Comparing Natural and Strong Typing Behavior for Keystroke Dynamics Multimodal Database Collection*

Markus Hoff Skudal[1], Matus Pleva[2][0000-0003-4380-0801], Štefan Korečko[2][0000-0003-3647-6855], Patrick Bours[1][0000-0001-5562-6957], and Daniel Hladek[2][0000-0003-1148-3194]

[1] Norwegian University of Science and Technology,
markuhs@stud.ntnu.no; patrick.bours@ntnu.no
[2] Technical University Kosice;
{Matus.Pleva;Stefan.Korecko;daniel.hladek}@tuke.sk

Abstract. Multimodal Keystroke Dynamics has been accurately shown to increase identification and authentication precision. When expanding this technique with EMG (Electromyography) analysis, a concern arose regarding a suspected low dynamic range of the EMG data. Thus a small-scale, multimodal experiment with five subjects was performed to display its potential, both with natural and strong typing behavior. This experiment lays the groundwork for a future multimodal Keystroke Dynamics project with 50–100 subjects. The results suggest that natural typing behavior gives high quality, and possibly better, data than strong typing. For 1-session-training, the evaluation shows user identification accuracy of 93% for natural and 85% for strong typing behavior. Hence, further research of scale could rely on natural typing behavior as the standard approach of recording.

Keywords: Keystroke Dynamics · User Identification · EMG · Timing Analysis · Typing Behavior.

1 Introduction

The experiment presented in this poster captured multimodal data in the form of (1) Keyboard timing, (2) EMG signals from Myo[3] armbands, (3) audio, and (4) video. More details on timing and audio keyboard analysis background can

* This work has been supported by the Slovak Research and Development Agency projects APVV-16-0202 & APVV-15-0517, by Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences under the research projects VEGA 1/0753/20 & VEGA 2/0165/21 and by Cultural and Educational Grant Agency of the Slovak Republic grant Nos. KEGA 009TUKE-4/2019 & KEGA 035TUKE-4/2019, funded by the Ministry of Education, Science, Research and Sport of the Slovak Republic.

[3] https://developerblog.myo.com/

be found in [4]. From this data, timing and EMG data sets were studied to determine typing behavior effectiveness. Recording of the Keystroke Dynamics data was exercised in four sessions for each of the five subjects, for each typing behavior as described in [3]. This resulted in 20 session recordings for each form of typing. On average these sessions lasted 80 seconds and recorded 25 correct typings of a predetermined word: "password". Ideally, the experiment should have had a larger subject base. This, however, proved difficult due to the COVID-19 pandemic and local, physical restrictions.

1.1 Keyboard Timing Analysis

In Keyboard Timing Analysis we measure when each key is pressed down and when it is released. From that, we can calculate the time a key is held down (called the duration of that key) and the time elapsed between typing two consecutive keys (called latency between the keys). Because the main goal of this preliminary research was to test the capability to use EMG for identification, we have only applied simple statistical analysis of the timing data, similar to what has been done by Pleva et al. in [4]. Besides the Scaled Manhattan Distance (SMD) that was used in [4], we also implemented the Scaled Euclidean Distance (SED), as these are rather similar in their implementations.

1.2 EMG Data Analysis

For analysis, the EMG data was loaded from CSV files into a data-handler for easy access and manipulation. With one Myo armband on each arm, and each armband producing eight EMG signals, each timestamp consisted of 16 data points. The sampling rate was found to vary between 165–175 Hz, though it originally was set to 200Hz. In order to compare the two forms of typing behavior, we chose MFCC (Mel-frequency cepstrum coefficients) [4, 5] with a 2-second window frame and 0.5-second step size to retrieve features. The MFCC was performed on each signal individually, with a dynamically adjusted sampling rate, and then the coefficients were combined in a 1x208 dimension array for each time frame (16 signals * 13 coefficients).In total, the experiment resulted in 2806 samples with 208 features, which stem from approximately 140 samples from each subject per session.

Despite our modest sample size, we hypothesized that the feature extraction from the MFCC would provide pattern rich samples for neural network analysis. We experimented with four Keras[4] models: GRU (Gated Recurrent Units), LSTM (Long Short-Term Memory), CNN 1D (one-dimensional Convolution Network), and a basic Feed-Forward-Network for reference. The hypothesis was demonstrated as plausible, as all networks cross-validated performed above 85%. The highest performing one-dimensional CNN-model was further used to compare natural and strong typing behavior. This model is composed of a Conv1D (32 nodes) and MaxPooling1D (5-sized kernel) layer, together with

[4] https://keras.io/

a Flatten, Dense (128 nodes), and Dense (5 nodes/classes, Softmax) layer. The following results are based on parameters batch size and epochs set to 64 and 30 respectively.

2 Results

2.1 Keyboard Timing Results

As mentioned above, we have used the SMD and SED distance metrics to evaluate the performance of the timing information. We have tested it on duration values only, latency values only, as well as the full set of duration and latency values together. The results are given in Table 1.

Table 1. Performance results from KD timing information

Typing Behavior:	**Natural**		**Strong**	
	SMD	SED	SMD	SED
Duration only	0.77	0.77	0.49	0.66
Latency only	0.55	0.63	0.74	0.76
All features	0.73	0.78	0.65	0.78

The results in Table 1, even though only based on the typing behavior of 5 persons, seem to indicate that duration values give a better performance when using natural typing, while latency features perform better for the strong typing behavior. When using the combination of duration and latency features we see that the natural typing behavior has a better performance when using SMD while for SED the performance is the same for both types of typing behavior.

Based on the results of the typing behavior there is no clear advantage in collecting strong typing behavior samples in addition to the normal typing behavior samples.

2.2 EMG Analysis Results

EMG Data Characteristics. From the raw EMG data, we retrieved important characteristics to describe its properties and compared the data from natural and strong typing behavior. Maximum and Minimum values were studied, together with the Median, Mean Absolute ($MA = \frac{1}{n}\sum_{t=1}^{n}|x_t|$), and Root Mean Squared ($RMS = \sqrt{\frac{1}{n}\sum_{t=1}^{n}x_t^2}$) [1, 2]. It shows that on average both were higher for *strong* typing behavior: MA 6.9 vs *7.4* and RMS 11.6 vs *12.4*.

Identification performance. When comparing the ability for identification using the best performed CNN 1D, both data sets provided above 85%. When training on one session the Natural typing behavior scored a good 92.8% while the Strong typing behavior gave an accuracy of 85.2% (see Fig. 1).

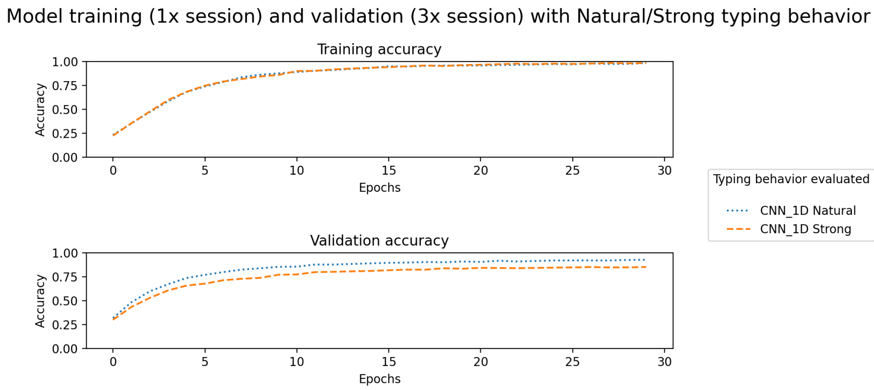

Fig. 1. The plot is the result of cross-validated session training on each type of typing behavior. The average of identification accuracy for validation gave a 92.8% vs. 85.2% with 1-session-training.

3 Discussion & Conclusion

As presented in this paper, the results suggest no clear benefit for further usage of Strong typing behavior in Multimodal Keystroke Dynamics Analysis. The indications, however, are not entirely aligned in this view. While EMG sensory identification shows beneficial results for the exclusive use of Natural typing behavior, the timing analysis is ambiguous. Based on these findings, we decided to efficiently allocate resources to Natural typing behavior only. Hence, further research with Multimodal Keystrokes Dynamics on 50–100 subjects will utilize this methodology in its recording.

References

1. Chong, H.J., Kim, S.J., Yoo, G.E.: Differential effects of type of keyboard playing task and tempo on surface EMG amplitudes of forearm muscles. Frontiers in Psychology **6**, 1277 (2015). https://doi.org/10.3389/fpsyg.2015.01277
2. Grabczyński, A., Szklanny, K., Wrzeciono, P.: Applying EMG technology in medial and lateral elbow enthesopathy treatment using Myo motion controller. Australas Phys Eng Sci Med. **42**(3), 701–710 (2019), doi.org/10.1007/s13246-019-00770-5
3. Korečko, Š., Pleva, M., Haluška, M., Skudal, M.H., Bours, P.: EMG input data collection for multimodal keystroke analysis. In: 12th IEEE Int. Conference on Cognitive Infocommunications, CogInfoCom 2021 Sept. 23-25. pp. 205–210. IEEE (2021)
4. Pleva, M., Bours, P., Ondáš, S., Juhár, J.: Improving static audio keystroke analysis by score fusion of acoustic and timing data. Multimedia Tools and Applications **76**(24), 25749–25766 (2017). https://doi.org/10.1007/s11042-017-4571-7
5. Rahim, M.A., Shin, J.: Hand movement activity-based character input system on a virtual keyboard. Electronics **9**(5), 774 (May 2020). https://doi.org/10.3390/electronics9050774