”ALGORITMI SI UN SET DE DATE CARE FOLOSESC DINAMICA TASTĂRII TASTELOR IN CAZUL TEXTULUI SCRIS LIBER PENTRU AUTENTIFICAREA CONTINUĂ IN PLATFORME EDUCATIONALE CARE CUPRIND CUSRURI ONLINE DESCHISE MASIVE (MOOC)” /

”FREE-TEXT KEYSTROKE DYNAMICS DATA SET AND ALGORITHMS FOR CONTINUOUS AUTHENTICATION IN EDUCATIONAL PLATFORMS WITH MASSIVE OPEN ONLINE COURSES (MOOC)”

Teză de doctorat – Rezumat

pentru obținerea titlului științific de doctor la

Universitatea Politehnica Timișoara

în domeniul de doctorat Calculatoare și Tehnologia Informației

autor ing. Augustin-Cătălin IAPĂ

conducător științific Prof.univ.emerit dr.ing. Vladimir-Ioan CREȚU

februarie 2021

Content:

FREE-TEXT KEYSTROKE DYNAMICS DATA SET AND ALGORITHMS FOR CONTINUOUS AUTHENTICATION IN EDUCATIONAL PLATFORMS WITH MASSIVE OPEN ONLINE COURSES (MOOC)

The paper focuses on the continuous authentication of a computer user based on keystroke dynamics, the way to type on the keyboard. During the research, an authentication algorithm based on keystroke dynamics was developed, a data set regarding the typing mode was collected from 80 volunteers, two modified metrics were proposed to obtain better performances of the authentication algorithm and a data structure was proposed to store the necessary information of the users.

This method of authentication justifies its attention especially in online educational platforms, platforms that experienced a very large increase in 2020, due to the relocation of most courses in the online environment, a restriction generated by the COVID-19 crisis.

1. INTRODUCTION

1.1 Thesis context

This thesis started from the need to develop additional ways to identify the identity of a user who uses a private account on a computer. This need is more pronounced in the case of courses or exams that take place online. The MOOC phenomenon (Massive Open Online Courses), courses attended by a large number of students from any corner of the world online, was born in 2008. This phenomenon reached a first maximum in 2012, and in 2020 there was an exponential increase in the number of students enrolled [IAP21b].

The year 2020 also led to radical changes in education systems as an outcome of the health crisis caused by the SARS-CoV-2 virus. This has resulted in an unprecedented push to online learning. Universities, primary schools or high schools have been pressed to adapt and move the entire classical education system from studying in the classroom, face to face, to distance platforms. In this context, it has become much more important to find methods to ensure that, for instance, during an exam, where both the teachers and students are in different locations, to ensure that the student, through easily accessible means, is the one who solves the subjects and receives a grade based on his knowledge and performance [IAP21b].
There are many ways and possibilities to identify and authenticate a user from an electronic account. The most common method is to retain a username and its password and based on these two, the user has access to the account. The use of physical cards, such as those used by banks, or fingerprints, retinal scanning or face recognition requires the existence of additional devices for retrieving data from users. For authentication during an exam, it is not enough to have an account and a password, in case the student wants to speculate by leaving someone else in his place to solve the subjects. In most cases, the camera and microphone must be turned on throughout the exam [IAP21b].

An effective method in solving the problem described above is continuous authentication using keystroke dynamics. Keystroke dynamics is the method by which a user can be identified or authenticated based on his or her particular way of typing text on the keyboard. This method does not require additional hardware, any computer or laptop that is equipped with a keyboard is accepted. Additionally, another advantage is represented by the fact that the identity verification can be done continuously, at any time when the user types on the keyboard. The password authentication cannot be done the same way presented before, being done usually only once when accessing the account, and along the way the user can change without the system to realize the change.

Another advantage of using identification or authentication using keystroke dynamics is that the user does not have to take additional steps. The participants just have to type and the system monitors the way of their typing. In this case, after an authentication in a system, if the user changes, the system will realize that someone else is at the computer and can signal this change.

Thousands of students can participate in MOOC courses at the same time. In the case of an exam with thousands of students, it becomes impossible to supervise through the video camera and the microphone, this method being effective when the number of students is reduced. In the case of keystroke dynamics, any number of students can be continuously authenticated, there is no such limitation in this regard.

The disadvantage of a system with authentication or identification of users through the keystroke dynamics method is the accuracy of the algorithm with which the user can be identified. Currently, systems that use this method do not reach error rates of 0%. They have performance that identifies the user with an error rate of less than 10%, or in some cases with

even higher accuracy, instead improving algorithms based on keystroke dynamics is still a challenge in this area. Along these, another challenge for scientific research in this field is the fact that in order to test the efficiency of the algorithms proposed in various researches, databases are needed that capture the typing mode, thus better simulating the real conditions. Within the scientific research made about the keystroke dynamics they were identified two different branches. The first would be when a user types a default text on the keyboard, such as a user, a standard password or phrase. The second one would be the typing of a free text on the keyboard without certain conditions being imposed [UMP85] [MES11]. The two methods are analyzed separately by different methods in the scientific literature on this subject. Both, however, involve a phase in which the system collects data about the user, the typing times, and the typing mode, thus, creating a profile of the user that he will use later in the continuous authentication phase. The first method has been more intensively explored and the results are more successful in this direction because it is the same text entered from the keyboard each time. The second method, when the user types a free text with the help of the keyboard, without conditions, has been researched especially in recent years, and the results are increasingly improved.

Only in the last 5 years over 10,000 scientific papers have been published about keystroke dynamics. Also, survey papers have been published as keystroke dynamics biometrics has drawn intense research interest the past couple of decades [ZHO15].

1.2 Thesis objectives

In this research project, the author set the following four objectives:

Objective 1, O1, The first objective of this thesis is to collect a database with the test pattern from at least 80 users, in order to test the authentication algorithm for this research, but also to make it available to other interested researchers.

Objective 2, O2, The second objective of this thesis is to implement an algorithm for authenticating the users of a computer based on the keystroke dynamics, the keyboard typing mode.

Objective 3, O3, The third objective of this thesis is to propose at least two new metrics for calculating the distances between two vectors that generate better performance compared to the Equal Error Rate (EER) performance indicator than the classical methods.

Objective 4, O4, The fourth objective of this thesis is to propose a data structure as efficient as possible, which should contain the most relevant information about the typing of a user.

1.3 Thesis structure

The thesis is organized as follows:

- Chapter 1 presents the thesis context, the thesis objectives and the thesis structure.
- Chapter 2 presents the state-of-the-art of the field to which this work is addressed.
- Chapter 3 presents the research methodology applied in this research project. The steps performed in the present scientific research are presented below: A. Development of the platform for the acquisition of input data, B. Acquisition and initial processing of input data from 80 volunteers (how typing on their keyboard), C. Processing the input data so as to generate a user pattern for each user, D. Development of an algorithm in the C programming language for calculating distances used in keystroke dynamics authentication, E. Simulation of system authentication by genuine users or impostors to measure the performance of the developed algorithm.
- Chapter 4 is about the data set collected for the present research. This chapter

addresses the validation of O1 from the first chapter of this thesis.

• Chapter 5 In this chapter it is presented the authentication algorithm based on free-text keystroke dynamics. First of all, the algorithm developed for processing the data obtained from the users is presented. The algorithm simulates user authentication based on keystroke dynamics and measures the obtained performances. The development of this algorithm is established by O2 from the first chapter of this thesis.

• Chapter 6 In this chapter it is presented a series of experiments performed to measure the performance of the written algorithm for the purpose of this research and to analyze the results obtained. This chapter addresses the validation of O3 and O4 from the first chapter of this thesis.

• Chapter 7 summarizes the conclusions drawn from the previous chapters and future research directions in this field, starting from the results presented in this paper. The author's own contributions to the field of keystroke dynamics are presented in this chapter. The personal contribution: the proposal of two new metrics for calculating the distance between two vectors in order to allow the approximation of the degree of similarity between two patterns from two different users or from the same user. Also, the data collected from the 80 users about how to type on the keyboard is a contribution to the advantage of future researches because they will be available to all researchers interested in conducting investigation in the field. Another own contribution is the proposal of a pattern in order to retain the minimum necessary data about a user so to obtain performances in the continuous authentication.

2. STATE-OF-THE-ART

2.1 Evolution of educational systems

In this subchapter the author presents the evolution of MOOC (Massive Open Online Courses) platforms. In 2020, in the Coursera Platform are involved nearly 69 million learners [VAN20]. The number of Massive Open Online Courses increased in the last years. Debates about future and evolution of eLearning and MOOC (Massive Open Online Courses) were in the last few years. In this chapter the author makes an introspection in evolution of Massive Open Online Courses with a comparison of the most important platforms of MOOC. Also, in the last years, researchers have paid attention to Learning Analytics field [IVA16]. We have more and more data from Learning Management Systems. There were noticeable additional challenges regarding the field of education in 2020. With the COVID-19 pandemic the authorities have not only introduced restrictions on the movement of citizens, but have also tightened the preventive measures implementing new regulations with reference to education. A decisive number of universities have had to adapt to the unfamiliar circumstances, moving all their activities to the online environment. These limitations have led to an unprecedented leap in online education. Suddenly, both teachers and students or pupils, were forced by the newly implemented conditions to move their entire activity to online educational platforms and thus continue their courses in this manner. This process has led to a development of the e-learning section, helping the growth of companies that are being active in this field and has forced those who have not used these systems so far to learn them in a very quick way [IAP14a].

The educational system has continually evolved due to technological innovations. In [DAN12] the author made an enumeration of innovations: In 1841 the blackboard, in 1940 the motion picture, in 1957 the television. Programmed learning and computers were another invention which contributed on education evolution. Internet and communication technologies could develop the format of education [IAP14a].

The MOOC evolution starts with the "Connectivism and Connective Knowledge" – CCK08 course in 2008 which had a large number of online participants. The course was facilitated by Downes and Siemens [DOW14] [IAP14a].
The MOOC starts in the 2008 but the year 2012 was declared the MOOC year. The next years after 2012 was good years for MOOC, with millions of learners and hundreds of partners involved to develop courses [IAP14a].

A record number of users turned to online learning in 2020. Since March, there were more than 69 million enrollments only on Coursera. About 430% increase compared to the same period last year [VAN20] [IAP21b].

2.2 Keystroke dynamics – literature review

Keystroke dynamics is a research field with more and more importance in network access control and cyber security [ZHO12] [IAP21b]. For now, only a few studies are about free-text keystroke dynamics, the way that the users type what text the user wants. Most of them are analyzed only fixed text, static text [ZHO12] [SAL10] [ZAC10]. Fixed content and fixed length data are usernames or passwords [MON02]. Free text requires two phases: the user enrollment phase in the system and the user verification phase [MON02].

First, the use for users identification was researched in the 1970`s [ZHO12]. Spillane wrote his conclusions about the first investigation in 1975 [FOR77] and Forsen, Nelson and Staron in 1977 [SPI75]. 'Fist of the Sender' was a methodology in World War II that was used to identify, by using the rhythm, the sender of the telegraph. [BAN12] [VAC07] [DUN08][IAP21a].

Keystroke dynamics have been studied mostly in connection to authentication, but some studies, such as [MES11], have also studied the detection of emotional states of the user who uses the keyboard. Other studies focus on predict users age and gender from unintentional traces, that left behind by use of keyboard and mouse [AVA17]. In [SAL18], the authors explored the relevance of individual and general keyboard and mouse interaction patterns and they had modeled user`s keystroke dynamics and mouse movements with data mining techniques to detect the emotion of users in real-world learning scenarios [IAP21a]. In [LIM14], the authors indicates that automatic analysis of human stress from mouse input and keyboard input is potentially useful for providing adaptation in e-learning systems [IAP21a].

Typing behavior for continuous authentication is a biometric modality proposed in [ROT14]. The authors collected a video database from 63 users with static text and free text typing and developed computer vision algorithms to extract hand movement from the video stream.

If most studies use only data retrieved from the keyboard, there are studies that use a mixed method of user identification, based on data retrieved from the keyboard, but also on data retrieved from the mouse [LOZ17]. Additional features, like pressure, are used in addition to time-based features, but to capture this data you need touch screens or other special devices [TEH13]. The stages that a research in this field goes through are: extracting the keyboard features, creating user profiles and updating them and identifying the efficiency criteria [KOC19] [IAP21a].

Commercial keystroke dynamic products exist. In 2003, the paper [ILO03] presents the company BioNet Systems which patented the BioPassword authentication system [ZIL98]. In Romania, Typing DNA is a company, a start-up, that received funds of 6.2 million euros in 2020 to create a typing identity for security [STE20].

Other studies, like [ARW17], incorporates the use of nonconventional typing features using free text typing dynamics. Semi-timing features along with the editing features were extracted from the users' typing flow and decision trees were used to classify each of the user

data.

Algorithms of dynamic authentication can be divided into three major groups: estimation of metric distances, statistical methods and machine learning. Methods of keyboard recognition used in the literature are: distance, neural networks, statistical, probabilistic, machine learning, clustering, decision tree, evolutionary computing, fuzzy logic or other [KOC19] [IAP21a].

Some limitations of keystroke dynamics previous research were: it took a long time to train the model, data were manual preprocessed by human or large database was required [YUE04]. The authors from [YUE04] conclude that use of keystroke dynamics can make a more secure system.

3. RESEARCH METHODOLOGY

The steps performed in the present scientific research are described below:

A. Development of the platform for the acquisition of input data

B. Acquisition and initial processing of input data from 80 volunteers (how typing on their keyboard)

C. Processing the input data so as to generate a user pattern for each user

D. Development of an algorithm in the C programming language for calculating distances used in keystroke dynamics authentication

E. Simulation of system authentication by genuine users or impostors to measure the performance of the developed algorithm

The first two steps of the research methodology, A. Development of the platform for the acquisition of input data and B. Acquisition and initial processing of input data from 80 volunteers (how typing on their keyboard), have the role of approaching O1, as described in the first chapter of the thesis: to collect a database with the test pattern from at least 80 users, in order to test the authentication algorithm for this research, but also to make it available to other interested researchers.

The third step of the research methodology, C. Processing the input data so as to generate a user pattern for each user, and the last step of the research methodology, E. Simulation of system authentication by real users or imposters to measure the performance of the developed algorithm, have the role of approaching O4, as described in the first chapter of the thesis: to propose a data structure as efficient as possible, which should contain the most relevant information about the typing of a user.

The third step of the research methodology, C. Processing the input data so as to generate a user pattern for each user, and the forth step of the research methodology, D. Development of an algorithm in the C programming language for calculating distances used in keystroke dynamics authentication, have the role of approaching O2, as described in the first chapter of the thesis: to implement an algorithm for authenticating the users of a computer based on the keystroke dynamics, the keyboard typing mode.

The last step of the research methodology, E. Simulation of system authentication by genuine users or impostors to measure the performance of the developed algorithm, has the role of approaching O3, as described in the first chapter of the thesis: to propose at least two new metrics for calculating the distances between two vectors that generate better performance compared to the Equal Error Rate (EER) performance indicator than the classical methods.

4. FREE-TEXT KEYSTROKE DYNAMICS DATA SET FOR CONTINUOUS AUTHENTICATION

To research in the field of keystroke dynamics biometrics the researchers need input data obtained from computer users in different real situations. The necessary data are represented by the keys typed on the keyboard but also by the times at which they are pressed. The time when a certain key is pressed, respectively the time when a certain key is raised. The difference between these times is the keystroke time. Another important piece of information is the time between two keys. The difference between the time a key was released and the time a next key was pressed [IAP21a].

This information can only be obtained in a restrained or controlled environment, with the consent of those participating to this experiment. The agreement of the participants is necessary because it exists a possibility to form the initial text that the user typed on the keyboard with access to this data, and if, for example, a user is monitored while sending e-mails or doing other activities, the information may be confidential.

In the literature there are several sets of data that are accessible for research purposes. In the first phase, the author used these data sets. Most are represented by texts in English, obtained from the educational environment, by researchers from their university colleagues or from students.

Some data sets are retrieved by a specific program, in a special environment made for this purpose. Others are made to monitor everything that is typed on a computer, regardless of the program used at one time by the user. It monitors everything typed on the keyboard and typing times whether the user is writing e-mails, writing in a Word, Excel document or programming on a computer in a certain programming environment.

On the other hand, for the purpose of the research the author developed their own environment to obtain data from volunteers. The author has created a web environment for taking over keys and typing times in JavaScript. A form is created that takes over the keys and typing times while completing a form on a web page [IAP21a]. The website was created on the sites.google.com platform. The web platform can be accessed at https://sites.google.com/view/cataliniapa.

To capture the keys and typing times the author created a web form through which users were invited to answer several generic questions. The text entered from the keyboard by each user should be written freely by each user, without the need to reproduce a specific predefined text. At each text box, a series of generic questions were formulated to guide the user to a certain topic in the text he completed. The questions asked were about the weather, the ideal day or the educational system. To form the database for research is not relevant the topic of the text, but the way it is written.

The text written by users is in Romanian. Most datasets in the literature are texts captured from users who have written in English [IAP21a].

The form created to purchase data sets for research purposes was completed by a number of 80 users. They handed over data for 410,633 key-events [IAP21a]. The average number per user is 5132 key events. The comprise time used by all 80 users to complete the form was 23 hours, 28 minutes and 19 seconds. The average time spent by users on the data collection platform is 17 minutes and 36 seconds.

5. ALGORITHM DEVELOPMENT FOR KEYSTROKE DYNAMICS AUTHENTICATION

The architecture of the keystroke dynamic authentication system has two important parts. The first is the system training phase part, part in which users enroll in the system providing data on how to type. In this phase a pattern is created for each user and is stored in the database to be used in the continuous authentication phase. The second part is the continuous authentication phase. In this phase the system continuously verifies the users connected with a valid username and password. Throughout the time a user is logged in to the account, the system takes data from it on the typing mode and continuously compares the resulting pattern with the pattern in the database. As long as there is acceptable similarity between the two patterns the user remains logged in to the system. When the system finds that the two patterns are no longer similar, the one taken from the user logged in to the account and the one from the database, the system generates an alarm signal and the user is removed from the account. He can re-enter the account by re-entering the username and password [IAP21a].

The collection and initial processing of input data are the first steps taken in order to obtain key events of each user. This data represents the input data for the continuous authentication algorithm based on keystroke dynamics. Once the sample size is set, the next step is to divide the user data into key event sequences. The algorithm transforms key events into information about keys and information about diagrams, and then forms the time vectors needed to calculate distances. After the steps described above have been completed, the distances between the vectors are calculated, in order to establish the similarity between two users. Four types of distances are used: Euclidean distance, Manhattan distance, R distance and A distance. With these distances calculated for each user in the database, we proceed to simulate the authentication in the system, in turn by each user in the database. Following the simulation of the authentication in the system, four performance indicators of the algorithm are generated: False Acceptance Rate (FAR), False Rejection Rate (FRR), True Acceptance Rate (TAR) and True Rejection Rate (TRR). Based on these, the Equal Error Rate (EER) can be calculated, the main indicator of the performance of the algorithms used in this thesis. Also, to view the performances, two graphs are generated: FAR and FRR chart and ROC curve.

6. EXPERIMENTS AND RESULTS - SIMULATION OF SYSTEM AUTHENTICATION BY GENUINE USERS OR IMPOSTORS

In this chapter it is presented a series of experiments performed to measure the performance of the written algorithm for the purpose of this research and to analyze the results obtained. Gradually, experiments with the keystroke time of a single key, in the subchapter 6.1, and experiments with di-graphs, in the subchapter 6.2, are presented. Both in the analysis of the characteristics with a single key and with a di-graph, the degree of Equal Error Rate (EER) is calculated in order to appreciate the performances of the algorithms. The results are presented in the case of experiments using Euclidean distance (in the subchapters 6.1.1 and 6.2.3), Manhattan distance (in the subchapters 6.1.2 and 6.2.4), R distance (in the subchapter 6.1.3) and A distance (in the subchapters 6.1.4 and 6.2.5). The chapter also investigates, in the subchapter 6.1.5 The sample size, the differences in performance if the pattern is built for each user with various sample sizes, starting from 200 key events / pattern and up to 3000 key events / pattern. At the end of the chapter, following all the experiments

performed and presented, the author proposes, in the subchapter 6.4, Proposing new metrics for calculating distances between users, the modification of two metrics obtaining new metrics for calculating the distances between two vectors that have higher performances than the classical calculation methods. For the two new metrics, the performances obtained in terms of Equal Error Rate (EER) are presented. By proposing these metrics, O3 is validated. It also proposes, in the subchapter 6.5, Proposed user pattern, a structure for retaining a user's pattern, a structure that takes up small memory and requires little time to perform all the necessary calculations in the algorithms. By proposing the user pattern, O4 is validated. In the end of the Chapter, in the subchapter 6.6 Comparison of the related works, the performances obtained in the present research are compared with those obtained by other authors in their researches.

7. CONCLUSIONS AND FUTURE WORKS

The four objectives formulated in the first chapter were approached and validated during the presented chapters.

7.1.1 The personal contributions

The personal contributions presented in this research are:

1. A free-text keystroke dynamics algorithm for continuous authentication has been developed. The algorithm can be found in Appendix 1 - Free-text keystroke dynamics algorithm for continuous authentication and it was presented in Chapter 4.
2. It was created a database with typing mode from 80 users, 410.000 key events, a total time of approximately 24 hours for the acquisition of the necessary data. Detailed in Chapter 5
3. A modify Manhattan distance metric has been proposed, calculated on the most used 14 letters. The proposed new distance metric improves performance coefficient EER from 7.13% to the value of 5.33%. This means a 25.24% improvement in performance. Details about the proposed new metric are in the subchapter 6.5 Proposing new metrics for calculating distances between users, 6.4.1 New metric for calculating distances based on individual key time.
4. A modify distance metric has been proposed, calculated on the most used 12 di-graphs. The proposed new distance metric improves performance coefficient EER from 5.23% to the value of 3.27%. This means a 37,47% improvement in performance. Details about the proposed new metric are in the subchapter 6.4 Proposing new metrics for calculating distances between users, 6.4.2 New metric for calculating distances based on di-graphs times.
5. A structure for user pattern with the efficiency of the space used but also with the premises to make the necessary calculations in a short time has been proposed. The total space occupied by such a pattern for a user is only 256 bytes (64 floats). The proposal formulated for the retention of the pattern is represented in subchapter 6.5 Proposed user pattern.

7.2 Future works

This thesis has reached its established objectives, and the conclusions presented in the previous subchapter open new possibilities to continue research in new directions, such as:

• Expanding the keystroke dynamics database by collecting data from a larger number

of users;

• Expanding the database by collecting data from the 80 users in new sessions in order to research the evolution of the typing pattern over time

• Analysis of new algorithms, based on different techniques compared to calculating distances between time vectors

• Applying the metrics proposed in this paper to other databases available from another scientific research

• Analysis of the particularities of the special characters from the Romanian language, which are not found in English: Ă, Î, Â, Ș, Ț.

• Character analysis punctuation, SPACE, ENTER, TAB, BACKSPACE etc.

• Changing data collection conditions: changing the keyboard, under stress, etc.

• Analysis of the word in which the di-graph appears

• Developing keystroke dynamics authentication algorithms based on tri-graphs

• Developing keystroke dynamics authentication algorithms based on n-graphs

REFERENCES

[ARW17] Arwa Alsultan, Kevin Warwick, Hong Wei, Non-conventional keystroke dynamics for user authentication, Pattern Recognition Letters, Volume 89, 2017, Pages 53-59, ISSN 0167-8655

[AVA17] Avar Pentel. 2017. Predicting Age and Gender by Keystroke Dynamics and Mouse Patterns. In Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization (UMAP '17). Association for Computing Machinery, New York, NY, USA, 381–385. DOI:https://doi.org/10.1145/3099023.3099105

[BAN12] Banerjee, Salil & Woodard, D.L.. (2012). Biometric Authentication and Identification Using Keystroke Dynamics: A Survey. Journal of Pattern Recognition Research. 7. 116-139. 10.13176/11.427.

[DAN12] J. Daniel. 2012. Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. Technical Report. Korea National Open University. http://www.tonybates.ca/wp-content/uploads/Making-Sense-of-MOOCs.pdf Retrieved February 2014

[DOW14] Downes, S. 2008. Places to go: Connectivism & Connective Knowledge. Innovate 5 (1). http://www.innovateonline.info/index.php?view=article&id=668 January 2014

[DUN08] T. Dunstone and N. Yager. Biometric System and Data Analysis: Design, Evaluation, and Data Mining. Springer, 1 edition, 2008.

[FOR77] G. Forsen, M. Nelson, and R. Staron, Jr. "Personal attributes authentication techniques", Technical Report RADC-TR-77-333, Rome Air Development Center, October 1977.

[IAP14a] Iapa, A.C. (2014), Outstanding research in MOOC and future development, Proceedings of the 10th International Scientific Conference "eLearning and Software for Education" Bucharest, Editura Universitatii Nationale de Aparare "Carol I" 2014 Volume 1, 251-254, DOI: 10.12753/2066-026X-14-035

[IAP21a] Iapa A.C., Cretu V.I., Modified Distance Metric That Generates Better Performance

For The Authentication Algorithm Based On Free-Text Keystroke Dynamics, IEEE 15th International Symposium on Applied Computational Intelligence and Informatics, Timisoara, Romania, 2021 – paper sent, unpublished

[IAP21b] Iapa A.C., Cretu V.I., Evaluating the performance of authentication algorithms based on keystroke dynamics used in online educational platforms, The 17th International Scientific Conference eLearning and Software for Education, Bucharest, Romania, 2021 – paper sent, unpublished

[ILO03] Ilonen, Jarmo. (2003). Keystroke dynamics. Advanced Topics in Information processing–lecture (2003).

[IVA16] Ivanova, Malinka, Holotescu, C., Grosseck, G., Iapa, C. "RELATIONS BETWEEN LEARNING ANALYTICS AND DATA PRIVACY IN MOOCs." The International Scientific Conference eLearning and Software for Education. Vol. 3. " Carol I" National Defence University, 2016.

[KOC19] Kochegurova, Elena & Luneva, Elena & Gorokhova, Ekaterina. (2019). On Continuous User Authentication via Hidden Free-Text Based Monitoring: Volume 2. 10.1007/978-3-030-01821-4_8.

[LIM14] Y. M. Lim, A. Ayesh and M. Stacey, "Detecting cognitive stress from keyboard and mouse dynamics during mental arithmetic", Proc. Sci. Inf. Conf. (SAI), pp. 146-152, Aug. 2014.

[LOZ17] Lozhnikov, Pavel & Sulavko, Alexey & Ekaterina, Buraya & Viktor, Pisarenko. (2017). Authentication of Computer Users in Real-Time by Generating Bit Sequences Based on Keyboard Handwriting and Face Features. Voprosy kiberbezopasnosti. 24-34. 10.21681/2311-3456-2017-3-24-34.

[MES11] A. Messerman, T. Mustafic, S. A. Camtepe and S. Albayrak. Continuous and non-intrusive identity verification in real-time environments based on free-text keystroke dynamics. Proceedings of IEEE International Joint Conference on Biometrics. 1–8, 2011

[MON02] Monrose F, Reiter MK, Wetzel S (2002) Password hardening based on keystroke dynamics. Int J Inf Secur 1(2):69–83

[ROT14] J. Roth, X. Liu and D. Metaxas, "On Continuous User Authentication via Typing Behavior," in IEEE Transactions on Image Processing, vol. 23, no. 10, pp. 4611-4624, Oct. 2014, doi: 10.1109/TIP.2014.2348802.

[SAL10] E. Al Solami, C. Boyd, A. Clark, and A. K. Islam, "Continuous Biometric Authentication: Can It Be More Practical?", IEEE Int'l Conf. on High Performance Computing and Communications (HPCC), pp. 647-652, 2010.

[SAL18] S. Salmeron-Majadas, R. S. Baker, O. C. Santos and J. G. Boticario, "A Machine Learning Approach to Leverage Individual Keyboard and Mouse Interaction Behavior From Multiple Users in Real-World Learning Scenarios," in IEEE Access, vol. 6, pp. 39154-39179, 2018, doi: 10.1109/ACCESS.2018.2854966.

[SPI75] R. Spillane, "Keyboard Apparatus for Personal Identification", IBM Technical Disclosure Bulletin, vol. 17, no. 3346, 1975.

[STE20] Stefan Koritar. (2020). Romanian startup Typing DNA raises €6.2 million in Series A funding to create ‘typing identity’ for security (2020).

[TEH13] Teh, Pin Shen & Teoh, Andrew & Yue, Shigang. (2013). A Survey of Keystroke Dynamics Biometrics. TheScientificWorldJournal. 2013. 408280. 10.1155/2013/408280.

[UMP85] D. Umphress and G. Williams, "Identity Verification through Keyboard Characteristics", Int'l J. Man-Machine Studies, Vol. 23, No. 3, pp. 263-273, 1985.

[VAC07] J. R. Vacca. Biometric Technologies and Verification Systems. Butterworth-Heinemann, 1 edition, 2007.

[VAN20] Vandenbosch, B., Most Popular Courses of 2020: A Year of Mental Health, Contract Tracing, and Job-Relevant Skills, Coursera Blog.

[YUE04] Yu, Enzhe & Cho, Sungzoon. (2004). Keystroke dynamics identity verification - Its problems and practical solutions. Computers & Security. 23. 428-440. 10.1016/j.cose.2004.02.004.

[ZAC10] R. Zack, C. Tappert, and S. Cha, "Performance of a long-text-input keystroke biometric authentication system using an improved k-nearest-neighbor classification method", IEEE Int'l Conf. on Biometrics: Theory Applications and Systems (BTAS), pp. 1-6, 2010.

[ZHO12] Y. Zhong, Y. Deng and A. K. Jain, "Keystroke dynamics for user authentication," 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Providence, RI, 2012, pp. 117-123, doi: 10.1109/CVPRW.2012.6239225.

[ZHO15] Zhong, Yu & Deng, Yunbin. (2015). A Survey on Keystroke Dynamics Biometrics: Approaches, Advances, and Evaluations. 10.15579/gcsr.vol2.ch1.

[ZIL98] Zilberman, A.G.: Security method and apparatus employing authentication by keystroke dynamics (1998) United States Patent 6,442,692.