Comparative Study of Various Features-Mining-Based Classifiers in Different Keystroke Dynamics Datasets

Soumen Roy, Utpal Roy and D.D. Sinha

Abstract Habitual typing rhythm or keystroke dynamics is a behavioural biometric characteristic in Biometric Science relates the issue of human identification/ authentication. In 30 years of on-going research, many keystroke dynamics databases have been created on various pattern of strings (“greyc laboratory”, “.tie5Roanl”, “the brown fox”, …) taking various combination of keystroke features (flight time, dwell time) and many features-mining classification algorithms have been proposed. Many have obtained impressive results. But in evaluation process, a classifier’s average Equal Error Rates (EERs) are widely varied from 0 to 37 % on different datasets ignoring typographical errors. The question may arise, which classifier is best on which pattern of keystroke databases? To get the answer, we have started our experiment and created our own five rhythmic keystroke databases on different daily used common pattern of strings (“kolkata123”, “facebook”, “gmail.com”, “yahoo.com”, “123456”) and executed various classification algorithms in R statistical programming language, so, we can compare the performance of all the classification algorithms soundly on different datasets. We have executed 22 different classification algorithms on collected data considering various keystroke features separately. In the observation, obtained best average EER of the classifier Lorentzian is 1.86 %, where 2.33 % for Outlier Count, 3.69 % for Canberra, 5.3 % for Naïve Baysian and 8.87 % for Scaled Manhattan by taking all five patterns of strings and all combination of features in the consideration. So the adaptation of keystroke dynamics technique in any existing system increases the security level up to 98.14 %.

S. Roy (✉) · D.D. Sinha
Department of Computer Science and Engineering, University of Calcutta,
92 APC Road, Calcutta 700009, India
e-mail: soumen.roy_2007@yahoo.co.in

D.D. Sinha
e-mail: devadatta.sinha@gmail.com

U. Roy
Department of Computer and System Sciences, Visva-Bharati,
Santiniketan 731235, India
e-mail: roy.utpal@gmail.com

S.C. Satapathy and S. Das (eds.), *Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems: Volume 1*,
Smart Innovation, Systems and Technologies 50, DOI 10.1007/978-3-319-30933-0_17

Keywords Keystroke dynamics · EER · Canberra · Chebyshev · Czekanowski · Gower · Intersection · Kulczynski · Lorentzian · Minkowski · Motyka · Ruzicka · Soergel · Sorensen · Wavehedges · Manhattan distance · Euclidean distance · Mahanobolis distance · Z Score · KMean · SVM · Naïve Baysian

1 Introduction

Knowledge-based user authentication technique is very popular for its simplicity characteristics and users are very comfortable on it. But today, passwords or PIN is not limited due to brute-force, shoulder surfing or key logger attack. It demands higher level of security keeping simplicity with giving better performance. Some of the words we type daily like our name, address, email ID, password, … and we are habituated to type it in same rhythm which is unique just like our signature as a behavioral biometric characteristics and can be used in human identification or verification technique. It cannot be lost or stolen or mimicry in addition with no extra security apparatus is needed. The accuracy level of keystroke dynamics characteristics is not much promising in practice. It demands higher level of synthesis, so this technique can be effectively implemented in real life.

In our experiment, we have implemented Java Applet program to get the raw data of key press and release times of various pattern of strings ("kolkata123", "facebook", "gmail.com", "yahoo.com", "123456"), we have not taken password type strings or hard strings because users are not habituated to type hard strings in same rhythm. So, in the training session, we have considered some daily used words where all the users are habituated to press the words in same rhythm. Here getTime() function return the time of key press and release events in ms unit. Then we have calculated the following features of keystroke dynamics also in ms unit: time duration between key press and release of single key or key hold time or dwell time or key duration time (KD), time duration between two subsequent release or up-up key latency (UU), time duration between release of a key and press of next key or up-down key latency (UD), time duration between press of a key and release of next key or di-graph time or down-up key latency (DU), time duration between two subsequent press or down-down key latency (DD), total time (ttime), tri-gap time (trigap) and four-gap time (4gap).

Rhythmic keystroke is a behavioral biometric characteristics measured in Keystroke Dynamics methods is not a new. In the year 1897, Bryan and Harter investigated keystroke dynamics. In 1975, Spillane described the concept of keystroke dynamics. After that many researchers collected keystroke dynamics databases and evaluated by different classification algorithms, where statistical methods are common, many distance-based algorithms and many machine-learning methods have been also applied, many have obtained impressive results. But in evaluation process the classification's average Equal Error Rates are varied widely.

We have collected press and release time of 12096 keystrokes of 1440 samples of patterns from 12 different individuals in 4 different sessions with minimum of 1 month interval for five different common words ("kolkata123", "facebook", "gmail.com", "yahoo.com", "123456") in our experiment. Then we have considered 8 different features and combination of features then we have executed 22 different classifiers on that collected data. In our observation, obtained best average EER of the classifier Lorentzian is 1.86 %, where 2.33 % for Outlier Count, 3.69 % for Canberra, 5.3 % for Naïve Baysian, 8.87 % for Scaled Manhattan by taking all five patterns of strings and all combination of features in the consideration. So the adaptation of keystroke dynamics technique in any existing system increases the security level up to 98.14 %.

2 Keystroke Dynamics

2.1 *Basic Idea*

Keystroke dynamics is a technique work on set of some timing or pressure data of keystroke which is generated at typing on keyboard which is unique and can be used to classify the users.

2.2 *Science and Features Selection*

Placement of fingers on keyboard, hand weight, length of finger, neuro-physiological factors are made typing style unique. We have calculated key press and release time P_i and R_i for key K_i which represent entered character set, where $6 \le i \le$ length of the entered pattern. The features of the keystroke dynamics as follows [1]:

$$\text{Key Duration}(T_1) = R_i - P_i \tag{1}$$

$$\text{Up Up Key Latency}\ (T_2) = R_{i+1} - R_i \tag{2}$$

$$\text{Down Down Key Latency}(T_3) = P_{i+1} - P_i \tag{3}$$

$$\text{Up Down Key Latency}(T_4) = P_{i+1} - R_i \tag{4}$$

$$\text{Down Up Key Latency}(T_5) = R_{i+1} - P_i \tag{5}$$

$$\text{Total Time Key Latency}(T_6) = R_n - P_1 \tag{6}$$

$$\text{Tri-graph Latency } (T_7) = R_{i+2} - P_i \tag{7}$$

$$\text{Four-graph Latency } (T_8) = R_{i+3} - P_i. \tag{8}$$

2.3 Keystroke Dynamics as User Authentication

There are different ways in which a user can be authenticated. However all of these ways can be categorized into one of three classes: "Something we know" e.g. password, "Something we have" e.g. token, "Something we are" e.g. biometric property. Here, keystroke dynamics is the combination of three, something we know that is the pattern of strings, and something we have that is our finger tips size or our hand weight and something we are that is our typing style what we have learned in our life.

2.4 Factors Affecting Performance

Some of the factors which affect the way of keystroke Dynamics as follows: Text length, sequences of character types, word choice, and number of training sample, statistical method (mean or median) to create template, mental state of the user, tiredness or level of comfort, keyboard type, keyboard position and height of the keyboard, hand injury, weakness of hand mussel, shoulder pain, education level, computer knowledge, and category of users.

2.5 Algorithms

Following features mining algorithms [2, 3] can be applied on keystroke dynamics database. Here, P refers to the training set and Q refers to the test set. Mean and standard deviation is represented by μ and α respectively.

Canberra: $D_{car} = \sum_i^n \frac{\lvert P_i - Q_i \rvert}{P_i + Q_i}$	Minkowski: $D_{mink} = \sqrt[p]{\sum_i^n \lvert P_i - Q_i \rvert^p}$	Euclidean Distance: $D_{eu} = \sqrt[2]{\sum_i^n (\lvert P_i - Q_i \rvert)^2}$
Chebyshev: $D_{cheb} = \sum_i^n \max \lvert P_i - Q_i \rvert$	Motyka: $D_{mot} = \frac{\sum_i^n \max(P_i,) \rvert}{\sum_i^n (P_i + Q_i)}$	Mahanobolis Distance: $D_{maha} = \sqrt[2]{\sum_i^n ((\lvert P_i - Q_i \rvert)/\alpha_i)^2}$
Czekanowski: $D_{cze} = \frac{\sum_i^n \lvert P_i - Q_i \rvert}{\sum_i^n (P_i + Q_i)}$	Ruzicka: $D_{ruz} = 1 - \frac{\sum_i^n \min(P_i, Q_i)}{\sum_i^n \max(P_i, Q_i)}$	Z Score: $D_z = \sum_{i=1}^n (\lvert P_i \rvert - \mu(\lvert Q_i \rvert))/\alpha_i$

(continued)

(continued)

Gower: $D_{gow} = \frac{1}{n}\sum_i^n \lvert P_i - Q_i \rvert$	Soergel: $D_{soe} = \frac{\sum_i^n \lvert P_i - Q_i \rvert}{\sum_i^n \max(P_i, Q_i)}$	Lorentzian: $D_{lor} = \sum_i^n \ln(1 + \lvert P_i - Q_i \rvert)$
Intersection: $D_{ins} = \frac{1}{2}\sum_i^n \lvert P_i - Q_i \rvert$	Sorensen: $D_{sor} = \frac{\sum_i^n \lvert P_i - Q_i \rvert}{\sum_i^n (P_i + Q_i)}$	Manhattan Distance: $D_{man} = \sum_{i=1}^n (\lvert P_i - Q_i \rvert)$
Kulczynski: $D_{kuld} = \frac{\sum_i^n \lvert P_i - Q_i \rvert}{\sum_i^n \min(P_i, Q_i)}$	Wavehedges: $D_{wv} = \frac{\sum_i^n \lvert P_i - Q_i \rvert}{\sum_i^n \max(P_i, Q_i)}$	

Table 1 Background of keystroke dynamics

Authors	Classifiers	Length of the pattern	Features	EER (%)
Joyce and Gupta [4]	Manhattan	33	UD	0.25–16.36
Bleha et al. [5]	Euclidian	11–17	UD	2.8–8.1
Haider et al. [6]	Nural Network	7	UD	16.1
Yu and Cho [7]	SVM	6–7	UD	10.2
Killourly S. [8]	Manhattan (Scaled)	10+	UD	9.6
Kang et al. [9]	K mean	7–10	KD, UD	3.8
Giot et al. [10]	SVM	100	KD, UD	15.28

3 Background Details

In 30+ years of experience, many researchers have proposed their algorithms, taking various features and various length of pattern strings (Table 1).

4 Experimental Setup and Database

We have collected key press and release time as raw data in ms unit of all the entered keys for all five patterns using Java Applet program from 12 individuals during 12 months. Then we have calculated 8 different keystroke features using Eqs. 1–8. Then we have implemented 22 different classification algorithms (Canberra, Chebyshev, Czekanowski, Gower, Intersection, Kulczynski, Lorentzian, Minkowski, Motyka, Ruzicka, Soergel, Sorensen, Wavehedges, Manhattan Distance, Euclidean Distance, Mahanobolis Distance, Z Score, K-Mean, SVM, Naïve Baysian) in R Statistical language and evaluated the performance on the databases of the different pattern of strings “kolkata123”, “facebook”, “gmail.com”, “yahoo.com”, “123456” and combination of all strings.

5 Experimental Results

The following line chart represents the time stamps of 6 sample of same subjects and we can see the ratio of time tamps are almost same (Fig. 1).

The following line chart represents the 6 sample of time stamps of 6 different subjects and we can see the different time tamps ratios for 6 different subjects (Fig. 2).

In our experiment, we have seen that most of the time some distance-base algorithms achieved impressive results given bellow. Here, EERs are calculated in % (Table 2).

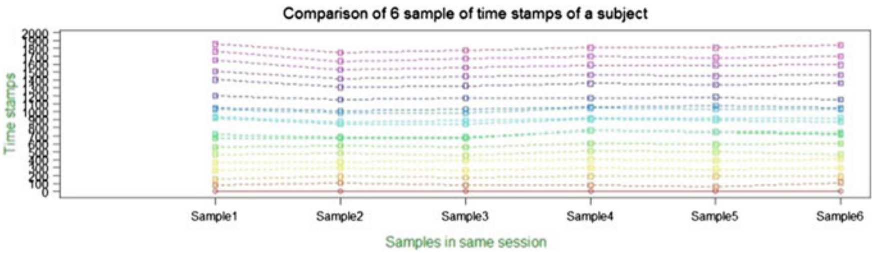

Fig. 1 Comparisons of 6 samples of time stamps of a subject

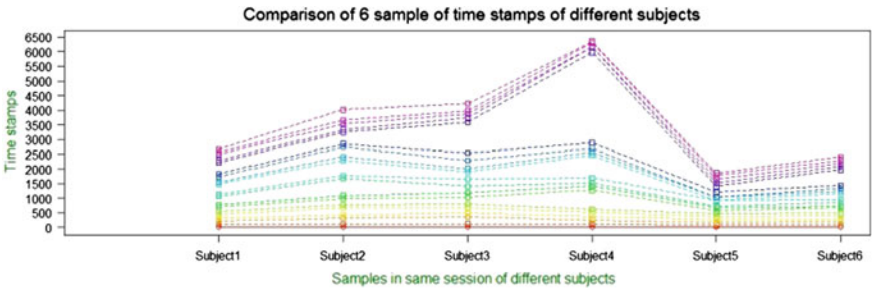

Fig. 2 Comparisons of 6 ample of time stamps of different subjects

Table 2 Obtained average EERs of different patterns for different classification algorithms

Classifiers	All strings	kolkata123	Facebook	gmail.com	yahoo.com	123456	.tie5Roanl [8]
Canberra	**3.69**	**8.93**	**10.83**	**12.47**	**12.92**	**15.18**	**29.74**
Cheby	12.53	13.48	15.85	16.16	20.04	19.24	25.68
Czekanowski	11.40	14.36	14.61	15.88	19.26	18.53	30.77
Gower	51.36	52.56	53.25	53.82	53.25	49.94	62.32
Intersection	60.21	63.67	54.04	54.67	55.49	59.12	41.44
Kulczynski	11.40	14.36	14.61	15.88	19.26	18.53	30.77
Kulczynskis	11.40	14.36	14.61	15.88	19.26	18.53	30.77
Lorentzian	**1.86**	**9.09**	**9.94**	**10.64**	**13.10**	**16.38**	**25.27**
Minkowski	23.99	18.78	19.07	22.89	21.40	20.49	31.99
Motyka	11.40	14.36	14.61	15.88	19.26	18.53	30.77
Ruzicka	11.40	14.36	14.61	15.88	19.26	18.53	30.77
Soergel	11.40	14.36	14.61	15.88	19.26	18.53	30.77
Sorensen	11.33	14.36	14.61	15.88	19.26	18.53	30.72
Wavehedges	11.40	14.36	14.61	15.88	19.26	18.53	30.77
Euclidean	22.82	16.00	18.88	19.07	22.38	20.20	29.11
Manhattan	12.53	13.48	15.85	16.16	20.04	19.24	25.68
ScaledManh	**8.87**	**11.77**	**9.75**	**11.90**	**14.74**	**17.61**	**15.45**
OulierCount	**2.33**	**9.85**	**9.97**	**13.42**	**12.15**	**16.30**	**17.12**
Mahalanobis	26.17	15.40	16.29	16.32	26.77	31.72	26.77
KMeans	18.40	15.21	13.19	13.89	16.98	17.96	16.62
SVM	18.18	15.06	11.30	14.49	16.24	16.82	17.16

6 Evaluation, Analysis and Comparison

In this subsection, comparisons on the basis of average EERs have been made between top most distance-based classification algorithms on different patterns of strings. Lorentzian classification algorithm achieved impressive results on 5 patterns of strings. It achieved only 1.86 % of average EER. Lorentzian algorithm represented in red colour in the following bar chart (Figs. 3 and 4).

In the following line chart, average EERs for different features-mining techniques for different patterns of strings have been represented to get the answer, which pattern of strings are suitable and which classification algorithms can be applied on keystroke database. Line in blue colour in Fig. 5 represents all strings and most of the time, it achieves optimum results. Twenty two different classification algorithms have been evaluated on different datasets but some of that algorithms achieved impressive results. In the following line chart, we can see that Lorentzian achieved 1.86 % of average EER, Outlier Count achieved 2.33 % and Canberra achieved 3.69 %.

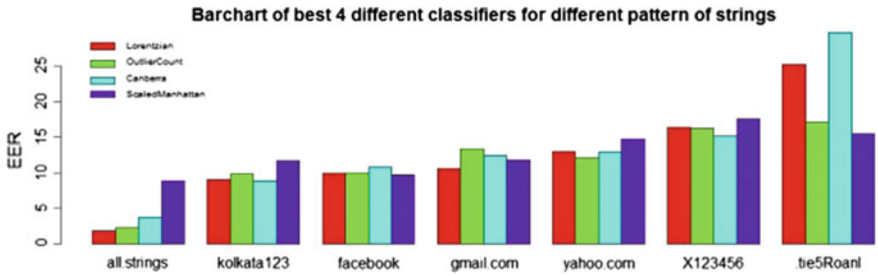

Fig. 3 Comparisons between top 4 classification algorithms on different patterns

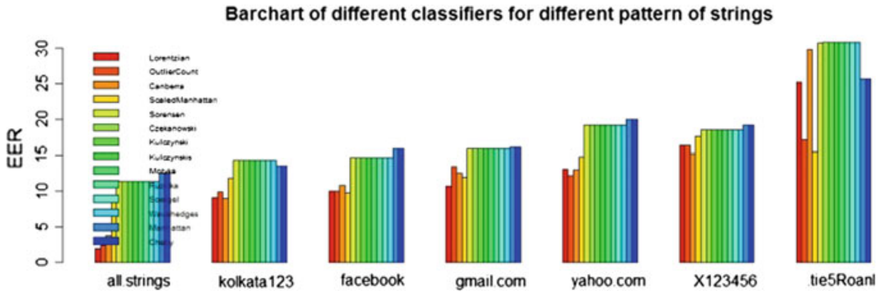

Fig. 4 Comparisons between top classification algorithms basis on average EER

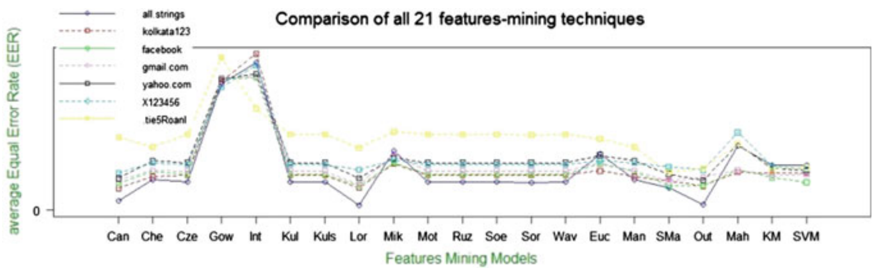

Fig. 5 Line chart of all 22 classifiers on different patterns

In the bellow figure, we have represented the EERs for different patterns of strings for all classification algorithms in histogram and we have seen except 2–4 classification algorithms most of the time we got impressive results (Fig. 6).

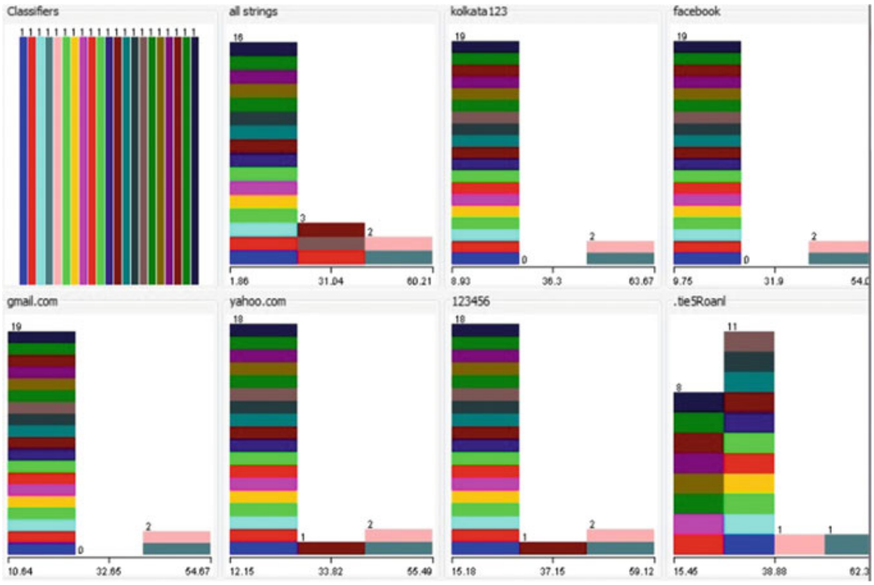

Fig. 6 Histogram of EERs of different patterns of strings

7 Conclusion

We have evaluated 22 different classification algorithms and we have seen most of the time obtained results are impressive. But Lorentzian classification algorithms achieved best average EER only 1.86 %, where Ourlier Count achieved only 2.33 %. In our evaluation process, we have worked on different databases on keystroke dynamics for verification but we have seen hard type of password type patterns are not suitable to train the system, we have to use some daily-used common words to train the system. We have also seen the pattern of string like “123456” are also not suitable in keystroke dynamics. Keystroke dynamics characteristics may change over time depending on mental state or muscle pain or tiredness of the person. Length of the string or PIN, which is used in authentication and type of that string, affect the way of regular typing rhythm. Position of the keyboard and type of key board also affect the way of typing style. Need much more experiment on it, like key pressure, finger placement etc. can also be calculated and considered to optimize the performance in accuracy of keystroke dynamics.

References

1. Roy, S., Roy, U., Sinha, D.D.: Enhanced knowledge-based user authentication technique via keystroke dynamics. Int. J. Eng. Sci. Invention (IJESI) **3**(9), 41–48 (2013)
2. Monrose, F., Rubin, A.D.: Keystroke dynamics as a biometric for authentication. Future Gener. Comput. Syst. **16**(4), 351–359 (2000)
3. Cha, S.H.: Comprehensive survey on distance/similarity measures between probability density functions. Int. J. Math. Models Methods Appl. Sci. **4**(1), 300–307 (2007)
4. Joyce, R., Gupta, G.: Identity authorization based on keystroke latencies. Commun. ACM **33** (2), 168–176 (1990)
5. Bleha, S., Slivinsky, C., Hussien, B.: Computer-access security systems using keystroke dynamics. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1217–1222 (1990)
6. Haider, S., Abbas, A., Zaidi, A.K.: A multi-technique approach for user identification through keystroke dynamics. In: IEEE International Conference on Systems, Man, and Cybernetics. vol. 2, pp. 1336–1341. Nashville, TN (2000)
7. Yu, E., Cho, S.: Novelty detection approach for keystroke dynamics identity verification. In: Intelligent Data Engineering and Automated Learning, vol. 2690, pp. 1016–1023. Springer, Berlin (2003)
8. Killourhy, K.S.: A scientific understanding of keystroke dynamics, PhD thesis, Computer Science Department, Carnegie Mellon University, Pittsburgh, US (2012)
9. Kang, P., Hwang, S.S., Cho, S.: Continual retraining of keystroke dynamics based authenticator. In: Advances in Biometrics, Proceedings, vol. 4642, pp. 1203–1211. Springer, Berlin (2007)
10. Giot, R., El-Abed, M., Rosenberger, C.: GREYC keystroke: a benchmark for keystroke dynamics biometric systems. In: Proceedings of the IEEE 3rd International Conference on Biometrics: Theory, Applications and Systems (BTAS '09), pp. 1–6 (2009)