: some preliminary results. DTIC Document (1980)
4. Killourhy, K.S.: The role of environmental factors in keystroke dynamics. In: IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2009) Supplemental Volume (Student Forum), pp. 125–134 (2009)
5. Killourhy, K.S., Maxion, R.A.: Comparing anomaly-detection algorithms for keystroke dynamics. In: IEEE/IFIP International Conference on Dependable Systems and Networks. DSN 2009, pp. 125–134 (2009)
6. Killourhy, K.S.: A scientific understanding of keystroke dynamics. DTIC Document (2012)
7. Teh, P.S., Teoh, A.B.J., Yue, S.: A survey of keystroke dynamics biometrics. The Scientific World Journal **2013** (2013)
8. Karnan, M., Akila, M., Krishnaraj, N.: Biometric personal authentication using keystroke dynamics: A review. Applied Soft Computing **11**, 1565–1573 (2011)
9. Bhatt, S., Santhanam, T.: Keystroke dynamics for biometric authentication—A survey. In: 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), pp. 17–23. IEEE (2013)
10. Banerjee, S.P., Woodard, D.L.: Biometric authentication and identification using keystroke dynamics: A survey. Journal of Pattern Recognition Research **7**, 116–139 (2012)
11. Giot, R., El-Abed, M., Rosenberger, C.: Keystroke Dynamics Authentication. Biometrics chapitre 8 (2011)
12. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology **14**, 4–20 (2004)
13. Bours, P., Mondal, S.: Continuous Authentication with Keystroke Dynamics (2015)
14. Maxion, R.: Making experiments dependable. In: Jones, C.B., Lloyd, J.L. (eds.) Dependable and Historic Computing. LNCS, vol. 6875, pp. 344–357. Springer, Heidelberg (2011)

15. Killourhy, K.S., Maxion, R.A.: The effect of clock resolution on keystroke dynamics. In: Lippmann, R., Kirda, E., Trachtenberg, A. (eds.) RAID 2008. LNCS, vol. 5230, pp. 331–350. Springer, Heidelberg (2008)
16. Lee, J.-W., Choi, S.-S., Moon, B.-R.: An evolutionary keystroke authentication based on ellipsoidal hypothesis space. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation, pp. 2090–2097. ACM (2007)
17. Killourhy, K., Maxion, R.: Why did my detector do *That*?! In: Jha, S., Sommer, R., Kreibich, C. (eds.) RAID 2010. LNCS, vol. 6307, pp. 256–276. Springer, Heidelberg (2010)
18. Bello, L., Bertacchini, M., Benitez, C., Pizzoni, J.C., Cipriano, M.: Collection and publication of a fixed text keystroke dynamics dataset. In: XVI Congreso Argentino de Ciencias de la Computación (2010)
19. Villani, M., Tappert, C., Ngo, G., Simone, J., Fort, H.S., Cha, S.-H.: Keystroke biometric recognition studies on long-text input under ideal and application-oriented conditions. In: 2006 Computer Vision and Pattern Recognition Workshop, CVPRW 2006, p. 39. IEEE (2006)
20. Roy, S., Roy, U., Sinha, D.: Text-based Analysis of Keystroke Dynamics in User Authentication. International Journal of Computer Sciences and Engineering **3** (2015)
21. Lee, P.-M., Chen, L.-Y., Tsui, W.-H., Hsiao, T.-C.: Will user authentication using keystroke dynamics biometrics be interfered by emotions? In: Zhong, Y., Deng, Y. (eds.) Recent Advances in User Authentication Using Keystroke Dynamics Biometrics, GCSR vol. 2, pp. 71–81. Science Gate Publishing (2015)
22. Syed Idrus, S.Z., Cherrier, E., Rosenberger, C., Mondal, S., Bours, P.: Keystroke dynamics performance enhancement with soft biometrics. In: IEEE International Conference on Identity, Security and Behavior Analysis (ISBA), Hong Kong, China (2015)
23. Zhong, Y., Deng, Y., Jain, A.K.: Keystroke dynamics for user authentication. In: 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 117–123. IEEE (2012)
24. Lee, P.-M., Tsui, W.-H., Hsiao, T.-C.: The influence of emotion on keyboard typing: an experimental study using visual stimuli. Biomedical Engineering Online **13**, 81 (2014)
25. Deng, Y., Zhong, Y.: Keystroke dynamics user authentication using advanced machine learning methods. In: Zhong, Y., Deng, Y. (eds.) Recent Advances in User Authentication Using Keystroke Dynamics Biometrics, GCSR vol. 2, pp. 23–40. Science Gate Publishing (2015)
26. Microsoft: Timers, Timer Resolution, and Development of Efficient Code (2010). http://www.microsoft.com/whdc/system/pnppwr/powermgmt/Timer-Resolution.msp