A LOOK AT DIGITAL TRANSFORMATION: CAN SOCIETY BE PERSUADED TO A NEW AUTHENTICATION METHOD – BEHAVIORAL BIOMETRICS?

Emanuel Krzysztoń

Faculty of Mechatronics Kazimierz Wielki University (POLAND)

Abstract

The 21st century brings many technological novelties and modifications to existing solutions. The digital transformation that we undergo in every aspect of life often affects us against our will, with varying results. Thousands of years ago, when a human rubbed two stones together, he caused a spark that started a chemical phenomenon, but above all - a revolution. Today, lighting a bonfire does not arouse great emotions, we are faced with many other technological novelties aimed at making our lives easier. In the past, when human didn't need the fire, he had a choice. The present shows that we are becoming strongly dependent on technology and our choice becomes very limited or impossible. As a result, we agree to the terms offered. It is worth to investigate whether we take a bold decision overnight? What guides us in making such key decisions in our lives?

The main goal of this paper is to take a look at safeguards used to protect our goods, in information technology. Until now, the security measures used to authenticate the user were in the form of i.a. login and passwords, two-factor authentication (2FA), universal 2nd Factor (U2F), biometrics. Each of these solutions has its advantages and disadvantages. It would seem that these are the best solutions, and using them together would provide a multi-layer security. However, does it not become a significant obstacle in everyday use? Then behavioral biometrics comes to the rescue, which aims to learn our behaviour using machine learning and to clearly confirm in real time whether the person using the service or device is the owner. Using this method, many doubts and concerns arise about the technology itself. Based on the available research, literature and other sources, I will look for answers to the question „Can the public be persuaded to use this method of authentication?". Justifying the reply, I will analyse the available tools and refer to the document of General Data Protection Regulation.

Keywords: Authentication; behavioral biometric; cybersecurity; machine learning; keystroke.

INTRODUCTION

Continuous technological development is causing a modification of existing solutions. With these changes, it is necessary to introduce new methods that can significantly improve the functionality and security of the system. Making changes at a very fast pace, arouses incomprehension and even fear in public opinion, which creators can't calm down. In the paper, I will present an authentication using behavioral biometrics, mainly focusing on keystroke method. User authentication is the process of verifying the declared identity (Padmavathi & Shanmugapriya, 2009). So far, the following division has been adopted for user verification in the system: what you know (personal identification numbers (PIN), passwords, pattern locks, graphical passwords); what you have (security keys, magnetic/chip cards); what you are (biometrics incl. facial features, retina, fingerprint) and what you do (behavioral biometrics incl. keystroke, mouse motion, touch motion). Use of a single-factor authentication (SFA) in order to access the system is not good practice. Therefore, nowadays it is standard to use at least two-factor authentication (2FA), when performing key services i.e. bank transfers, modification

of entitlements and such like. In theory, the use of multi-layer authentication (MFA) in which the system defines as a minimum three layers of verification should provide sufficient security. However, each of these methods has its advantages and disadvantages, and the authentication methods are divided into more and less user-friendly for a user. Despite these multi-stage verification methods, the system is still vulnerable because entry-point authentication can't guarantee continuous verification. Therefore, ensuring continuous authentication is necessary in order to create a more secure system. On the other hand, in order to achieve continuity of authentication, it will be necessary to create a model that interprets all the anomalies, in the result access to the system may be not possible on the basis of these deviations.

Using behavioral biometrics - keystroke, will we need additional devices for authentication? What information will be needed to create a unique user profile? Will user privacy be compromised? Behavioral biometrics raises a number of questions and doubts. Are these fears really justified? Can society be persuaded to use this authentication method? The main purpose of this paper is to review selected research studies that will be analyzed. I will explain and give you the potential application of this method, the advantages, problems and challenges for this technology.

The paper is structured as follows: Section 2 is a general outline of behavioral biometrics. Section 3 provides a simplified outline of the development of the biometric model. Section 4 presents the technical content of keystroke behavioral biometrics. Section 5 is a review and analysis of keystroke on the basis of the available publications with regard to disadvantages, advantages and new challenges. Section 6 focuses on the legal aspect of biometrics. Section 7 is a summary, taking into account the current problems and challenges for this authentication.

BEHAVIORAL BIOMETRICS

Behavioral biometrics is a set of characteristic features and behaviors that define a given user, which can be used to authenticate the user. In this method, a unique profile is created, which becomes a reference model for subsequent verification and identification attempts, on the basis of rules defined by the system. Behavioral biometrics are already being implemented in online banking, e-commerce and markets with high levels of authentication (https://www.ibia.org/download/datasets/3839/Behavioral). Traditionally, two types of biometric data are distinguished: physiological and behavioral (Ali et al., 2016; Moskovitch et al., 2009).

Table 1. *Division of biometrics.*

BIOMETRICS	
PHYSIOLOGICAL	BEHAVIORAL
FACE	HANDWRITING
SHAPE OF THE EAR	SIGNATURE
ODOR	VOICE
RETINA	MOVEMENT OF THE LIPS
IRIS	GAIT
HANDPRINT	KEYSTROKE

FINGERPRINT	MOUSE DYNAMICS
DNA	TOUCH DYNAMICS
ECG	BEHAVIORAL PROFILE

Behavioral biometry is closely related to user behaviour, which is divided into two categories due to data collection: active and passive.

Table 2. *Categories of behavioral biometrics.*

BEHAVIORAL BIOMETRICS BREAKDOWN BY DATA COLLECTION METHOD	
ACTIVE (WITH A MACHINE)	PASSIVE (WITHOUT A MACHINE)
KEYSTROKE	HANDWRITING
MOUSE DYNAMICS	SIGNATURE
TOUCH DYNAMICS	VOICE
BEHAVIORAL PROFILE	GAIT

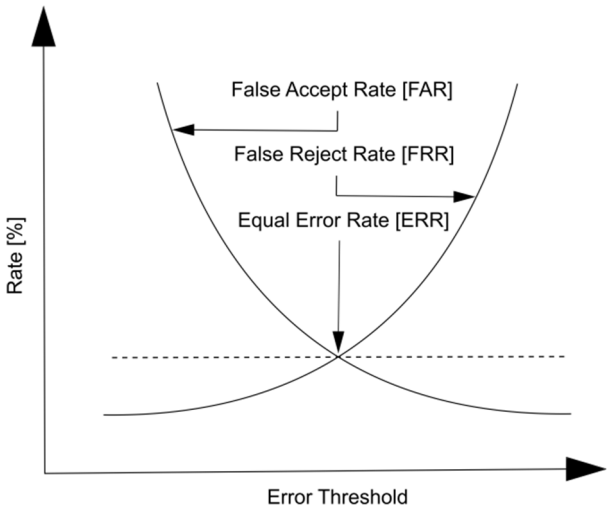
The active method requires interaction with the system, while the passive method can be recorded without interaction with the system (Monaco, 2014). Based on the collected data with the selected method, a user profile is created, which will serve as a reference model during verification and / or identification. The main objective of behavioral biometrics is to focus on how the user performs specific activities. To put it simply, according to the standard adopted by the algorithm, for example when the user is entering data to log into the system, there are significant anomalies in which the system benchmarks suggest that they are dealing with a fraud. In such a case, it is noticed by the system, because the way the user entered the data deviates from the predefined model. As a result of these events, the system is blocked.

HOW CREATED IS A MODEL OF BIOMETRICS?

The process of creating a biometric model is a multi-stage cycle. The first stage is data acquisition in order to collect them, the amount of data needed depends on the method used. After obtaining the necessary samples from the user, they will be included in the dataset, to the assigned pattern. However, the mere acquisition of data may not mean much in the case of a method in which frequent changes occur and they may determine the recognition of the right pattern. Therefore, static and dynamic functions shall be used during data acquisition (Revett, 2008, p. 12). The final point of the first step is to generate a biometric information record (BIR) for each user of the system (Revett, 2008, p.13). The next step is to transform the collected data into a useful set of models that provide the best accuracy. Thanks to this, it will be possible to identify, i.e. the process during which the biometric pattern will be attempted to determine the correct identity based on specific biometric features for a given user (Padmavathi & Shanmugapriya, 2009). At this point in our dataset there are many models and even more data, so an important step is to design an automated deciding mechanism. It is a key role in deciding whether to reject or accept a user trying to authenticate. Most often, this mechanism is based on selected features that are compared with previously obtained data stored in BIR. The

changes taking place in the samples stored in the BIR are very important. In book *Behavioral Biometrics: a remote access approach*, K. Revett referred to the ongoing changes in the authentication system taking place in keystroke. Due to various circumstances (e.g. faster reaction in the morning), this model may transform, therefore the real user may be falsely rejected, and therefore the created pattern cannot be implemented permanently. In such cases, a reference model is needed that collects a defined number of successful verifications and refers to them each time. When the number of collected data has reached its maximum value, it is updated starting with the removal of the oldest record in BIR. In order to be able to confirm the user's identity for fear of impersonating a real person, indicators that assess the performance of behavioral biometrics are necessary to evaluate this event. The following indicators have been taken over: false rejection rate (FRR) rejection by the system of a real user; false acceptance rate (FAR) the fraudster is accepted by the system; equal error rate (EER) this is the value in which the FRR and FAR indices are used, the intersection of both indices will determine the ERR value, if the ERR value is close to 'o', in such a case this indicates high performance and safety; true positive rate (TPR) these data are interpreted as a percentage of correctly identifying the owner; false positive rate (FPR) indicates the percentage of false positive results; receiver operating characteristic curve (ROC) an indicator showing performance in terms of TPR and FPR in the graph. More typically used indicators are presented in the paper by A. Alzubaidi and J. Kalita (2016), titled *Authentication of Smartphone Users Using Behavioral Biometrics*.

Figure 1. *Relationship between FAR, FRR AND ERR.*



BEHAVIORAL BIOMETRICS - KEYSTROKE

One of the methods of behavioral biometrics is keystroke. The history of this method dates back to World War II, when on telegraphic machines the operators of the Morse alphabet transmitted a message, then they were identified on the basis of the rhythm, pace and timing of the tap (Banerjee & Woodard, 2012). Since those years, technology has changed. Today, the main purpose of this authentication method is to protect the real user from impersonators. In this verification method, no additional devices are needed, moreover, the implementation of this method may provide the possibility of phased and / or continuous authentication to the system. Keystroke is based on the collection and processing of information obtained from the interaction between the user and the keyboard in order to automate the authentication and / or user identification (Revett, 2008, p.73). Each user uses the keyboard in a specific way, and the authentication process itself is done by observing a change in the user's typing pattern. The unique keystroke profile can be constructed from various writing functions such as writing

speed, time between keystrokes, pressure applied to the keys and finger positions on the keys (Ali et al., 2016). Most often, the authentication model using the keystroke method consists of six elements: data acquisition and collection, feature extraction, classification/matching algorithm, decision making, retraining and evaluation.

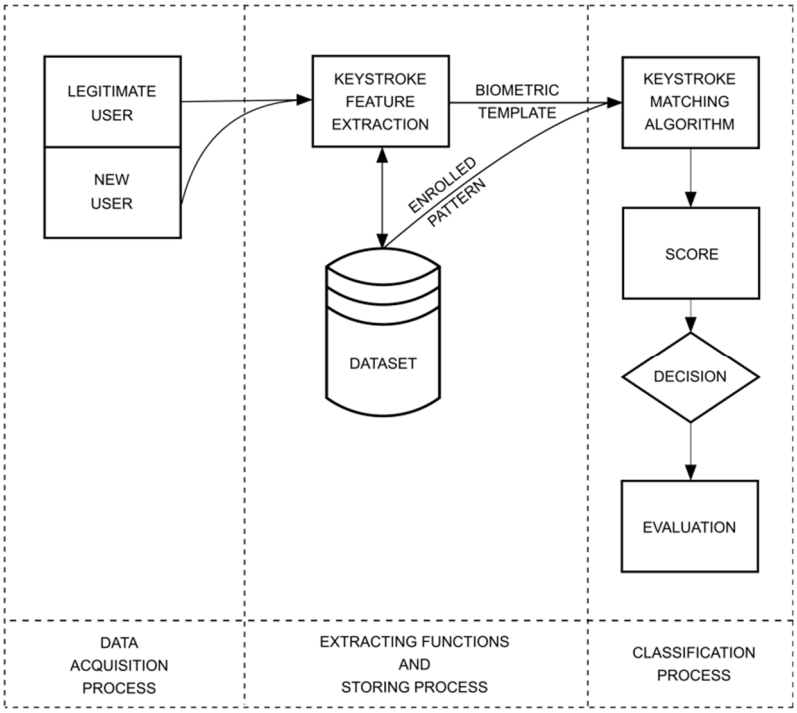


Figure 2. *Sample authentication model for keystroke method based on (Amin et al., 2014).*

Analysing previous studies contained in the paper (Padmavathi & Shanmugapriya, 2009) keystroke can be grouped into several data acquisition techniques and typing metrics on which keystroke system analysis is based: static at login; periodic dynamic; continuous dynamic; keyword-specific; application-specific. However, the keystroke system usually falls into two categories (Acien et al., 2020), fixed-text and free-text. In the case of fixed text, the user is prompted to enter a predetermined number of characters that were entered during the testing / learning phase. Free-text is characterized by the fact that you can enter any number of characters, then in both cases the system compares previously entered samples. These techniques are based on specific functions that are derived from the writing rhythm. Among the functions extracted from the user's writing rhythm depending on the technique used, the most common are time-measuring functions incl.: hold time, inter time, inter key, release latency, distance, speed (Alzubaidi & Kalita, 2016; Acien et al., 2020). In order to use the abovementioned functions, the following techniques shall be used to measure their performance for keystroke: statistical algorithms, artificial neural networks, fuzzy logic, support vector machine and others (Ahmad et al., 2013; Mohd Noorulfakhri Yaacob et al., 2020).

KEYSTROKE ADVANTAGES, DISADVANTAGES AND CHALLENGES

By analysing various scientific papers in terms of the strengths and weaknesses of the keystroke method. I can conclude that user authentication using this behavioral biometric method, regardless of the data acquisition technique, can be implemented in several ways: Server deployment in that case, only system host collects data, then this information is sent to the server where the rest of the process takes place. Hybrid deployment in this case, only writing rhythm functions are extracted from the client's. Client deployment here all processes are performed on the client's architecture.

The mentioned examples of keystroke implementation have some limitations. In work (Moskovitch et al., 2009), the authors stressed that, in the case of server deployment, in addition to the continuous load on the network, data security is also questionable. In addition, the researchers highlight the following problems: creation of an appropriate database template; use of different devices and different types of keyboards; the changing behaviour of the user, at different intervals when we are full of power, and when we have no energy; privacy and data security; scalability of this method. Moreover, the authors refer to the studies carried out and small data sets, as they may not reflect the real application.

To the advantages of behavioral biometrics keystroke researchers (Padmavathi & Shanmugapriya, 2009) they pass: running in the background; inexpensive implementation conditions, as the only hardware requirement in addition to the keyboard is the computer. Among the challenges for this method, researchers note: irregularity and inconsistency in keystroke operations; using different keyboards and keyboard layouts (i.e. Qwerty or dvorak); user's position while typing (sitting, lying and standing).

In paper (Banerjee & Woodard, 2012) the authors mention among the problems of behavioral biometrics keystroke: possible problems with privacy; the type of keyboard used; languages other than English aren't tested; no standardization of keystroke evaluation protocols; lack of database standardization. Among the main advantages of the use of behavioral biometrics keystroke the authors mentioned: low cost and define it as user-friendly; use in cybersecurity; information can be collected in keystroke without the user's knowledge. Among the challenges to behavioral biometrics keystroke researchers highlight: failure to conduct research in uncontrolled environments where conditions reflect reality; no categorization of the algorithms used to classify users; system resistance to time inaccuracies; determining the variability of the user's emotional state; improving classifiers, creating and adapting patterns to specific groups of people; time needed to train and execute algorithms for the pattern; size and duration of data storage.

Among the further advantages of the keystroke method, researchers (Ali et al., 2016) they exchange: low implementation cost compared to other authentication systems using biometrics; software implementation; little dependence on computer hardware; customization to the user; user-friendliness; the uniqueness of keystroke patterns; ensuring continuous monitoring. As a major drawback, the authors point to low accuracy, compared to other verification systems using biometrics. As a challenge in authenticating using this behavioral biometric method, they define achieving greater accuracy.

Authors (Alzubaidi & Kalita, 2016) identify the following challenges for behavioral biometrics: improving the accuracy of research; conducting research in conditions reflecting reality, as laboratory tests don't take into account many factors; focus on resource consumption, i.e. processor, memory and battery consumption, as most studies do not include this data.

In work (Patel et al., 2019) they drew attention to the excessive learning process and the time needed for the data processing in statistical and machine learning algorithms. In order to create a more robust and secure behavioral biometric system, researchers suggest complementing the keystroke behavioral biometric method with another method, such as mouse movement. In addition, they also pay attention to the users ability to copy and paste information and the use of password managers. Such activities may limit the effectiveness of the indicators.

By analyzing papers (Padamavathi & Shanmugapriya, 2009; Banerjee & Woodard, 2012; Ali et al., 2016; Mohd Noorulfakhri Yaacob et al., 2020), in which a review of studies focusing on different performance measurement techniques for keystroke was carried out. It can be seen that out of the studies selected by the authors, only in one sample the number of participants exceeded 1000, in most studies the number of participants ranged from 3 to 315. Moreover, there is a large discrepancy in the results achieved for controlled and uncontrolled

environments and data collection techniques. Analysing the results of these studies, one can come to the bold conclusion that the larger the number of participants and the set of data, the more difficult it is to identify the user.

Recent research (Acien et al., 2020) carried out on a database of 100 000 users, based on the Siamese Recurrent Neural Network (RNN) is a breakthrough. By analyzing the obtained results, for the free-text keystroke technique, which ranges from 9.53% to 3.33% ERR (depending on the amount of user data enrolled). The obtained results emphasize the new face and potential of keystroke operation. Moreover, the authors pointed out that the increase in the number of test users does not significantly affect the performance of this method.

BEHAVIORAL BIOMETRICS, AND GDPR

Behavioral biometrics are increasingly being used in institutions with a high level of security. Is this trend being followed by legislation on personal data protection? According to Article 4 (14) of GDPR quote „(...) biometric data means personal data which result from specific technical processing relating to the physical, physiological or behavioral characteristics of an individual and which enable or confirm the unambiguous identification of that individual, such as a facial image or dactyloscopic data (...)" (Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016). They should provide an adequate level of security for economic operators and individuals when using these authentication methods. On the other hand, they should establish minimum security standards for the storage and processing of biometric data. European Union legislation is aimed at strict protection of biometric data. Regulation of the European Parliament and of the Council (UE) 2016/679 of 27 April 2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data and repealing the Directive 95/46/WE (general data protection regulation) introduces in Article 9(1) a general prohibition to process, inter alia, biometric data in order to clearly identify an individual (Pyka, 2019). At the same time, this regulation has indicated a number of conditions under which this ban doesn't apply. It is worth paying particular attention to the fact that the ban in question does not apply, among other things, in the event that:

- explicit consent has been given for the processing of the data by the data subject;
- processing is necessary for the protection of the interests of the data subject and the performance of the administrator's duties.

According to the GDPR, this is sensitive data and should only be used in justified cases. Consent must be voluntary and mustn't be forced. The administrator must assess the appropriateness of the use of biometric data.

- In assessing it, it must answer the questions raised:
- Is biometrics necessary to achieve the objective?,
- Is there a less innovative solution?,
- Can a different solution be offered to the data subject?.

Behavioral biometrics is becoming an increasingly common element in improving system security. The specific features of this verification method make it suitable for use in the economic environment, such as banks and companies. It should also be stressed that, according to Article 9(4) of that Regulation, Member States may retain or introduce further conditions, including restrictions on the processing of, inter alia, biometric data. Therefore, the EU legislator has left quite a large scope of freedom in the processing of information collected for the purposes of this method. In addition, Member States may introduce additional provisions in personal data protection law governing the use of biometrics. The arrangements should provide the possibility to increase the level of security for economic entities, offices through the use of various types of biometrics, and establish minimum security standards for the storage and processing of information using these methods. These standards should protect data

against loss of confidentiality, availability, integrity and, in particular, against attacks on these systems.

CONCLUSION

The digital transformation in which we take part introduces new technologies whose actions we may not understand. In turn, with that, the danger of ignorance increases. The continuing trend of threats in cyberspace, which is aimed at crafting an attack and exploiting the naivety of human nature, is one of the themes of those involved in cyber security. In order to limit the occurring cyber threats, multi- layered security is used, including user authentication. Until now, entry-point verification methods have been used, the use of which doesn't guarantee the continuity of security in the system. In this paper, I have described a method of behavioral biometrics, which changes the existing face of user authentication methods, introducing real time verification mainly focusing on the keystroke method, i.e. active data acquisition using a machine. Although this method has been known for several decades, it has not found commercial use since its inception. Today, it is being successfully tested and implemented in many banks and markets that rely on a high level of system security. The main advantage of this method is low cost, or even lack of it, as only the keyboard and computer are required for proper operation. Furthermore, the user doesn't need to take time to implement and become familiar with this method, as it can work in the background without the user being aware of the processes that are taking place. Keystroke perfectly complements other authentication methods, protecting users from various attacks using social engineering. Technological progress and the experience gained from previous research by scientists in the field of keystroke behavioral biometrics abound in the development of this method. By analyzing the latest research, we can see a new face and potential in security for systems using this authentication method, in large-scale operation. The results of these works allow us to look with optimism about their improvement and development. Keystroke is not without flaws, and implementing this method is a difficult task. Therefore, analyzing the results of various studies, we can see large discrepancies due to, among other things, the choice of data collection techniques, the conditions under which the studies were carried out, the functions used to extract the data and the techniques measuring their performance. It is worth mentioning that Gartner's prediction for 2020-2022 indicates that the purchasing power of traditional PCs will be constantly decreasing, while mobile devices will be constantly at the top and they will set trends for the years to come. Therefore, all research should take into account user-side resource usage, i.e. CPU, memory, network / data transfer. In turn, excessive load can have a negative impact on the length of operation and, ultimately, even on the battery life of mobile devices. Furthermore, the issue of data security and privacy remains open. Most research focuses on creating and developing efficient systems using behavioral biometrics, but user safety is often overlooked. Data that are extracted by different functions must be adequately protected and leak-proof, for every way of implementing this authentication method. An important element of the protection of this information will be the construction of a common security policy that will be consistent with the applicable legal provisions. Based on the foundation of all actions taken in ICT systems, the so-called CIA triad, which includes properties such as confidentiality, integrity and availability.

To sum up, "Can society be persuaded to use this security method?" in the chemical phenomenon mentioned at the outset, which is fire. The primitive man probably didn't realise that he would have a whole range of new opportunities using this combustion process, thanks to which he will be able to create new tools and speed up everyday processes, making his life easier and also increasing his sense of security. However, as a result of many events, homo sapiens knew very well how dangerous fire is when he will lose control over it.

REFERENCES

Acien, A., Monaco, J. V., Morales, A., Vera-Rodriguez, R., & Fierrez, J. (2020). TypeNet: Scaling up Keystroke Biometrics. In *IAPR/IEEE International Joint Conference on Biometrics (IJCB)*. Retrieved from: https://arxiv.org/pdf/2004.03627.pdf

Ahmad, N., Szymkowiak, A., & Campbell, P. (2013). Keystroke dynamics in the pre-touchscreen era. *Medicine Frontiers in Human Neuroscience*. doi: 10.3389/fnhum.2013.00835

Ali, M. L., Monaco, J. V., & Tappert, C. C. et al. (2016). Keystroke Biometric Systems for User Authentication. *J Sign Process Syst., 86*, 175–190. doi: 10.1007/s11265-016-1114-9

Alzubaidi, A., & Kalita, J. (2016). Authentication of Smartphone Users Using Behavioral Biometrics. In *IEEE Communications Surveys & Tutorials*, vol. 18(3) (pp. 1998-2026). doi: 10.1109/COMST.2016.2537748

Amin, R., Gaber, T., ElTaweel, G., & Hassanien, A. E. (2014). Biometric and Traditional Mobile Authentication Techniques: Overviews and Open Issues. In A. Hassanien, T. H. Kim, J. Kacprzyk, & A. Awad (Eds.), *Bio-inspiring Cyber Security and Cloud Services: Trends and Innovations. Intelligent Systems Reference Library*, *Vol. 70* (pp. 423-446). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-662- 43616-5_16

Banerjee, S. P., & Woodard, D. (2012). Biometric Authentication and Identification Using Keystroke Dynamics: A Survey. *Journal of Pattern Recognition Research, JPRR, 7* (1). doi: 10.13176/11.427

Behavioral Biometrics, International Biometrics Identity Association IBIA. Retrieved from: https://www.ibia.org/download/datasets/3839/Behavioral

Mohd Noorulfakhri Yaacob, Syed Zulkarnain Syed Idrus, Wan Nor Ashiqin Wan Ali, Wan Azani Mustafa, Mohd Aminudin Jamlos, & Mohd Helmy Abd Wahab. (2020). Decision Making Process in Keystroke Dynamics*. Journal of Physics: Conference Series*, *Vol. 1529*, *The 2nd Joint International Conference on Emerging Computing Technology and Sports (JICETS)*. doi: 10.1088/1742-6596/1529/2/022087

Monaco, J. V. (2014). Classification and authentication of one-dimensional behavioral biometrics. In *IEEE International Joint Conference on Biometrics*, *(IJCB).* doi: 10.1109/BTAS.2014.6996253

Moskovitch, R. et al. (2009). Identity theft, computers and behavioral biometrics. In *IEEE International Conference on Intelligence and Security Informatics* (pp. 155-160). doi: 10.1109/ISI.2009.5137288

Padmavathi, G., & Shanmugapriya, D. (2009). A Survey of Biometric keystroke Dynamics: Approaches, Security and Challenges. *International Journal of Computer Science and Information Security, IJCSIS*, *5*(1), 115-119. Retrieved from: https://arxiv.org/abs/0910.0817

Patel, Y., Ouazzane, K., Vassilev, V. T., Faruqi, I., & Walker, G. L. (2019). Keystroke Dynamics using Auto Encoders. In *International Conference on Cyber Security and Protection of Digital Services (Cyber Security)* (pp. 1-8). Oxford, United Kingdom. doi: 10.1109/CyberSecPODS.2019.8885203

Pyka, A. (2019). Przetwarzanie danych osobowych do celów badań naukowych. Aspekty prawne. *Studia Prawa Publicznego*, *4* (28), 79-101. Retrieved from: https://doi.org/10.14746/spp.2019.4.28.4

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L 119, 4 May 2016 with later amendments. Retrieved from: https://eur-lex.europa.eu/eli/reg/2016/679/2016-05-04

Revett, K. (2008). *Behavioral Biometrics: A Remote Access Approach*. Chichester, United Kingdom: John Wiley & Sons, Ltd.