Keystroke Dynamics

CSCI 688- AI & Cybersecurity Final Paper

Eli Gnesin

May 14, 2022

1 Introduction

In recent years, the rise of personal computing has led to an increased need to protect personal accounts and Personally Identifiable Information online. Originally, a password might have sufficed for protection, but advancements by adversaries has necessitated improvements in security, the most prominent of which has been the introduction of biometrics. Biometrics are, in general, behavioral or physical attributes or properties that can be used to identify or authenticate an individual's profile [13]. The most common of these biometrics include fingerprinting and facial recognition, but biometrics are not only limited to these features.

One emerging biometric method is Keystroke Dynamics. Keystroke Dynamics, also called keystroke recognition, is the set of measures related to a person's typing rhythm, more specifically the exact timing of each key-press, hold, and release in a person's typing pattern [13]. These timing patterns can also include sequences of keys, called digraphs or trigraphs. Most commonly recorded are the hold time, or "dwell time", the down/down time, or "press-press", and the up/down time, or "flight time," depicted in Figure 1.

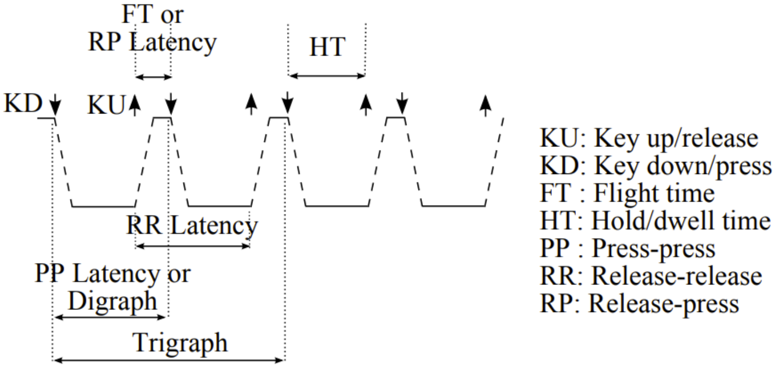


Figure 1: The different timings that can be recorded for keystroke dynamics. [7]

Unlike other biometric methods, Keystroke dynamics are generally considered to be lightweight and unobtrusive, requiring no extra hardware beyond a standard keyboard, and being able to augment current protection methods [7]. Given this, this paper will first explore the history of Keystroke Dynamics, then discuss the current state of Keystroke Dynamics in cybersecurity, and finally present a new research idea in the field.

2 A Brief History of Keystroke Dynamics

While the use of Keystroke Dynamics in security is a relatively new phenomenon, the concept of tracking individual typing patterns extends back to the mid-19th century. At that time, telegraph operators could distinguish and identify other operators based on the rhythm they typed with [7]. This continued into World War II, where the concept of the "Fist of the Sender" developed. During the war, telegraph operators could identify the typing patterns of other operators using the rhythm and pace, and then use that to track allied and enemy troop movements [7].

The focus on using keystroke dynamics for user identification and authentication picked up again in 1980 with several experiments conducted by the RAND Corporation, led by R. Stockton Gaines. With these experiments, Gaines hoped to demonstrate that "keystroke timing" could be used for authentication purposes for sensitive information [11]. More significantly, Gaines's research was the first to consider the concept of digraphs, or successive letters typed, for the purposes of authentication. In his results, Gaines and his team found statistically significant results for comparing typists to themselves and the other typists in the group using five different digraphs [11].

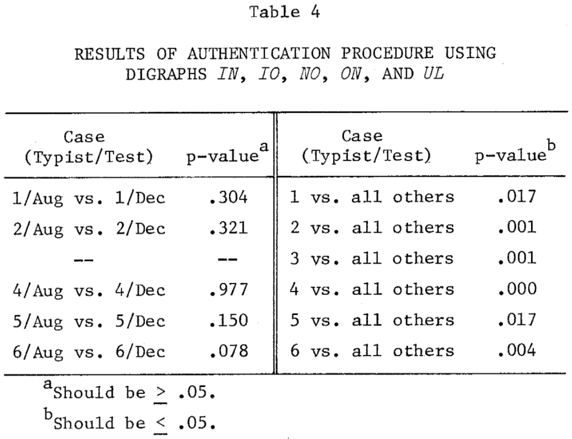


Table 4

RESULTS OF AUTHENTICATION PROCEDURE USING DIGRAPHS *IN*, *IO*, *NO*, *ON*, AND *UL*

Case (Typist/Test)	p-value[a]	Case (Typist/Test)	p-value[b]
1/Aug vs. 1/Dec	.304	1 vs. all others	.017
2/Aug vs. 2/Dec	.321	2 vs. all others	.001
--	--	3 vs. all others	.001
4/Aug vs. 4/Dec	.977	4 vs. all others	.000
5/Aug vs. 5/Dec	.150	5 vs. all others	.017
6/Aug vs. 6/Dec	.078	6 vs. all others	.004

[a] Should be ≥ .05.

[b] Should be ≤ .05.

Figure 2: The digraph results from the Gaines experiment. [11]

In the 1990s, the focus on Keystroke Dynamics shifted firmly into the cybersecurity sphere and began to incorporate the growing field of Artificial Intelligence and Machine Learning [5]. In 1990, Bleha et al. explored keystroke dynamics for computer access purposes using a Bayes-classifier method, with findings that were 95% accurate overall in recognition [8]. Seven years later, Monrose and Rubin honed in on the concept of authentication using keystroke dynamics even further, using clustering in C++ to identify and authenticate users [14].

Moving forward into the 2000s, and up into recent years, most research on Keystroke Dynamics has continued to focus on its uses in authentication, particularly with classification algorithms from machine learning. In the late 2000s, this focus shifted to also include the rise of touchscreens and mobile phones. In 2009, Saevanee et al. explored user authentication on mobile phones incorporating both keystroke dynamics and other features such as finger pressure. They found that finger pressure was the most useful feature for classification with their methods, but that the interaction of other keystroke methods was also a strong classifier [15]. These fields have been the primary drivers of research in Keystroke Dynamics up until recent years, where the focus has broadened to new topics and fields, which will be discussed in the Literature Survey (Section 4).

3 Keystroke Dynamics in Cybersecurity

3.1 Current Threat Detection

Currently, the focus on keystroke dynamics as a cybersecurity tool has centered around using the biometric for user authentication purposes. Most of this recent research has focused on using keystrokes at the password level. When an individual types a password, their typing rhythm can be recorded and compared, through any classification algorithm or matching scheme, to a "profile" to either affirm or reject the user as authentic. Most recently, these methods have been extended towards "continuous authorization." In this sense, a user's typing pattern is regularly, or consistently, matched against their profile, and any flag (or series of flags) could either be sent to a cybersecurity analyst, or trigger an automatic locking mechanism on the user's computer.

3.2 Products, Services, and Vendors

Due to its use as a non-intrusive, simple biometric method, the market for Keystroke Dynamics software is set to grow significantly in the coming years. Market research firm Allied Market Research, for example, predicts the market for Keystroke Dynamics will grow to over $700 million by 2025 [1]. Currently, there are two prominent companies that focus solely on using Keystroke Dynamics for user authentication and authorization tasks: KeyTrac and TypingDNA.

KeyTrac markets itself as a lightweight software solution for user authentication using keystroke dynamics. Their system records the "relative flight and dwell times" as a user types their password, then feeds that information back into their algorithm to produce a "match score," which gets passed back to the user's backend system [3].

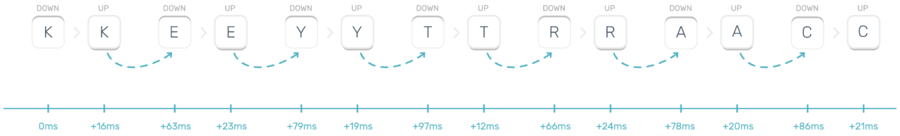


Figure 3: The "relative flight and dwell times" from KeyTrac. [3]

Unlike some other cybersecurity solutions, it is then up to the user's company to act that information, either with an analyst or with some other algorithmic prevention method if the match score is below some threshold. KeyTrac primarily emphasizes the ease of the system, that it can be immediately integrated into a website or web service with no extra overhead hardware. This touches upon one of the major benefits of keystroke dynamics as a cybersecurity solution; its relative unobtrusiveness and lightness makes it easy for small companies to improve their security with minimal overhead.

TypingDNA, in contrast to KeyTrac, has a much wider set of solutions to offer. TypingDNA offers an authentication solution, like KeyTrac, but also has solutions in the sphere of continuous authorization, which is relatively nascent as a keystroke dynamics field. For authentication, the TypingDNA authenticationAPI is largely similar to the KeyTrac offering. It similarly offers a lightweight software that can be easily integrated and requires no extra overhead in hardware [4]. Unlike KeyTrac, however, the TypingDNA system is also available on mobile devices and can incorporate mobile sensor data [4]. In contrast to the authentication systems, the TypingDNA continuous authorization system is slightly more complicated. The system, called *ActiveLock*, prevents device sharing by detecting whether a user is legitimate based on typing pattern. Further, *ActiveLock* also has prevention measures if a user is not legitimate by automatically locking the device, preventing any further damage from being done [2]. In the coming years, given the potential market growth for keystroke dynamics, other companies will certainly begin to occupy this sphere, but TypingDNA appears to currently have the continuous authorization market largely to itself for now.

4 Literature Review

Current research in the field of keystroke dynamics in connection with both cybersecurity and Machine Learning/Artificial Intelligence has touched upon four main areas overall: Feature Importance, Free Text, Deep Learning, and Mobile Devices. In this section, current research in each of these areas will be introduced.

4.1 Feature Importance/User Classification

A small subset of recent research has focused on the dual questions of what features of keystroke dynamics are most important, and can those features be used to classify users, as opposed to simply authenticating them. One recent example of this was a study by Tsimperidis et al. titled "Age and Gender as Cyber Attribution Features in Keystroke Dynamics-Based User Classification Processes" [17]. The dataset for this study was a set of 387 logfiles of keystrokes, separated by age/gender subsets (such as Male 18-25, Female 18-25, Male 26-35, etc.). First, the researchers used Information Gain, measuring entropy to determine which features were most significant in gender-age classification, finding that the hold or dwell time for most keys was generally containing the most information. With this complete, the researchers shifted to using five Machine Learning Algorithms, both with the full set of features and with the "most significant" features as determined by the information gain.

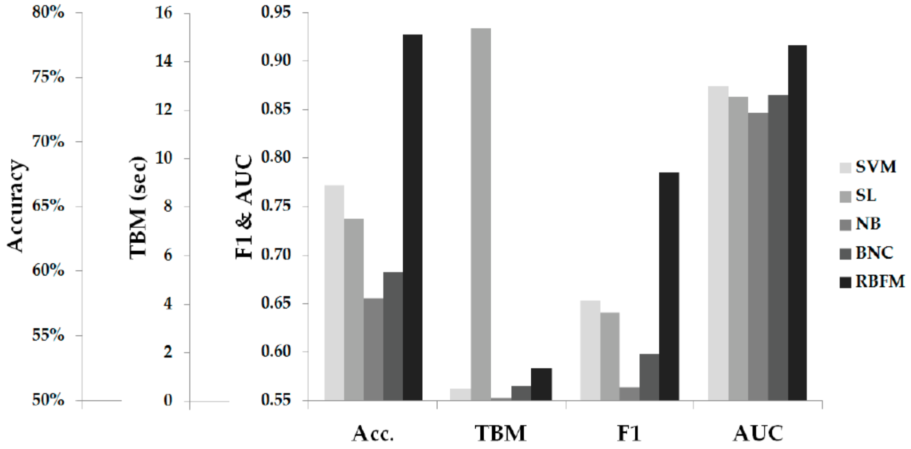


Figure 4: The results for the machine learning algorithms in the Tsimperidis et al. study. [17]

Having determined that the Radial Basis Function Network (mislabeled RBFM on Figure 4) was the best algorithm, they then used AdaBoost to improve the accuracy of the model, increasing the accuracy to over 80% and F1 score to over 0.8 [17].

4.2 Free Text CAPTCHA

A second recent field of research in keystroke dynamics has focused on free text data. Initially, most user authentication has centered around passwords; an individual types a password, their typing rhythm is checked using an algorithm, and they are authenticated or rejected. Theoretically, keystroke dynamics should be character-agnostic, however. Which keys are typed should not necessarily affect the overall rhythm, and we would not want to inadvertently record the keystrokes of an individual, which could be prone to data leaks. To combat this, in their 2022 study "Investigation of Using CAPTCHA Keystroke Dynamics to Enhance the Prevention of Phishing Attacks," Alamri et al. applied the concepts of keystroke dynamics to random alphanumeric CAPTCHA strings [6]. In this study, the researchers had participants "enroll" with seven random alphanumeric strings, and in the experimental phase, the "phishing" group had 6 attempts, also with random alphanumeric strings, to match a profile significantly enough (with the matching done through Euclidean distance between the means of the hold time, up/down time, and down/down time) to fool the system into being seen as "genuine." In the end, of the 45 illegitimate users, 17 were blocked due to exceeding the number of attempts (despite the small number of attempts used to build the user profile to begin with). In all, the study demonstrated that such a system, due to its lightweight nature and ease of use, could be applied as an extra layer of security on traditional attacks, while simultaneously demonstrating that free-text was a possible route for future research in keystroke dynamics.

4.3 Deep Learning

The bulk of recent research in keystroke dynamics has focused on different applications of deep learning and neural network based classification for the purposes of authentication. One growing aspect of this field is a focus on Recurrent Neural Networks (RNNs), including a 2021 paper by Zheng and Elmaghraby titled "Cybersecurity Enhancement Using Recurrent Neural Networks and Keystroke Dynamics" [10]. In this study, Zheng and Elmaghraby explored using RNNs, specifically because of their use of sequential information (in the context of sequential key-presses). They used the Carnegie Mellon Benchmark dataset, which contains 51 individuals each typing the same 10-character password string 400 times (over eight sessions). They then proposed using one-class and two-class Support Vector Machines (SVMs) and Long Short-Term Memory Neural Networks (LSTMs), and they judged results based on the Equal Error Rate (EER). They also used the Keystroke 100 dataset, which had 100 users typing the same password, but also collected finger pressure along with the timing data. Overall, they found that the LSTM models produced the best results, in general, but at significant time costs compared to the SVM models.

In another study on LSTM models, Soni and Prabakar sought to build an improved LTSM-based authentication model, named KeyNet. Their paper, titled "KeyNet: Enhancing Cybersecurity with Deep Learning-Based LSTM on Keystroke Dynamics for Authentication" explains their process for doing so [16]. Similarly to Zheng and Elmaghraby's study, Soni and Prabakar also used the Carnegie Mellon Benchmark dataset and proposed using a LSTM model to make predictions for authentication. However, they focused also on model optimization by tuning the batch size and epoch hyper-parameters [16]. In the end, they achieved approximately 90% recall and precision, and 99% accuracy, using a LSTM model with a SoftMax activation function in the output layer, an Adam optimization function, and a batch size of 64. These studies, and similar studies, show the increasing interest in using neural networks to harness the sequential nature of key-press events and improve authentication using keystroke dynamics.

4.4 Mobile Devices

Perhaps the most interesting new route of research in the field is a new focus on authentication, and more importantly authorization, in mobile devices. Unlike desktops, and even laptops, mobile devices are much easier to lose or misplace, so having continuous authentication or authorization becomes more important. Methods of continuous authentication were the focus of a 2021 study by de-Marcos et al. titled "Comparing Machine Learning Classifiers for Continuous Authentication on Mobile Devices by Keystroke Dynamics" [9]. In this study, the researchers used a public phone Hand Movement, Orientation, and Grasp (HMOG) dataset consisting of 100 users and 712,000 key-press events. They then built a training set for each participant using their keystroke events and a random subset of the events from the other 99 users. From there, they applied seven different Machine Learning algorithms, with reasonable results.

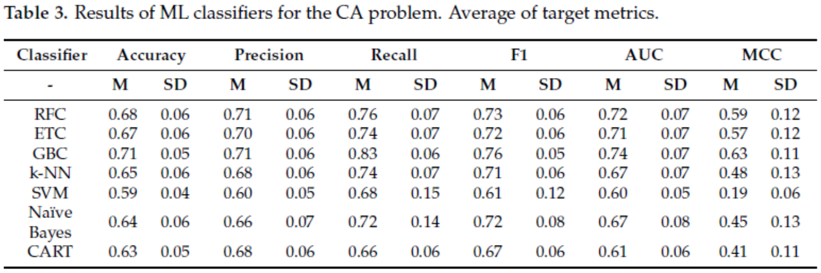


Table 3. Results of ML classifiers for the CA problem. Average of target metrics.

Classifier	Accuracy		Precision		Recall		F1		AUC		MCC	
-	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RFC	0.68	0.06	0.71	0.06	0.76	0.07	0.73	0.06	0.72	0.07	0.59	0.12
ETC	0.67	0.06	0.70	0.06	0.74	0.07	0.72	0.06	0.71	0.07	0.57	0.12
GBC	0.71	0.05	0.71	0.06	0.83	0.06	0.76	0.05	0.74	0.07	0.63	0.11
k-NN	0.65	0.06	0.68	0.06	0.74	0.07	0.71	0.06	0.67	0.07	0.48	0.13
SVM	0.59	0.04	0.60	0.05	0.68	0.15	0.61	0.12	0.60	0.05	0.19	0.06
Naïve Bayes	0.64	0.06	0.66	0.07	0.72	0.14	0.72	0.08	0.67	0.08	0.45	0.13
CART	0.63	0.05	0.68	0.06	0.66	0.06	0.67	0.06	0.61	0.06	0.41	0.11

Figure 5: The results of the de-Marcos et al. study [9]

The researchers argue, however, that although the results for a single key-press event were not highly accurate, the models are still useful. Instead, by stacking individual predictions, the probability of having consecutive false negatives (or positives) decreases significantly. As such, any implementation of this concept should only have the agent respond in the event of several consecutive flagged events, or a high percentage of flagged events in a short time period [9].

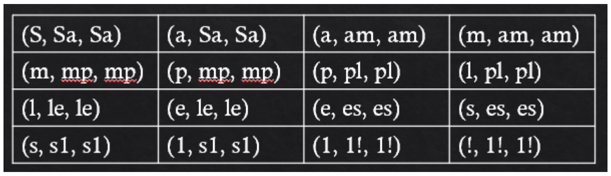
5 Original Research

Given the breadth of current research into the topic of Keystroke Dynamics in the realm of Artificial Intelligence or Machine Learning and Cybersecurity, especially with regards to deep learning and neural networks, there is a space for research that addresses questions of explainability and clarity with depictions of human typing patterns. Put simply, keystroke dynamics are currently presented as a list of times, corresponding to various hold times, flight times, and press-press times for a string of characters, and this information is then passed into a neural network or other classifier to produce a prediction. The average individual, however, will undoubtedly not understand, nor care, for a string of times that supposedly represent their individual typing pattern.

My future research idea is titled "Visual Mapping for User Authentication Convolutional Neural Networks (CNN) using Keystroke Dynamics." The research is two parts. First, I want to provide a simple visual method, similar to a heatmap, for understanding an individual's typing patterns. Then, I want to take those visuals as input for a CNN to be used for User Authentication. This neural network classifier would have one class per user, and would aim to take an input of a certain length and match it to the user in question (where a match means an authentic user and a non-match means a flagged user).

For this research, two datasets would be useful. The first is the Carnegie Mellon Benchmark Dataset [12]. The dataset contains 51 subjects, who each typed a 10 character password 400 times. In some sense, the second dataset that would be useful for this research is a modification of this first dataset. For this research, I would seek a dataset with some $N \approx 40$ typists, who each typed a string of $\frac{n^2}{2} + 1$ characters K times, where $K > 200$ and $n \geq 4, n\%2 = 0$. For each character, I would record the hold time for the key, as well as the up/down and down/down times for the two digraphs in which the key is included.

With this data, I can now construct the visual mapping of typing rhythm. For the sake of simplicity, consider the minimal case of $n = 4$, and thus we have a sequence of 9 letters. We then construct an 4×4 matrix, where each pixel is RGB coded by (hold, up/down, down/down) for each key, scaled 0-255 according to the minimum and maximum times in the dataset. The first and last letters in the sequence each have one pixel, and every intermediate letter has two pixels (one for each digraph the key is a part of). An example of this is depicted in Figure 6, below.



(S, Sa, Sa)	(a, Sa, Sa)	(a, am, am)	(m, am, am)
(m, mp, mp)	(p, mp, mp)	(p, pl, pl)	(l, pl, pl)
(l, le, le)	(e, le, le)	(e, es, es)	(s, es, es)
(s, s1, s1)	(1, s1, s1)	(1, 1!, 1!)	(!, 1!, 1!)

Figure 6: A sample matrix for the password Samples1! ($n = 4$)

With this mapping, an individual could see a simple heatmap that would represent their typing pattern given a string of characters. For someone without a strong grasp of data science or classification, having such a visual representation would make it much easier to understand what is happening behind the scenes. Finally, the method is character-agnostic, meaning it could be applied to something like a free-text CAPTCHA, where each individual string is a random set of alphanumeric characters.

With these visual mappings, the second part of the research would use the results as the training input into a convolutional neural network. The output of this network would have a probability that the input belonged to each of the N classes (one for each typist), and any probability over some predetermined threshold would be considered "authentic" (this threshold likely should be lower 50% due to the number of classes, but we also want to avoid an input potentially being considered "authentic" for multiple users, since that defeats the point of authentication). Overall, this second part is more of a proof of concept, to demonstrate that the visual mappings are useful beyond their explainability and simplicity for non-technical users.

6 Conclusion

The increased need for online account security in recent years has led to the growth of new security measures and biometrics. One of these growing biometrics is Keystroke Dynamics, which considers an individual's typing rhythm. In contrast to other biometric measures, keystroke dynamics are less intrusive, requiring less overhead and extra hardware. In recent years, the field has further grown, with vendors like KeyTrac and TypingDNA staking niches as User Authentication and Authorization softwares using keystroke dynamics. Likewise, there have been advances in research in recent years, primarily with free text keystrokes and with deep learning. With these, however, come issues of explainability and ease of understanding for non-technical users. To resolve this, I present a new "Visual Mapping" method of examining keystroke dynamics, with the ability to feed this mapping into neural networks for user authentication classification.

References

[1] Keystroke dynamics market size, share and analysis: Projection - 2025. https://www.alliedmarketresearch.com/keystroke-dynamics-market.

[2] Prevent unauthorized use of company computers with continuous authentication. ttps://www.typingdna.com/activelock-continuous-authentication.

[3] Technical details behind our solution. https://www.keytrac.net/en/technology.

[4] Typing biometrics authentication api. https://www.typingdna.com/authentication-api.htm.

[5] Nasir Ahmad, Andrea Szymkowiak, and Paul A. Campbell. Keystroke dynamics in the pre-touchscreen era. *Frontiers in Human Neuroscience*, 7, 2013. https://www.frontiersin.org/articles/10.3389/fnhum.2013.00835/full.

[6] Emtethal K. Alamri, Abdullah M. Alnajim, and Suliman A. Alsuhibany. Investigation of using captcha keystroke dynamics to enhance the prevention of phishing attacks. *Future Internet*, 14(3), 2022. https://www.mdpi.com/1999-5903/14/3/82.

[7] Salil Partha Banerjee and Damon Woodard. Biometric authentication and identification using keystroke dynamics: A survey. *Journal of Pattern Recognition Research*, 7(1):116–139, 2012. http://www.jprr.org/index.php/jprr/article/view/427/167.

[8] S. Bleha, C. Slivinsky, and B. Hussien. Computer-access security systems using keystroke dynamics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(12):1217–1222, 1990. https://ieeexplore.ieee.org/document/62613.

[9] Luis de Marcos, José-Javier Martínez-Herráiz, Javier Junquera-Sánchez, Carlos Cilleruelo, and Carmen Pages-Arévalo. Comparing machine learning classifiers for continuous authentication on mobile devices by keystroke dynamics. *Electronics*, 10(14), 2021. https://www.mdpi.com/2079-9292/10/14/1622.

[10] Adel S. Elmaghraby and Yufeng Zheng. Cybersecurity enhancement using recurrent neural networks and keystroke dynamics. In Sos S. Agaian, Vijayan K. Asari, Stephen P. DelMarco, and Sabah A. Jassim, editors, *Multimodal Image Exploitation and Learning 2021*, volume 11734, pages 90 – 98. International Society for Optics and Photonics, SPIE, 2021. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11734/117340D/Cybersecurity-enhancement-using-recurrent-neural-networks-and-keystroke-dynamics/10.1117/12.2585803.short.

[11] R. Stockton Gaines, William Lisowski, S. James Press, and Norman Shapiro. *Authentication by keystroke timing: Some preliminary results.* Rand, 1980. https://www.rand.org/pubs/reports/R2526.html.

[12] Kevin Killourhy and Roy Maxion. Keystroke dynamics - benchmark data set, 2009. https://www.cs.cmu.edu/~keystroke/.

[13] Rawlson King. Explainer: Keystroke recognition: Biometric update, Mar 2018. https://www.biometricupdate.com/201612/explainer-keystroke-recognition.

[14] Fabian Monrose and Aviel Rubin. Authentication via keystroke dynamics. *Proceedings of the 4th ACM conference on Computer and communications security - CCS '97*, 1997. https://dl.acm.org/doi/pdf/10.1145/266420.266434.

[15] H. Saevanee and P. Bhattarakosol. Authenticating user using keystroke dynamics and finger pressure. In *2009 6th IEEE Consumer Communications and Networking Conference*, pages 1–2, 2009. https://ieeexplore.ieee.org/document/4784783.

[16] Jayesh Soni and Nagarajan Prabakar. Keynet: Enhancing cybersecurity with deep learning-based lstm on keystroke dynamics for authentication. In Jong-Hoon Kim, Madhusudan Singh, Javed Khan, Uma Shanker Tiwary, Marigankar Sur, and Dhananjay Singh, editors, *Intelligent Human Computer Interaction*, pages 761–771, Cham, 2022. Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-030-98404-5_67.

[17] Ioannis Tsimperidis, Cagatay Yucel, and Vasilios Katos. Age and gender as cyber attribution features in keystroke dynamic-based user classification processes. *Electronics*, 10(7), 2021. https://www.mdpi.com/2079-9292/10/7/835.