


Keystroke Biometric Systems for User Authentication

Md Liakat Ali¹  · John V. Monaco¹ · Charles C. Tappert¹ · Meikang Qiu¹

Received: 1 October 2015 / Revised: 22 December 2015 / Accepted: 14 February 2016
© Springer Science+Business Media New York 2016

Abstract Keystroke biometrics (KB) authentication systems are a less popular form of access control, although they are gaining popularity. In recent years, keystroke biometric authentication has been an active area of research due to its low cost and ease of integration with existing security systems. Various researchers have used different methods and algorithms for data collection, feature representation, classification, and performance evaluation to measure the accuracy of the system, and therefore achieved different accuracy rates. Although recently, the support vector machine is most widely used by researchers, it seems that ensemble methods and artificial neural networks yield higher accuracy. Moreover, the overall accuracy of KB is still lower than other biometric authentication systems, such as *iris*. The objective of this paper is to present a detailed survey of the most recent researches on keystroke dynamic authentication, the methods and algorithms used, the accuracy rate, and the shortcomings of those researches. Finally, the paper identifies some issues that need to be addressed in designing keystroke dynamic biometric systems, makes

suggestions to improve the accuracy rate of KB systems, and proposes some possible future research directions.

Keywords Authentication · Identification · Keystroke biometrics · Classification · Behavioral biometrics

1 Introduction

User authentication is one of the most important and challenging aspects of controlling unauthorized access to a system. Authentication is the process by which the system verifies the user has a legitimate claim to access the system. There are three traditional modes used for authentication of a person, including possessions, knowledge, and biometrics. Figure 1 shows an ontology of various authentication modes, including biometric modalities. A *possession* is typically a unique physical item, such as a key, passport, or smartcard. These can be shared, duplicated, or even may be lost or stolen. *Knowledge* is some secret information, such as a password. Although widely used, many passwords are easy to guess, shared with others, or may be forgotten. *Biometrics* measures a unique human characteristic or trait. Biometric modalities can be classified as physiological and behavioral. Physiological modalities are related to the shape of human body such as fingerprint, face, DNA, hand geometry, iris, retina, ear shape, odor, and skin reflectance. Behavioral modalities are related to the behavior pattern of a person, such as signature, gait, lip motion, voice, keystroke, mouse movement, and stylometry. Biometric characteristics are generally difficult to reproduce and cannot be lost or forgotten [52].

Password-based authentication is the most common and widely used methods to protect data from intruders. Many

✉ Md Liakat Ali
ma03901n@pace.edu

John V. Monaco
jmonaco@pace.edu

Charles C. Tappert
ctappert@pace.edu

Meikang Qiu
mqiu@pace.edu

¹ Seidenberg School of CSIS, Pace University, 861 Bedford Road, Pleasantville, NY 10570, USA

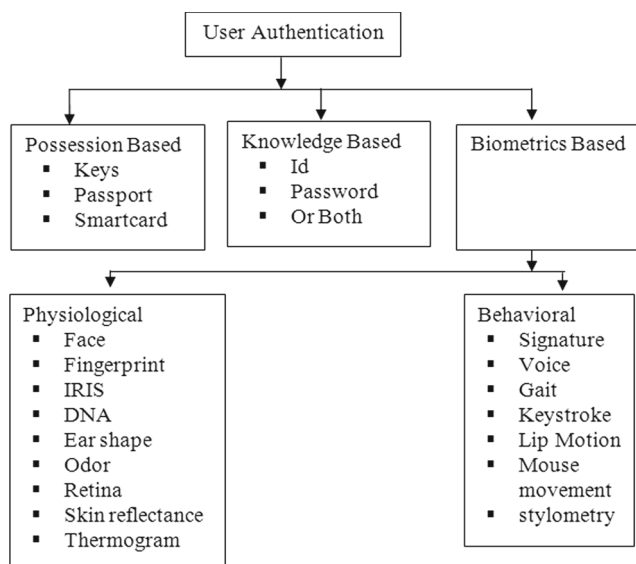


Figure 1 Ontology of authentication modes.

people choose their passwords using information from daily life, such as birth date, social security number, or a pet's name, which may be susceptible to dictionary attacks. Although some rules are designed to increase password entropy, such as combination of one or two capital letters, special character(s), and digit(s), these typically create an additional burden for the user. Additional mechanisms are needed to enhance the security and convenience of password-based authentication.

Like written signatures, user typing patterns have neuro-physiological factors that make them unique from others. A unique keystroke profile can be constructed from various typing features, such as typing speed, the duration between successive keys pressed, pressure applied on the keys, and finger positions on the keys. Recognition based on the unique typing pattern is non-intrusive, cost-efficient [63, 64, 93, 94, 130], and transparent to the user [81]. Moreover, it is very easy to capture data as keyboards are common and no special equipment is necessary. As the user types, an authentication decision can be made after each keystroke to provide continuous authentication [16, 18]. Ingo Deutschmann et al. [28] have tested a continuous authentication system on 99 users over a period of 10 weeks and determined that keystroke biometrics is appropriate for continuous authentication.

KB has a number of advantages, such as being low cost, transparent to the user, and non-invasive. However, the main disadvantage of KB is low accuracy when compared to other biometric systems [2–4, 76]. An ongoing challenge in KB authentication is to achieve greater accuracy. This work examines this effort by comparing the performances, data collection techniques, feature extraction, and classification algorithms of existing KB systems. By understanding the

performance and limitations of current systems, this work then proposes some guidelines to enhance the accuracy and performance of the existing KB biometric systems. The main objectives and contributions of this paper are listed below:

- Provide a meta-survey of earlier survey papers on KB systems.
- Offer a comprehensive survey of KB research efforts. Attributes include accuracy, features, hardware, training time, classification time, and memory usage.
- Most importantly, identify the shortcomings of existing KB systems and provide some direction for future research.

The structure of this paper is as follows. Section 2 provides a meta-survey of KB authentication systems. Section 3 describes the typical KB authentication system, including system components and evaluation criteria. Section 4 includes a comprehensive survey of KB systems and compares different methods used, as well as limitations. Finally, Section 5 concludes this work with recommendations and future research directions.

2 Meta-survey

KB authentication is quite and remains an active area of research area compared to other biometric systems. A limited number of researches have been conducted by the researchers to analyze keystroke authentication systems, and most of them utilize desktop or laptop computers [44, 122], with few studies conducted on ATM [42], mobile phone [21, 73], and even fewer on smartphones with touchscreen [31, 35, 49, 108]. Some of the researches have used timing features [46], some have used extra features such as pressure [71, 105], finger area [125] or a combination of [8, 49]. Few researches were conducted on free long text, [122] with most of the researches utilizing short input [52]. Stewart et al. [117] have used stylometry as a keystroke feature in an effort to develop a robust online examination authentication system. A limited number of researches have developed web based authentication [75, 112, 117, 131].

In mid 2004, A. Peacock et al. [89] surveyed KB researches. They have evaluated the previous researches based on classifier accuracy, usability, and other factors. They also discussed privacy and security issues that come with KB systems. It was found that performance of KB systems is greatly influenced by the number of training samples. They also recommended the creation of a public data set, introduction of schemes that ensure privacy of collected data, and data expiration under certain conditions.

D. Shanmugapriya et al. [110] conducted a survey in 2009 and discussed the performance evaluation of previous

researches. They also summarized different approaches, security, and challenges in KB systems. Another survey of KB was conducted by H. Crawford [24] in 2010. The author reviewed a representative subset of concurrent researches in KB and provided recommendations for future work. The survey found that high quality results from different researches were seen with neural network pattern classification. Although statistical classifiers are less computationally intensive, they do not provide a strong level of classification. The study also recommends not to use same participants for both authorized and unauthorized populations in system evaluation.

An extensive work surveying KB research was conducted by S. P. Banerjee [12] in 2012. The survey compared different researches based on different algorithms used, discussed explicitly the factors that affect system performance, and finally gave their recommendations. The study concluded that KB authentication systems have potential to grow in the area of cybersecurity and biometric monitoring since it is both non-intrusive and cost-effective.

A similar kind of research on surveying KB research was conducted by P. S. Teh et al. [124] in 2013, where the authors claim that KB biometric is unlikely to replace the existing authentication system entirely and cannot be a sole biometric authenticator. However, some properties, such as the ability to operate in stealth mode, low implementation cost, high user acceptance, and ease of integration with existing security systems, makes KB authentication promising and can play a significant rule in enhancing overall system security. Similar studies in KB authentication include A. Alsultan et al. [6], and M. Kanimozhi et al. [56], which vary by the length and depth of the review, the number of research papers covered, the way that the resources were presented, and their recommendations. In [2], Ali et al. have compared KB research those have used Hidden Markov Model (HMM) and in [4], the authors have compared several KB research those have used Neural Networks as classification method.

The following section describes the overview of a KB system—how data was collected, what features was extracted, what are the different classification methods were used by the researchers, and how performance of different experiments were evaluated in various KB authentication systems.

3 Keystroke Biometric System Overview

The KB system involves measuring and assessing a person's typing rhythm on some digital devices such as a mechanical keyboard or mobile device touchscreen, and creating a unique signature to identify the legitimate user [86]. According to [80], KB refers to the way a person types and

not what is typed. Although according to H. Crawford [24], the first studies on desktop keyboards were conducted by R. Spillance [116] in 1975. Gaines et al. [34] have studied the habitual patterns in user typing behavior for identification, however the idea of keystroke pattern for user authentication dates back to World War II. During this time, Morse code operators on a telegraph machine were identified by the rhythm, pace, and syncopation of the taps [58, 89]. Figure 2 shows the basis components of the typical KB authentication system. A typical KB authentication system mainly consists of six components including data collection, feature extraction, feature classification/matching, decision making, retraining, and evaluation. Some of the research have all of those six components while some have not considered decision making and/or retraining the system.

3.1 Data Collection

Data collection is the first step in the KB system, where raw data is collected via various input devices. Data acquisition varies in different systems, ranging from normal keyboard [75, 80] to pressure sensitive keyboards [68, 85]. The common choice for the mechanical keyboard is the *QWERTY* keyboard and built-in laptop keyboard [88]. References [118] and [23] have modified existing devices to sense the pressure applied to each key. Other experiments also have used a special purpose number-pad [42, 62], cellular phone [21, 103] and smart phone with touchscreen [8, 49, 74, 126]. Due to the limited number of publicly available standard data sets [5, 32, 38, 41, 60, 128], many researchers have generated their own data set. For example, there were 1254 participants involved in S. Douhou et al. [30] experiment, 118 participants in research [127], and only 3 participants were involved in I. V. McLoughlin [73]. In most works, 10–20 participants were involved [124]. Text input can be divided into two groups: short and long input.

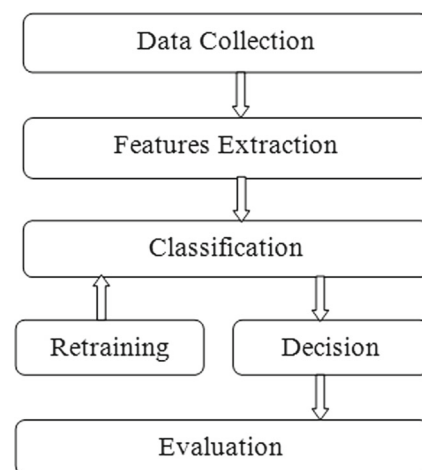


Figure 2 Overview of general KB Authentication System [3].

Table 1 Comparison of publicly available KB databases.

Database	Feature	Data size	Input	Scope	Keyboard layout	Text type
Jugurta et al. [33]	Timing	32	–	–	Brazilian	Static, dynamic
Giot et al. [38]	Timing	133	greyc laboratory	Genuine	AZERTY	Static
Killourhy et al. [60]	Timing, pressure	51	.tie5Roanl	Genuine	U.S.	Static
Allen et al. [5]	Timing	104	Jeffrey Allen drizzle	Genuine, imposter	U.S.	Static
Bello et al. [13]	Timing	54	–	–	–	Static
Idrus et al. [47]	Timing	110	–	–	AZERTY and QWERTY	Dynamic

Short input can be a username [87], password [66], or text phrase [14], while long input can be a paragraph [134]. Most of the experiments have used character-based texts, but some have used purely numerical inputs [72], and others have used mixed input [49]. Only a limited number of researches had free-text input, and in many researches data was collected in one session. Table 1 compares some of the publicly available KB databases.

3.2 Feature Extraction

After raw data is collected, it needs to be processed, normalized, and stored for classification. Several feature extraction methods have been used in KB research [100, 115, 120]. The most common features are timing measurements. When a key is pressed, it creates a hardware interrupt to the processor and generates a time stamp, typically measured with microsecond precision. Using the time stamps, the durations of and intervals between keystrokes can be calculated. The timing features applied in various researches can be classified into two groups: *di-graph* and *n-graph*.

Di-graph is the timing information between two consecutive keystrokes [25, 112] and can be grouped into two classes, including *dwelt time (DT)* and *flight time (FT)*. Figure 3 shows the comparison between *dwelt time* and *flight time*. *dwelt time* corresponds to the hold time, which is the time interval between key press and key release and *flight time* is the time interval between releasing one key and pressing the next key. The keystroke latency is the combination of hold and flight time. Key Press Latency (KPL) is the time interval between two consecutive keys press and Key Release Latency (KRL) is the time interval between two consecutive keys release.

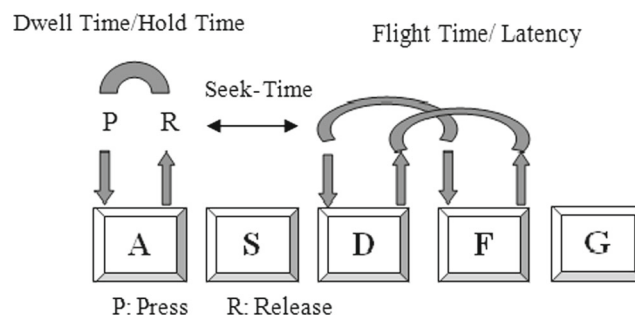
N-graph is the timing information from three or more consecutive keystrokes. It is usually taken as the time elapsed between a key and the *n*th key while typing. Most of the researches have used two consecutive key time events (*di-graph*).

Other spatial features extracted from keystroke data include: pressure, position on the touchscreen, length and orientation of major and minor axes of finger-press area on

the touch screen, frequency of word error, keystroke sound [101], typing rate, and text correction features [17].

3.3 Classification

Classification is a critical part of any KB system. During the *classification phase*, extracted features are categorized to make decision. Numerous pattern recognition approaches, such as statistical and machine-learning methods, have been proposed with the common goal of increasing accuracy rate. A survey of KB literature suggests that in earlier time, most of the classification methods were statistical approaches; however, in modern times, researchers tend to concentrate on machine-learning approaches. At present, Support Vector Machine (SVM) is the most popular classification method used by several researchers due to good out-of-box performance with reasonable computational complexity. Despite this, large feature sets used in conjunction with SVM tend to overfit and reduce performance. A combination of different machine-learning approaches, especially Artificial Neural Network (ANN) with evolutionary computing, Two-Class classifier [61, 90] have been met with some success. Statistical approaches used by researchers in keystroke biometrics include: mean, median, standard deviation measures, statistical t-test for similarity, and Bayesian modeling. Machine learning approaches include ANNs [29], decision trees [53], fuzzy logic [78], evolutionary computing, and SVM. Common ANN architectures are the multi-layer perceptron

**Figure 3** Dwell time and flight time.

(MLP) [91], radial basis function network (RBFN), learning vector quantization (LVQ), and self-organizing map (SOM). Evolutionary computing techniques include genetic algorithms (GA), particle swarm optimization (PSO), and colony optimization. Taculin et al. [121] have explored an algorithmic solution based on chemical reactions of molecules.

3.4 Decision

During the decision phase, the classifier outputs are aggregated and compared to a decision threshold. In authentication, the template of the claimant is compared to one or more reference templates, all of which may result in classifier output scores. The final decision is a binary value: either accept or reject the claimant [12, 124].

3.5 Retraining

During the retraining phase, the reference template of a user is updated to reflect a change in environment or behavior. Although most researchers do not explicitly consider retraining, some researchers, such as D. Hosseinzadeh et al. [46], have proposed algorithms that constantly renew the reference template. This is necessary because a user's typing pattern may change with time and environment. Other proposed retraining techniques include a growing window, moving window, retraining with impostor patterns, and adaptive thresholds [124].

3.6 Evaluation

Biometric systems mainly have two functions: *verification* and *identification*. In the *verification* stage, the system will accept or reject the identity claimed by the user. *Identification*, also called as recognition, is where a system classifies the input pattern into one of the N known classes [8]. The performance of the biometric verification system is usually characterized by its receiver operating characteristic (ROC) curve. The ROC quantifies the tradeoff between the false acceptance rate (FAR) and false rejection rate (FRR). Figure 4 shows the relationship among FAR, FRR, and EER. FAR (also referred to as false match rate [36] or type 2 error) is the rate at which the system incorrectly accepts a sample that is provided by impostor. A system with lower FAR indicates an impostors is less likely to be accepted. FRR (also referred to as false non-match rate or type 1 error) is the rate at which the system rejects a sample provided by a genuine user. Smaller FRR indicates a small number of genuine users rejected. Overall system performance is described by the equal error rate (EER), the point at which FAR and FRR are equal. The EER is sometimes referred as the *crossover error rate* (CER). The lower EER indicates

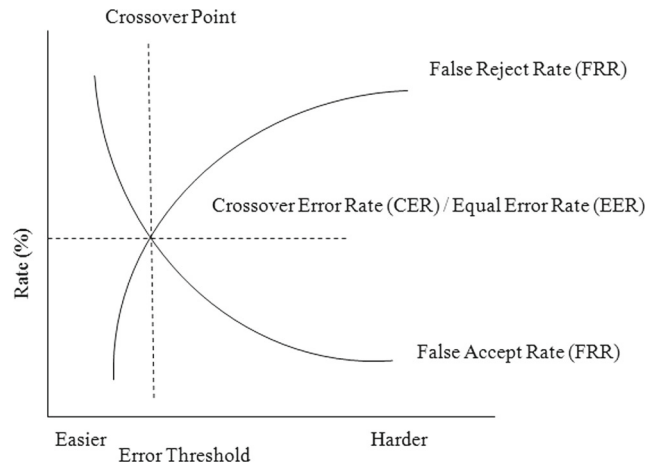


Figure 4 Relationship between FAR, FRR, and EER.

better performance. Other system evaluation criteria include efficiency, adaptability, robustness, and convenience.

Advantages and Drawbacks There are several advantages of KB systems. Typing patterns are believed to be unique to an individual. KB systems generally have lower implementation and deployment costs compare to other biometric authentication systems. The typically KB system can be fully implemented by software and has low dependency on specialized hardware. From the user's perspective, KB systems are transparent and non-invasive. They offer increased password strength and lifespan and continuous monitoring [79, 124]. The main disadvantage of KB authentication systems is lower accuracy compared to other biometrics authentication systems. There various factors that contribute to lower accuracy rate, such as variations in the typing rhythm, template aging, user injury, and environmental factors.

The following section compares different authentication and identification methods used in different research. It also compares different algorithms used in KB systems based on the training time, classification time and memory usage by the different algorithms. Finally it shows the limitations of existing KB systems.

4 Comparison of Research Works Based on Different Classification and Verification Methods

Statistical Approaches Table 2 shows the works that have used statistical methods and distance-based classifiers for authentication and identification. Table 3 shows the works that have used other statistical methods, including ensemble methods in which several algorithms are combined to achieve greater accuracy.

Table 2 Classification based on most common statistical approaches.

Paper	Year	Participants	Features	Input freedom	Input type	Device	Result%
Mean and STD							
[52]	1990	33	FT	Yes	Short	MK	FAR: 0.25, FRR:16.36
[80]	1997	31	DT, FT	Yes	–	MK	Accuracy: 90
[43]	2005	205	FT	Yes	Long	MK	FAR: 0.5, FRR: 5
[21]	2009	30	–	–	Text	MP	EER: 13
[30]	2009	1254	DT, FT	No	Short	MK	FAR: 16, FRR: 1
[73]	2009	3	–	–	Digit	MP	Accuracy: 90
[135]	2012	51	–	No	Short	MK	EER: 8.4
[125]	2013	152	FT, DT, P, FA	–	Digit	MP	FAR: 4.19, FRR: 4.59
KNN							
[81]	2000	63	DT, FT	Yes	–	MK	Accuracy: 83.2–92.1
[48]	2002	7	DT, FT	–	Digit	NP	EER: 78–99
[104]	2008	10	P	–	Digit	TS	EER: 1.00
[133]	2010	120	DT, FT	–	–	MK	EER: 1.00
[122]	2010	100	DT, FT	Yes	Text	MK	EER:2.7
[117]	2011	30	DT,FT	No	Digit	–	EER: 0.5
[77]	2013	40	–	–	–	MK	Accuracy: 88.2–91.5
[102]	2013	–	DT	Yes	Long	MK	EER: 6.1
Euclidian distance							
[127]	2006	118	DT, FT	Yes	Long	–	Accuracy: 97.9
[55]	2007	21	DT, FT	Yes	Short	MK	EER: 3.8
[42]	2008	30	DT, FT, P	–	Digit	NP	FAR: 15, FRR: 0, EER: 10
[106]	2009	112	DT, FT	No	Long	MK	Accuracy: 100
[39]	2009	16	DT,FT	No	Short	MK	EER: 4.28
[122]	2010	100	DT,FT	Yes	Text	MK	EER: 2.7
[107]	2011	189	DT, FT	No	Long	MK	FAR: 0.01, FRR: 3
[54]	2011	51	FT	No	Long	MK	EER: 0.84
[114]	2011	20	FT	No	Long	MK	FAR: 2, FRR: 4
[77]	2013	40	–	–	–	MK	Accuracy: 88.2–91.5
Degree of disorder method							
[44]	2005	31	FT	Yes	Long	MK	FAR: 1.99, FRR: 2.42
[43]	2005	205	FT	Yes	Long	MK	FAR: 0.5, FRR: 5
[26]	2009	21	–	Yes	Long	–	FAR:0.14, FRR: 1.59
[95]	2011	50	FT	No	Long	MK	EER: 10
[131]	2011	186	–	Yes	Long	–	FAR: 1.65, FRR: 2.75
[51]	2012	42	DT, FT, P	–	Digit	MK, TS	Better: MK (timing feature)

FT - Flight Time, DT - Dwell Time, P - Pressure, FA - Finger Area, FAR - False Acceptance Rate, FRR - False Reject Rate, EER - Equal Error Rate, TPR - True Positive Rate, FPR - False Positive Rate, MK - Mechanical Keyboard, MP - Mobile Platform, TS - Touchscreen, NP - Number Pad

Gunetti et al.'s experiment focused on using the degree of disorder [44] and combination of degree of disorder, mean, and standard deviation [43]. By combining mean and standard deviation, they achieved lower FAR but higher FRR. Kang et al. [55] have used a combination of k-means and Euclidian distance and achieved 3.8 % EER. Giot et al. [39] have used Bayesian decision theory and Euclidean distance and achieved 4.28 % EER. In another work, they achieved

EER 6.96 % using Bayesian, Euclidean, and Hamming distance [36].

A comparison between KB on personal computers and smartphones was performed by Johansen [51]. Their experiments utilized data collected from 42 participants on both standard keyboards and smartphones. By using the degree of disorder, the study concluded that the performance on smartphones is less than standard keyboard if only timing

Table 3 Mixed and other statistical methods used in different researches.

Paper	Year	Participants	Features	Input freedom	Input type	Method	Device	Result%
[34]	1980	7	FT	No	Long	t-Test	MK	Accuracy: 95
[14]	1990	26	FT	Yes	Short	Baysian, Minimum Distance classifier	MK	FAR: 2.8, FRR: 8.1
[46]	2004	41	DT, FT	Yes	Short	Gaussian mixture modeling	MK	FAR: 4.3, FRR 4.8, EER: 4.4
[62]	2005	9	DT, FT, P	–	Digit	ANOVA	MK	EER: 2.4
[69]	2006	100	DT, FT, P	–	–	Dynamic time warping	MK	FAR, FRR, EER: 1.4
[99]	2006	20	–	–	Digit	Hidden Markov model	MK	EER: 3.6
[20]	2006	20	–	–	Digit	Euclidian, Mahalanobis, FF-MLP	TS	FAR: 0, FRR: 2.5
[103]	2009	25	DT, FT	Yes	Short	Gauss, Parzen, K-NN, K-mean	MK	EER: 1.00
[40]	2009	100	DT, FT	No	Short	Bayesian, Euclidean, Hamming distance	MK	EER: 6.96
[59]	2010	51	DT, FT	No	Short	Manhattan distance	MK	EER: 7.1
[119]	2010	35	FT	No	Long	Kolmogorov-Smirnov	MK	EER: 7.55
[123]	2011	100	DT, FT	No	Short	Gaussian PDF, Direction similarity measure	MK	EER: 1.401
[75]	2011	55	–	Yes	Long	Spearmans's foot rule distance metric	–	FAR: 2.02, FRR: 1.84
[11]	2011	33	DT, FT	No	Long	Naive Bayesian	MK	EER: 1.72
[125]	2013	152	FT, DT, P, FA	–	Digits	k-mean	MP	FAR: 4.19 FRR: 4.59
[50]	2013	10	DT, FT, FA	Yes	Short	Bayesian	TS	FAR: 0.02 FRR: 0.178
[17]	2014	30	DT, FT	No	Digit	SMD, SED	MK	EER: 26

FT - Flight Time, DT - Dwell Time, P - Pressure, FA - Finger Area, FAR - False Acceptance Rate, FRR - False Reject Rate, EER - Equal Error Rate, TPR - True Positive Rate, FPR - False Positive Rate, MK - Mechanical Keyboard, MP - Mobile Platform, TS - Touchscreen, SMD- Scaled Manhattan Distance, SED- Scaled Euclidean Distance

features are used. Moreover, if additional smartphone sensors are used, performance increases over the standard keyboard. The data consisted of a numerical password on a 12-key mobile phone keyboard. Trojahn et al. [125] performed a similar experiment with 152 participants and a 17-digit password. In a single session, each participant typed the password ten times during a single session. They determined that the pressure and size of the finger, in addition to timing features, can reduce the error rate of a mobile KB system. The study found 4.19 % FAR, and 4.59 % FRR by using hold time combined with *di-graph*, and *tri-graph* features.

Machine Learning Approaches Table 4 shows the comparison between various works that have utilized machine learning approaches, such as random forest decision trees, SVM, and ANN. Table 5 shows works that have used other methods in their experiments. Some researchers have implemented new approaches and combined several algorithms together for better accuracy.

N. Harun et al. [45] have discussed various issues to enhance the security of password authentication schemes by using KB for user authentication. The authors have used

the time interval between keystrokes as input to a multilayer perceptron neural network (MLP-NN) trained by back propagation (BP). Equalization histogram and principal component analysis (PCA) were used for preprocessing. This resulted in greater than 80 % correctly classified users.

M. Antal et al. [8] have examined keystroke data with and without touchscreen features, such as pressure and finger area. Data was collected from mobile Android devices, and both time and touchscreen-based features were extracted. Classification accuracies were obtained by several machine learning classification algorithms, including Nave Bayes, Bayesian Network, C4.5(J48), K-NN(IBk), SVM, random forest, and MLP. The best performances were obtain by random forest, Bayesian nets, and SVM, respectively. The addition of touchscreen-based features to the timing feature set resulted in an increase of over 10 % accuracy for each classifier. Verification results were obtain using Euclidean, Manhattan, and Mahalanobis distance metrics. The study showed that the lowest error (12.9 %) was obtained by the Manhattan distance using both timing and touchscreen-based features. The research concluded that touchscreen-based features increase classification and verification accuracy.

Table 4 Classification based on most common machine-learning approaches.

Paper	Year	Participants	Features	Input freedom	Input type	Device	Result%
Random forest decision tree							
[85]	2004	41	DT, FT	No	Short	MK	EER: 2
[112]	2010	10	FT, DT	–	–	MK	FAR: 0.41, FRR: 0.63, EER: 0.53
[113]	2010	21	DT, FT	Yes	Long	MK	FAR:3.47, FRR:0, EER: 1.73
[72]	2010	28	DT, FT	–	Digit	MK	FAR: 0.03, FRR: 1.51, EER:1
ANN							
[15]	1993	24	FT	Yes	Short	MK	FAR: 8, FRR: 9
[86]	1995	15	DT, FT	Yes	Short	MK	EER: 0
[22]	1999	10	DT, FT	–	Short	MK	Accuracy: 97
[70]	1999	–	FT	–	Short	MK	MLP (AR: 84, RR: 69), k-mean (AR: 85, RR: 85)
[48]	2002	7	DT, FT	–	Digit	NP	Accuracy: 78–99
[132]	2003	21	DT, FT	–	Short	MK	FAR: 0, FRR: 0.814
[88]	2007	100	DT, FT	No	Short	MK	FAR:1, EER: 8
[67]	2007	100	DT, FT, P	–	–	MK	EER:11.78
[23]	2007	32	FT	Yes	Long	MP	EER: 12.8
[105]	2009	10	P	–	Digit	TS	EER: 1
[118]	2009	30	DT, FT,P	–	–	MK	FAR: 2
[1]	2009	7	P	–	–	MK	FAR:0, FRR: 0
[57]	2010	25	FT, DT	–	Digit	MK	Accuracy: 92.8
[84]	2010	20	DT, FT, P	–	–	MK	FAR: 4.12, FRR: 5.55
SVM							
[132]	2003	21	DT, FT	Yes	Short	MK	FAR:0, FRR: 0.814
[9]	2007	24	DT, FT	Yes	Long	MK	FAR: 0.76, FRR: 0.81, EER: 1.57
[92]	2007	61	–	No	Short	–	FAR: 14.5, FRR: 1.78
[71]	2007	5	P	–	Digit	MK	FAR: 0.95,FRR: 5.6
[83]	2011	24	–	–	Digit	MK	EER:2
[65]	2011	117	DT, FT	Yes	Short	MK	EER: 11.83
[37]	2011	100	DT, FT	No	Short	MK	EER: 15.28
[35]	2014	300	FT, DT	–	Short	TS	TPR = 92, FPR = 1
[49]	2014	30	FT, DT, P, FA	No	Digit	MK, TS	EER: 10.5(MK data), EER: 2.8 (TS data)

FT - Flight Time, DT - Dwell Time, P - Pressure, FA - Finger Area, FAR - False Acceptance Rate, FRR - False Reject Rate, EER - Equal Error Rate, TPR - True Positive Rate, FPR - False Positive Rate, MK - Mechanical Keyboard, MP - Mobile Platform, TS - Touchscreen, NP - Number Pad, AR - Acceptance Rate, RR - Rejection Rate

Similar to M. Antal et al., L. Jain et al. [49] have studied the timing and sensor features on a mobile device. Data was collected from 30 participants over several days, and a one-class SVM was used to obtain results. The experiment yielded a 10.5 % EER using timing features only, 3.5 % EER using non-timing touchscreen features, and 2.8 % EER with combined timing and touchscreen features. The results suggest that performance on mobile touchscreen devices is superior to that on hardware keyboards due to additional sensors.

M. Brown et al. [19] have developed a system to identify users via keystroke by using simple MLP-NN with BP. D. T. Lin [66] worked on the M. Brown system, modifying the architecture and parameters of the neural network.

He achieved a 1.1 % FAR, and 0 % FRR. N. Capuano et al. [22] proposed an authentication system that used MLP with radial basis function, and found a 0 % FAR and 3 % FRR. M. S. Obaidat et al. [87] proposed a technique to verify the identity of legitimate users using KB with several neural network and pattern recognition algorithms. The study showed that the combination of hold times and flight times as features, with fuzzy ARTMAP, radial basis function networks and learning vector quantization neural network paradigms, achieved a 5.8 % EER.

L. K. Maisuria et al. [70] compared keystroke classification performance using MLP and k-means cluster algorithm and found 16 % and 15 % FRR, and 31 % and 15 % FAR, respectively. Sulong et al. [118] have proposed a system

Table 5 Other less common machine-learning approaches used in different researches.

Paper	Year	Participants	Features	Input freedom	Input	Method	Device	Result%
[98]	1998	10	DT, FT	Yes	Short	Inductive learning classifier	MK	FAR: 9, EER: 10
[111]	2005	43	DT, FT	No	Long	Decision trees, Monte Carlo	MK	FAR: 0.88, FRR: 9.62
[68]	2005	–	DT, FT, P	–	–	Fuzzy ARTMAP	MK	FAR: 0.87, FRR: 4.4
[96]	2005	100	DT	No	Short	Genetic Algorithm	MK	Accuracy: 95
[134]	2006	–	DT, FT	No	Long	Decision tree c4.5.j48	MK	Accuracy: 93.3
[97]	2007	30	DT, FT	No	Short	Sequence alignment algorithms	MK	FAR: 0.2, FRR: 0.2, EER: 0.4
[10]	2013	30	–	Yes	Long	Dichotomy classifier	–	EER: 8.7 (Pass), 6.1 (Num)
[129]	2014	–	DT, FT	–	Short	–	MK	FAR: 4.8, FRR: 3.1, EER: 5
[108]	2014	10	P	–	Digits	–	TS	EER: 15.2
[31]	2014	13	FT, DT, P, FA	–	–	–	TS	FAR: 14 FRR: 2.2
[8]	2014	42	FT, DT, P, FA	No	Short	Naive, Bayesian, C4.5(J48), KNN, SVM, MLP, Random forest	MP	EER: 12.9
[7]	2015	42	FT, DT, P, FA	No	Short	Two-class	MP	EER: 3

FT - Flight Time, DT - Dwell Time, P - Pressure, FA - Finger Area, FAR - False Acceptance Rate, FRR - False Reject Rate, EER - Equal Error Rate, TPR - True Positive Rate, FPR - False Positive Rate, MK - Mechanical Keyboard, MP - Mobile Platform, TS - Touchscreen

combining maximum pressured applied on the keyboard and latency between keystrokes as input to a radial basis function network. They achieved 100 % classification rate with 22.4 s average training time. Based on FRR and FAR, the authors claimed that the proposed system is effective for biometric-based security systems. Robert S. Zack et al. [133] developed a long-text input keystroke biometric system that consists of three components: raw keystroke data collection over the Internet, a feature extractor, and a pattern classifier. The system was tested with 120 participants and achieved approximately 1 % EER. The system showed higher performance with a closed system of known users than an open system, as well as performance variations with the number of enrolled users.

A similar experiment was conducted by R. A. Maxion et al. [72], where 28 users typed the same 10 digit number using only the right-hand index finger. The authors used a random forest classifier and have achieved a 10 % EER. H. Saevanee et al. [105] studied timing features combined with finger pressure and used notebooks with touchpads as a touchscreen. Data was collected from 10 users, who entered their 10-digit cell phone number. The experiment yielded 99 % accuracy using the finger pressure features. A limitation of this approach is lack of impostor data due to each participant having a different phone number. In this case, only FRR was measured.

B. Draffin et al. [31] performed experiments utilizing soft keyboard data collected from 13 participants over 3 weeks. The study used key-press duration, finger area, drift, pressure, and keyboard orientation as features, and achieved a 14 % FAR and 2.2 % FRR. A similar study performed by S. Sen et al. [108] used pressure as a feature, with 4-digit

input from 10 participants. The study presented verification results based on a special impostor mode in addition to the typical performance measures.

Summary of Classification Algorithms Based on the review of KB research, there are several algorithms such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO) [109] are most commonly used algorithm in KB systems. Killourhy and Maxion [60] performance study of the fourteen existing keystroke dynamics algorithm shows that Manhattan distance outperform other algorithms with an equal error rate of 0.096 and nearest neighbor classifier using Mahalanobis distance with an equal error rate of 0.10 [135]. Table 6 summarizes the most common algorithms used in KB systems. This table also depicts the classification time, training time, and the memory usage of each algorithm. Most works do not consider the time and space requirements of the classification methods. In practice, the problem constraints may limit which methods can be used. Y. Deng et al. [27] have introduced two new algorithms: Gaussian mixture model with universal background model (GMM-UBM) and deep belief nets (DBN). These two approaches leverage data from background users and enhance the model's discriminative capability without using impostor's data at training time. The authors claimed that these two new algorithms make no assumption about underlying probability distribution and are fast for training and testing.

Limitations In general, the shortcomings of existing KB systems can be summarized as:

Table 6 Comparison of different algorithms based on the training time, the classification time and memory usage.

Algorithm	Type	Training time	Classification time	Memory usage
Linear regression	Parametric	Fast	Very fast	Very low
Logistic regression	Parametric	Medium	Very fast	Very low
ANN (back propagation)	Parametric	Slow	Very fast	Low
RBF ANN	Parametric	Medium	Medium	Medium
KNN	Nonparametric	Slow	High	
Gaussian mixture	Parametric	Medium/slow	Medium	Medium
Kk-means clustering	Nonparametric	Medium	Fast/medium	Medium
Estimate-maximized clustering	Nonparametric	Medium	Medium	Medium
Learning vector quantization (LVQ)	Nonparametric	Slow	Medium	Medium
Group method of data handling (GMDH)	Nonparametric	Fast/medium	Fast	Low
Genetic algorithm (GA)	Nonparametric	Slow	Medium	–
Particle Swarm Optimization (PSO)	Nonparametric	Fast	Fast	–

- There is currently not so much emphasis on real-world conditions and many studies have been conducted under laboratory settings.
- To date, there have been no studies that quantify KB system performance for users with low typing proficiency. Almost all works have utilized data collected from participants in an academic environment.
- Touchscreen features, such as pressure and finger area, have not been studied across different devices and platforms.
- In most studies, data was collected in only one session and ignore template aging. Longitudinal studies are needed to determine KB system performance over time.
- The number of standard and publicly available KB datasets remains limited compared to other disciplines.
- Many studies utilize the same keyboard for enrollment and testing.
- In some works, the justification for a particular classification method remains low.
- There is generally very little discussion of negative results (i.e. random-choice accuracy). Non-significant results should be reported in some way, such as through a journal that aims to publish such results of well-designed experiments.

5 Discussion

The current KB literature suggests that most of the experimental subjects involved in the research were from institutes, students, academicians or supporting staffs. As KB has a strong psychological basis, this population is not able to represent the greater population. A deeper understanding of typing behavior of people from different ages, genders, and labels of society may enhance the accuracy of existing keystroke biometric systems.

Classification and feature extraction algorithms may continue to be improved to achieve greater accuracy of KB systems. Some works have combined several methods to obtain better results, while others have used platform-dependent features to increase accuracy. Feature engineering remains a viable direction for future work to improve KB system performance. Additionally, advances in feature selection and feature fusion methods can increase the quality of KB systems.

Many researchers have not paid attention to the time and space required by an algorithm for training and testing. The reason behind this may be due to the dramatic increase in processor speed and memory size. Despite this, there are practical constraints that limit the choice of classification algorithm, such as in embedded systems.

Several studies have utilized long text input, although they have used only English as the primary language of communication. KB system performance with non-English input and non-US keyboard layouts remains largely untested.

With the increasing popularity of mobile devices, KB performance on mobile platforms should also continue to be investigated. Modern mobile devices are equipped with multi-touch screens, pressure sensitive panels, accelerometers, and gyroscopes, all of which may provide additional information to a KB system. Recent research has found that KB performance on the mobile devices is generally superior to that on hardware keyboards. However, many works have used fixed input, such as passcodes and phone numbers.

Data collection in works has been performed over a relatively short period of time. Like other biometrics, KB is subject to template aging as the typing behavior of individuals may change due to age, health conditions, and long term behavioral changes. Moreover, most works have not considered the dominant hand in data collection on mobile devices. The development of adaptive templates to

reduce false rejection is another area worthy of future investigation.

Majority of the researchers have not considered security issues in KB system. Like other authentication system, KB systems also vulnerable to various attacks and can be circumvented by a skillful imposter. So mitigation and proper countermeasures of those attacks are essential to gain acceptance of KB technology.

Lastly, the development of KB benchmark datasets will continue to encourage research in this area. While a number of benchmark datasets exist, these are limited to several scenarios. Future datasets should control for variables such as mechanical and touchscreen keyboard, input type, and demographics. The development of standardized protocols evaluating and comparing KB system performance remains an ongoing effort, as demonstrated by recent work [82].

6 Conclusion

This work presented a comprehensive review of keystroke biometrics research and an overview of keystroke biometric systems, which can be used as a starting point newcomers to the field. Moreover, a number of references, comparisons based on different techniques, and methods used in different researches has been presented and arranged in chronological order. This serves as a point of reference for other researchers to gauge their work identify future research directions.

Keystroke biometric systems are relatively unexplored compared to other disciplines, and a very limited number of studies have been conducted compared to other biometric systems. Despite generally lower accuracies than other biometric modalities, KB has a number of advantages, such as being low cost, transparent, noninvasive to the user, and offers ability to continuously monitor a system.

Acknowledgments The project is partially supported by NSF under grant 1457506 (Meikang Qiu). Md Liakat Ali would like to thank Dr. Charles C. Tappert, Dr. Lixin Tao, Dr. Meikang Qiu, and Pace University for Graduate Assistantship in his doctoral study and also for their support.

References

1. Ali, H., Wahyudi, W., & Salami, M.J.E. (2009). Keystroke pressure based typing biometrics authentication system by combining ann and anfis-based classifiers. In *Proceedings of the 5th international colloquium on signal processing and its applications (CSPA '09)* (Vol. 1, pp. 198–203). Kuala Lumpur. doi:10.1109/CSPA.2009.5069216.
2. Ali, M.L., Monaco, J., & Tappert, C. (2015). Hidden markov models in keystroke dynamics. In *Proceedings of student-faculty research day*. New York: Pace University.
3. Ali, M.L., Monaco, J., Tappert, C., & Qiu, M. (2015). Authentication and identification methods used in keystroke biometric systems. In *2015 IEEE international symposium on big data security on cloud (BigDataSecurity 2015)* (pp. 1424–1429). IEEE.
4. Ali, M.L., Thakur, K., & Tappert, C. (2015). User authentication and identification using neural network. *i-manager's Journal on Pattern Recognition*, 2(2), 28–39.
5. Allen, J.D. (2010). An analysis of pressure-based keystroke dynamics algorithms. Master's thesis, Southern Methodist University, Texas. <http://search.proquest.com/docview/608728043>.
6. Alsultan, A., & Warwick, K. (2013). Keystroke dynamics authentication: a survey of free-text methods. *International Journal of Computer Science Issues*, 10(4), 1–10.
7. Antal, M., & Szabó, L.Z. (2015). An evaluation of one-class and two-class classification algorithms for keystroke dynamics authentication on mobile devices. In *2015 20th international conference on control systems and computer science (CSCS)* (pp. 343–350). IEEE.
8. Antal, M., Szabo, L.Z., & Laszlo, I. (2014). Keystroke dynamics on android platform. In *INTER-ENG 2014, 8th international conference inerdisciplinary in engineering* (pp. 114–119). Tirgu Mures: Elsevier.
9. Azevedo, G.L.F., Cavalcanti, G.D.C., & Filho, E.C.B. (2007). An approach to feature selection for keystroke dynamics systems based on PSO and feature weighting. In *Proceedings of the IEEE congress on evolutionary computation (CEC '07)* (pp. 3577–3584). Singapore. doi:10.1109/CEC.2007.4424936.
10. Bakelman, N., Monaco, J.V., Cha, S.H., & Tapper, C.C. (2013). Keystroke biometric studies on password and numeric keypad input. In *2013 European intelligence and security informatics conference (EISIC)* (pp. 204–207). IEEE.
11. Balagani, K.S., Phoha, V.V., Ray, A., & Phoha, S. (2011). On the discriminability of keystroke feature vectors used in fixed text keystroke authentication. *Pattern Recognition Letters*, 32(7), 1070–1080.
12. Banerjee, S., & Woodard, D. (2012). Biometric authentication and identification using keystroke dynamics: a survey. *Journal of Pattern Recognition Research*, 7(1), 116–139.
13. Bello, L., Bertacchini, M., Benitez, C., Pizzoni, J.C., & Cipriano, M. (2010). Collection and publication of a fixed text keystroke dynamics dataset. Congreso Argentino de Ciencias de la Computación (CACIC'10).
14. Bleha, S., Slivinsky, C., & Hussien, B. (1990). Computer-access security systems using keystroke dynamics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(12), 1217–1222.
15. Bleha, S.A., & Obaidat, M.S. (1993). Computer users verification using the perceptron algorithm. *IEEE Transactions on Systems Man and Cybernetics*, 23(3), 900–902.
16. Bondada, M.B., & Bhanu, S. (2014). Analyzing user behavior using keystroke dynamics to protect cloud from Malicious insiders. In *2014 IEEE international conference on cloud computing in emerging markets (CCEM)* (pp. 1–8). IEEE.
17. Bours, P., & Masoudian, E. (2014). Applying keystroke dynamics on one-time pin codes. In *2014 International workshop on biometrics and forensics (IWBF)* (pp. 1–6). IEEE.
18. Bours, P., & Mondal, S. (2015). Performance evaluation of continuous authentication systems. *IET Biometrics*.
19. Brown, M., & Rogers, S.J. (1993). User identification via keystroke characteristics of typed names using neural networks. *International Journal of Man-Machine Studies*, 39(6), 999–1014.
20. Buchoux, A., & Clarke, N. (2006). Deployment of keystroke analysis on a smartphone. In *Proceedings of the 6th Australian information security management conference*. Perth.

21. Campisi, P., Maiorana, E., Bosco, M.L., & Neri, A. (2009). User authentication using keystroke dynamics for cellular phones. *IET Signal Processing*, 3(4), 333–341.
22. Capuano, N., Marsella, M., Miranda, S., & Salerno, S. (1999). User authentication with neural networks. In *Proceedings of the 5th international conference on engineering applications of neural networks (EANN 99)* (pp. 200–205). Warsaw.
23. Clarke, N.L., & Furnell, S.M. (2007). Authenticating mobile phone users using keystroke analysis. *International Journal of Information Security*, 6(1), 1–14.
24. Crawford, H. (2010). Keystroke dynamics: characteristics and opportunities. In *Proceedings of the 8th international conference on privacy, security and trust (PST '10)* (pp. 205–212).
25. Darabseh, A., & Siami Namin, A. (2015). Keystroke active authentications based on most frequently used words. In *Proceedings of the 2015 ACM international workshop on international workshop on security and privacy analytics* (pp. 49–54). ACM.
26. Davoudi, H., & Kabir, E. (2009). A new distance measure for free text keystroke authentication. In *Proceedings of the 14th International CSI computer conference (CSICC '09)* (pp. 570–575). Tehran. doi:10.1109/CSICC.2009.5349640.
27. Deng, Y., & Zhong, Y. (2013). Keystroke dynamics user authentication based on gaussian mixture model and deep belief nets. *ISRN Signal Processing*, 2013.
28. Deutschmann, I., Nordstrom, P., & Nilsson, L. (2013). Continuous authentication using behavioral biometrics. *IT Professional*, 15(4), 12–15.
29. D'Lima, N., & Mittal, J. (2015). Password authentication using keystroke biometrics. In *2015 International conference on communication, information & computing technology (ICCICT)* (pp. 1–6). IEEE.
30. Douhou, S., & Magnus, J.R. (2009). The reliability of user authentication through keystroke dynamics. *Statistica Neerlandica*, 63(4), 432–449.
31. Draffin, B., Zhu, J., & Zhang, J. (2014). *KeySens: passive user authentication through micro-behavior modeling of soft keyboard interaction* (Vol. 130). Springer International Publishing.
32. El-Abed, M., Dafer, M., & El Khayat, R. (2014). Rhu keystroke: a mobile-based benchmark for keystroke dynamics systems. In *2014 International Carnahan conference on security technology (ICCST)* (pp. 1–4). IEEE.
33. Filho, J.R.M., & Freire, E.O. (2006). On the equalization of keystroke timing histograms. *Pattern Recognition Letters*, 27(13), 1440–1446.
34. Gaines, R.S., Lisowski, W., Press, S.J., & Shapiro, N. (1980). Authentication by keystroke timing: some preliminary results. Tech. rep., R-2526-NFS, Rand Corporation, Santa Monica.
35. Gascon, H., Uellenbeck, S., Wolf, C., & Rieck, K. (2014). Continuous authentication on mobile devices by analysis of typing motion behavior. In *Proceedings of GI conference "Sicherheit"* (Vol. P-228, pp. 1–12).
36. Giot, R., Dorizzi, B., & Rosenberger, C. (2011). Analysis of template update strategies for keystroke dynamics. In *Proceedings of the IEEE workshop on computational intelligence in biometrics and identity management (CIBIM '11)* (pp. 21–28).
37. Giot, R., El-Abed, M., Hemery, B., & Rosenberger, C. (2011). Unconstrained keystroke dynamics authentication with shared secret. *Computers and Security*, 30(6–7), 427–445.
38. Giot, R., El-Abed, M., & Rosenberger, C. (2009). 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). doi:10.1109/BTAS.2009.5339051.
39. Giot, R., El-Abed, M., & Rosenberger, C. (2009). Keystroke dynamics authentication for collaborative systems. In *Proceedings of the international symposium on collaborative technologies and systems (CTS '09)* (pp. 172–179).
40. Giot, R., El-Abed, M., & Rosenberger, C. (2009). Keystroke dynamics with low constraints svm based passphrase enrollment. In *Proceedings of the IEEE 3rd international conference on biometrics: theory, applications and systems (BTAS '09)* (pp. 1–6). Washington, DC. doi:10.1109/BTAS.2009.5339028.
41. Giot, R., El-Abed, M., & Rosenberger, C. (2012). Web-based benchmark for keystroke dynamics biometric systems: a statistical analysis. In *2012 eighth international conference on intelligent information hiding and multimedia signal processing (IIH-MSP)* (pp. 11–15). IEEE.
42. Grabham, N.J., & White, N.M. (2008). Use of a novel keypad biometric for enhanced user identity verification. In *Proceedings of the IEEE international instrumentation and measurement technology conference (IMTC'08)* (pp. 12–16). IEEE.
43. Gunetti, D., & Picardi, C. (2005). Keystroke analysis of free text. *ACM Transactions on Information and System Security*, 8(3), 312–347.
44. Gunetti, D., Picardi, C., & Ruffo, G. (2005). Keystroke analysis of different languages: a case study. In *Advances in intelligent data analysis VI, vol. 3646 Lecture notes in Computer Science* (pp. 133–144). Berlin: Springer.
45. Harun, N., Dlay, S.S., & Woo, W.L. (2010). Performance of keystroke biometrics authentication system using multilayer perceptron neural network (mlp nn). In *7th International symposium on communication systems networks and digital signal processing (CSNDSP' 2010)* (pp. 711–714). IEEE, Newcastle upon Tyne.
46. Hosseinzadeh, D., & Krishnan, S. (2008). Gaussian mixture modeling of keystroke patterns for biometric applications. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 8(6), 816–826.
47. Idrus, S.Z.S., Cherrier, E., Rosenberger, C., & Bours, P. (2013). Soft biometrics database: a benchmark for keystroke dynamics biometric systems. In *2013 International conference of the biometrics special interest group (BIOSIG)* (pp. 1–8). IEEE.
48. Mäntyjärvi, J., Koivumäki, J., & Vuori, P. (2002). Keystroke recognition for virtual keyboard. In *International conference on multimedia and expo, 2002. ICME'02. Proceedings. 2002 IEEE* (Vol. 2, pp. 429–432). IEEE.
49. Jain, L., Monaco, J.V., Coakley, M.J., & Tappert, C.C. (2014). Passcode keystroke biometric performance on touchscreen is superior to that on hardware keyboards. *International Journal of Research in Computer Applications and Information Technology*, 2(4), 29–33.
50. Jeanjaitrong, N., & Bhattarakosol, P. (2013). Feasibility study on authentication based keystroke dynamic over touch-screen devices. In *2013 13th international symposium on communications and information technologies (ISCIT)* (pp. 238–242). IEEE.
51. Johansen, U.A. (2012). Keystroke dynamics on a device with touch screen. Master's thesis, Gjøvik University College, Norway. <http://hdl.handle.net/11250/143992>.
52. Joyce, R., & Gupta, G. (1990). Identity authentication based on keystroke latencies. *Communications of the ACM*, 33(2), 168–176.
53. Kaganov, V., Korolyov, A., Krylov, M., Mashechkin, I., & Petrovskiy, M. (2014). Hybrid method for active authentication using keystroke dynamics. In *2014 14th international conference on hybrid intelligent systems (HIS)* (pp. 61–66). IEEE.
54. Kaneko, Y., Kinpara, Y., & Shiomi, Y. (2011). A hamming distance-like filtering in keystroke dynamics. In *Proceedings of the ninth annual international conference on privacy, security and trust (PST '11)* (pp. 93–95). Montreal. doi:10.1109/PST.2011.5971969.
55. Kang, P., Seob Hwang, S., & Cho, S. (2007). Continual retraining of keystroke dynamics based authenticator. In *ICB'07*

- proceedings of the 2007 international conference on advances in biometrics* (Vol. 4642, pp. 1203–1211). Berlin.
56. Kanimozhi, M., Puvirajasingam, K., & Avitha, M.S. (2014). Survey on keystroke dynamics for a better biometric authentication system. *International Journal of Emerging Technologies and Engineering (IJETE)*, 1(9), 116–139.
 57. Karnan, M., & Akila, M. (2010). Personal authentication based on keystroke dynamic using soft computing techniques. In *Second international conference on communication software and networks (CCSN '10)* (pp. 334–338). Singapore. doi:10.1109/ICCSN.2010.50.
 58. Karnan, M., Akila, M., & Krishnaraj, N. (2011). Biometric personal authentication using keystroke dynamics: a review. *Applied Soft Computing Journal*, 11(2), 1565–1573.
 59. Killourhy, K., & Maxion, R. (2010). Why did my detector do that?!: predicting keystroke-dynamics error rates. In *Proceedings of the 13th international conference on recent advances in intrusion detection (RAID '10)* (pp. 256–276). Ottawa.
 60. Killourhy, K.S., & Maxion, R.A. (2009). Comparing anomaly detection algorithms for keystroke dynamics. In *Proceedings of the IEEE/IFIP international conference on dependable systems and networks (DSN '09)* (pp. 125–134).
 61. Kolakowska, A. (2015). Recognizing emotions on the basis of keystroke dynamics. In *2015 8th international conference on human system interactions (HSI)* (pp. 291–297). IEEE.
 62. Kotani, K., & Horii, K. (2005). Evaluation on a keystroke authentication system by keying force incorporated with temporal characteristics of keystroke dynamics. *Behaviour and Information Technology*, 24(4), 289–302.
 63. Li, J., Qiu, M., Niu, J.W., Yang, L.T., Zhu, Y., & Ming, Z. (2013). Thermal-aware task scheduling in 3d chip multiprocessor with real-time constrained workloads. *ACM Transactions on Embedded Computing Systems (TECS)*, 12(2), 24.
 64. Li, Y., Dai, W., Ming, Z., & Qiu, M. (2015). Privacy protection for preventing data over-collection in smart city. *IEEE Transactions on Computers*, 99, 1.
 65. Li, Y., Zhang, B., Cao, Y., Zhao, S., Gao, Y., & Liu, J. (2011). Study on the beihang keystroke dynamics database. In *Proceedings of the international joint conference on biometrics (IJCB '11)* (pp. 1–5).
 66. Lin, D.T. (1997). Computer-access authentication with neural network based keystroke identity verification. In *Proceedings of the 1997 IEEE international conference on neural networks* (Vol. 1, pp. 174–178).
 67. Loy, C.C., Lai, W.K., & Lim, C.P. (2007). Keystroke patterns classification using the artmap-fd neural network. In *Proceedings of the 3rd international conference on intelligent information hiding and multimedia signal processing (IIHMSP '07)* (Vol. 1, pp. 61–64). Kaohsiung. doi:10.1109/IIH-MSP.2007.218.
 68. Loy, C.C., Lai, W.K., & Lim, C.P. (2005). The development of a pressure-based typing biometrics user authentication system. Asean virtual instrumentation applications contest submission, National Instruments, Austin.
 69. Lv, H.R., & Wang, W.Y. (2006). Biologic verification based on pressure sensor keyboards and classifier fusion techniques. *IEEE Transactions on Consumer Electronics*, 52(3), 1057–1063. doi:10.1109/TCE.2006.1706507.
 70. Maisuria, L.K., Ong, C.S., & Lai, W.K. (1999). A comparison of artificial neural networks and cluster analysis for typing biometrics authentication. In *Proceedings of the international joint conference on neural networks (IJCNN '99)* (Vol. 5, pp. 3295–3299). Washington, DC: IEEE.
 71. Martono, W., Ali, H., & Salami, M.J.E. (2007). Keystroke pressure-based typing biometrics authentication system using support vector machines. In *Proceedings of the 2007 international conference on computational science and its applications: volume part II* (pp. 85–93). Kuala Lumpur.
 72. Maxion, R.A., & Killourhy, K.S. (2010). Keystroke biometrics with number-pad input. In *Proceedings of the IEEE/IFIP international conference on dependable systems and networks (DSN '10)* (pp. 201–210).
 73. McLoughlin, I.V., & Naidu, M.S.O.N. (2009). Keypress biometrics for user validation in mobile consumer devices. In *Proceedings of the IEEE 13th international symposium on consumer electronics (ISCE'09)* (pp. 280–284).
 74. de Mendizabal-Vazquez, I., de Santos-Sierra, D., Guerra-Casanova, J., & Sanchez-Avila, C. (2014). Supervised classification methods applied to keystroke dynamics through mobile devices. In *2014 International Carnahan conference on security technology (ICCST)* (pp. 1–6). IEEE.
 75. Messerman, A., Mustafic, T., Camtepe, S.A., & Albayrak, S. (2011). Continuous and non-intrusive identity verification in realtime environments based on free-text keystroke dynamics. In *Proceedings of the international joint conference on biometrics (IJCB '11)* (pp. 1–8). Washington.
 76. Monaco, J., Ali, M.L., & Tappert, C. (2015). Spoofing key-press latencies with a generative keystroke dynamics model. In *2015 7th international conference on biometrics: theory, applications and systems (BTAS 2015)*. IEEE.
 77. Monaco, J.V., Stewart, J.C., Cha, S., & Tappert, C.C. (2013). Behavioral biometric verification of student identity in online course assessment and authentication of authors in literary works. In *Proceedings of IEEE sixth international conference on biometrics: theory, applications and systems (BTAS '2013)* (pp. 1–8). Arlington. doi:10.1109/BTAS.2013.6712743.
 78. Mondal, S., & Bours, P. (2014). Continuous authentication using fuzzy logic. In *Proceedings of the 7th international conference on security of information and networks* (p. 231). ACM.
 79. Mondal, S., & Bours, P. (2015). Continuous authentication in a real world settings. In *2015 eighth international conference on advances in pattern recognition (ICAPR)* (pp. 1–6). IEEE.
 80. Monrose, F., & Rubin, A. (1997). Authentication via keystroke dynamics. In *Proceedings of the 4th ACM conference on computer and communications security* (pp. 48–56). Zurich.
 81. Monrose, F., & Rubin, A.D. (2000). Keystroke dynamics as a biometric for authentication. *Future Generation Computer Systems*, 16(4), 351–359.
 82. Morales, A., Falanga, M., Fierrez, J., Sansone, C., & Ortega-Garcia, J. (2015). Keystroke dynamics recognition based on personal data: a comparative experimental evaluation implementing reproducible research. In *Biometrics: theory, applications and systems (BTAS)*. IEEE.
 83. Ngugi, B., Kahn, B.K., & Tremaine, M. (2011). Typing biometrics: impact of human learning on performance quality. *Journal of Data and Information Quality*, 2(2).
 84. Nguyen, T.T., Le, T.H., & Le, B.H. (2010). Keystroke dynamics extraction by independent component analysis and bio-matrix for user authentication. In *Proceedings of the 11th pacific rim international conference on trends in artificial intelligence* (Vol. 6230, pp. 477–486). Daegu.
 85. Nonaka, H., & Kurihara, M. (2004). Sensing pressure for authentication system using keystroke dynamics. In *Proceedings of the international conference on computational intelligence* (pp. 19–22). Istanbul.
 86. Obaidat, M.S. (1995). Verification methodology for computer systems users. In *Proceedings of the 1995 ACM symposium on applied computing* (pp. 258–262).
 87. Obaidat, M.S., & Sadoun, B. (1997). Verification of computer users using keystroke dynamics. *IEEE Transactions on Systems, Man, and Cybernetics B*, 27(2), 261–269.

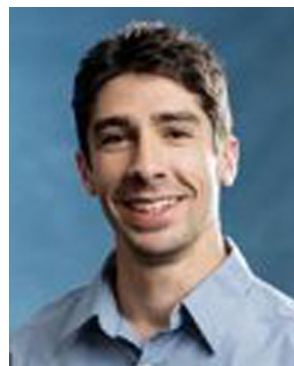
88. Pavaday, N., & Soyjaudah, K.M.S. (2007). Investigating performance of neural networks in authentication using keystroke dynamics. In *Proceedings of the IEEE AFRICON 2007 conference* (pp. 1–8).
89. Peacock, A., Ke, X., & Wilkerson, M. (2004). Typing patterns: a key to user identification. *IEEE Security and Privacy*, 2(5), 40–47.
90. Pisani, P.H., Lorena, A.C., & de Carvalho, A.C. (2015). Adaptive approaches for keystroke dynamics. In *2015 international joint conference on neural networks (IJCNN)* (pp. 1–8). IEEE.
91. Popovici, E.C., Guta, O.G., Stancu, L., Arseni, S.C., Fratu, O., et al. (2013). Mlp neural network for keystroke-based user identification system. In *2013 11th international conference on telecommunication in modern satellite, cable and broadcasting services (TELSIKS)* (Vol. 1, pp. 155–158). IEEE.
92. Pusara, M. (2007). An examination of user behavior for user re-authentication. Ph.D. thesis, Purdue University, West Lafayette.
93. Qiu, M., Ming, Z., Li, J., Liu, J., Quan, G., & Zhu, Y. (2013). Informer homed routing fault tolerance mechanism for wireless sensor networks. *Journal of Systems Architecture*, 59(4–5), 260–270.
94. Qiu, M., Ming, Z., Li, J., Liu, S., Wang, B., & Lu, Z. (2012). Three-phase time-aware energy minimization with dvfs and unrolling for chip multiprocessors. *Journal of Systems Architecture*, 58(10), 439–445.
95. Rahman, K.A., Balagani, K.S., & Phoha, V.V. (2011). Making impostor pass rates meaningless: a case of snoop-forge-replay attack on continuous cyber-behavioral verification with keystrokes. In *Proceedings of the IEEE computer society conference on computer vision and pattern recognition workshops (CVPRW '11)* (pp. 31–38). Colorado Springs. doi:10.1109/CVPRW.2011.5981729.
96. Revett, K., de Magalhães, S.T., & Santos, H. (2005). Data mining a keystroke dynamics based biometrics database using rough sets. In *Proceedings of the Portuguese conference on artificial intelligence (EPIA '05)* (pp. 188–191). Covilha. doi:10.1109/EPIA.2005.341292.
97. Revett, K., de Magalhães, S.T., & Santos, H.M.D. (2007). On the use of rough sets for user authentication via keystroke dynamics. In *Proceedings of the Portuguese conference on artificial intelligence (EPIA '07)* (pp. 145–159). Guimarães.
98. Robinson, J.A., Liang, V.M., Chambers, J.A.M., & MacKenzie, C.L. (1998). Computer user verification using login string keystroke dynamics. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 28(2), 236–241.
99. Rodrigues, R.N., Yared, G.F.G., do, N., Costa, C.R., Yabu-Uti, J.B.T., Violaro, F., & Ling, L.L. (2006). Biometric access control through numerical keyboards based on keystroke dynamics. In *Proceedings of the 2006 international conference on advances in biometrics (ICB'06)* (Vol. 3832, pp. 640–646). Hong Kong.
100. Roth, J., Liu, X., & Metaxas, D. (2014). On continuous user authentication via typing behavior. *IEEE Transactions on Image Processing*, 23(10), 4611–4624.
101. Roth, J., Liu, X., Ross, A., & Metaxas, D. (2013). Biometric authentication via keystroke sound. In *2013 international conference on biometrics (ICB)* (pp. 1–8). IEEE.
102. Rybnik, M., Tabledzki, M., Adamski, M., & Saeed, K. (2013). An exploration of keystroke dynamics authentication using non-fixed text of various length. In *2013 international conference on biometrics and Kansei engineering (ICBAKE)* (pp. 245–250). IEEE.
103. Hwang, S.S., Cho, S., & Park, S. (2009). Keystroke dynamics-based authentication for mobile devices. *Computers & Security*, 28(1), 85–93.
104. Saevanee, H., & Bhatarakosol, P. (2008). User authentication using combination of behavioral biometrics over the touchpad acting like touch screen of mobile device. In *Proceedings of the international conference on computer and electrical engineering (ICCEE '08)* (Vol. 2, pp. 82–86).
105. Saevanee, H., & Bhatarakosol, P. (2009). Authenticating user using keystroke dynamics and finger pressure. In *6th IEEE consumer communications and networking conference (CCNC'09)* (pp. 1–2). IEEE, Las Vegas. doi:10.1109/CCNC.2009.4784783.
106. Samura, T., & Nishimura, H. (2009). Keystroke timing analysis for individual identification in Japanese free text typing. In *Proceedings of the ICROS-SICE international joint conference (ICCAS-SICE '09)* (pp. 3166–3170). Fukuoka.
107. Samura, T., & Nishimura, H. (2011). Keystroke timing analysis for personal authentication in Japanese long text input. In *Proceedings of the 50th annual conference on society of instrument and control engineers (SICE '11)* (pp. 2121–2126). Tokyo.
108. Sen, S., & Muralidharan, K. (2014). Putting pressure on mobile authentication. In *Seventh international conference on mobile computing and ubiquitous networking (ICMU'2014)* (pp. 56–61). IEEE, Singapore. doi:10.1109/ICMU.2014.6799058.
109. Senathipathi, K., & Batri, K. (2014). An analysis of particle swarm optimization and genetic algorithm with respect to keystroke dynamics. In *2014 international conference on green computing communication and electrical engineering (ICGC-CEE)* (pp. 1–11). IEEE.
110. Shanmugapriya, D., & Padmavathi, G. (2009). A survey of biometric keystroke dynamics: approaches, security and challenges. *International Journal of Computer Science and Information Security*, 5, 115–119.
111. Sheng, Y., Phoha, V.V., & Rovnyak, S.M. (2005). A parallel decision tree-based method for user authentication based on keystroke patterns. *IEEE Transactions on Systems, Man and Cybernetics Part B: Cybernetics*, 35(4), 826–833.
112. Shimshon, T., Moskovitch, R., Rokach, L., & Elovici, Y. (2010). Clustering di-graphs for continuously verifying users according to their typing patterns. In *Proceedings of the IEEE 26th convention of electrical and electronics engineers in Israel (IEEEI '10)* (pp. 445–449).
113. Shimshon, T., Moskovitch, R., Rokach, L., & Elovici, Y. (2010). Continuous verification using keystroke dynamics. In *Proceedings of the international conference on computational intelligence and security (CIS'10)* (pp. 411–415).
114. Singh, S., & Arya, K.V. (2011). Key classification: a new approach in free text keystroke authentication system. In *Proceedings of the 3rd Pacific-Asia conference on circuits, communications and system (PACCS '11)* (pp. 1–5). Wuhan. doi:10.1109/PACCS.2011.5990168.
115. Singh, S., & Sinha, M. (2013). Pattern construction by extracting user specific features in keystroke authentication system. In *2013 4th international conference on computer and communication technology (ICCCCT)* (pp. 181–184). IEEE.
116. Spillane, R. (1975). Keyboard apparatus for personal identification. *IBM Technical Disclosure Bulletin*, 17(3346), 3346.
117. Stewart, J.C., Monaco, J.V., Cha, S., & Tappert, C.C. (2011). An investigation of keystroke and stylometry traits for authenticating online test takers. In *Proceedings of the international joint conference on biometrics (IJCB'11)* (pp. 1–7). Washington, DC.
118. Sulong, A., Wahyudi, W., & Siddiqi, M.D. (2009). Intelligent keystroke pressure-based typing biometrics authentication system using radial basis function network. In *Proceedings of*

- the 5th international colloquium on signal processing and its applications (CSPA '09) (pp. 151–155).
119. Sunghoon, P., Jooseoung, P., & Sungzoon, C. (2010). User authentication based on keystroke analysis of long free texts with a reduced number of features. In *Proceedings of the 2nd international conference on communication systems, networks and applications (ICCSNA '10)* (Vol. 1, pp. 433–435). Hong Kong. doi:10.1109/ICCSNA.2010.5588979.
 120. Syed, Z., Banerjee, S., & Cukic, B. (2014). Leveraging variations in event sequences in keystroke-dynamics authentication systems. In *2014 IEEE 15th international symposium on high-assurance systems engineering (HASE)* (pp. 9–16). IEEE.
 121. Taculin, A.R.F., Abuhan, D.M., Cruz, J.R.B., Santos, M.L., & Crisostomo, R.V. (2014). Using molecular algorithm in keystroke dynamics. In *2014 International conference on computer, communications, and control technology (I4CT)* (pp. 193–197). IEEE.
 122. Tappert, C., Cha, S., Villani, M., & Zack, R. (2010). A keystroke biometric system for long-text input. *International Journal of Information Security and Privacy*, 4(1), 32–60.
 123. Teh, P.S., Teoh, A.B.J., Tee, C., & Ong, T.S. (2011). A multiple layer fusion approach on keystroke dynamics. *Pattern Analysis and Applications*, 14(1), 23–36.
 124. Teh, P.S., Teoh, A.B.J., & Yue, S. (2013). A survey of keystroke dynamics biometrics. *The Scientific World Journal*, 2013, 1–24. doi:10.1155/2013/408280.
 125. Trojahn, M., Arndt, F., & Ortmeier, F. (2013). Authentication with keystroke dynamics on Touchscreen Keypads—effect of different N-Graph combinations. In *MOBILITY 2013, The third international conference on mobile services, resources, and users* (pp. 114–119).
 126. Tsai, C.J., Chang, T.Y., Tsai, W.J., Peng, C.C., Chiang, M.L., & Wu, H.S. (2014). Work in progress: a new approach of changeable password for keystroke dynamics authentication system on smart phones. In *2014 9th international conference on communications and networking in China (ChinaCOM)* (pp. 353–356). IEEE.
 127. Villani, M., Tappert, C., Ngo, G., Simone, J., Fort, H.S., & Cha, S. (2006). Keystroke biometric recognition studies on long-text input under ideal and application-oriented conditions. In *Proceedings of the conference on computer vision and pattern recognition workshops (CVPRW'06)* (p. 39).
 128. Vural, E., Huang, J., Hou, D., & Schuckers, S. (2014). Shared research dataset to support development of keystroke authentication. In *2014 IEEE international joint conference on biometrics (IJCB)* (pp. 1–8). IEEE.
 129. Wankhede, S.B., & Verma, S. (2014). Keystroke dynamics authentication system using neural network. *International Journal of Innovative Research and Development*, 3(1), 157–164.
 130. Wu, G., Zhang, H., Qiu, M., Ming, Z., Li, J., & Qin, X. (2013). A decentralized approach for mining event correlations in distributed system monitoring. *Journal of parallel and Distributed Computing*, 73(3), 330–340.
 131. Xi, K., Tang, Y., & Hu, J. (2011). Correlation keystroke verification scheme for user access control in cloud computing environment. *The Computer Journal*, 54(10), 1632–1644. doi:10.1093/comjnl/bxr064.
 132. Yu, E., & Cho, S. (2003). Novelty detection approach for keystroke dynamics identity verification. In *Intelligent data engineering and automated learning* (Vol. 2690, pp. 1016–1023). Berlin: Springer.
 133. Zack, R.S., Tappert, C.C., & Cha, S.H. (2010). Performance of a long-text-input keystroke biometric authentication system using an improved k-nearest-neighbor classification method. In *Fourth IEEE international conference on biometrics: theory applications and systems (BTAS '2010)* (pp. 1–6). Washington, DC.
 134. Zhao, Y. (2006). Learning user keystroke patterns for authentication. In *Proceedings of the world academy of science, engineering and technology* (Vol. 14, pp. 65–70). Karnataka.
 135. Zhong, Y., Deng, Y., & Jain, A.K. (2012). Keystroke dynamics for user authentication. In *2012 IEEE computer society conference on computer vision and pattern recognition workshops (CVPRW)* (pp. 117–123). IEEE.



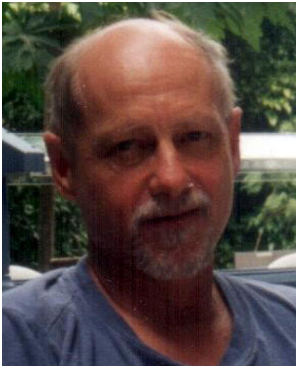
Md Liakat Ali is a doctoral student and Adjunct Professor at Pace University, New York, USA. He completed his M.Sc in Computer Science (Major in Security Engineering) and M.Sc in Electrical Engineering (Major in Telecommunication) during 2007 and 2008 respectively, from Blekinge Institute of Technology, Sweden. Mr. Ali received a quality thesis award for his Master's thesis, which was one of the top ten theses out of fifty in the academic year of 2007. From

February'2009 to June' 2012 he taught undergraduate Computer Science course at two different universities in Bangladesh. Mr. Ali also an Adjunct Professor at Rowan University, and County College of Morris, New Jersey, USA. Ali's research interest includes Pattern Recognition, Machine Learning Biometrics, Data Mining and Big Data Analytics, and Cyber Security.



John V. Monaco is an Adjunct Professor at Pace University. Monaco earned both his BS in Computer Science and Mathematics and his MS in Computer Science and Ph.D in Computer Science from Pace University. In 2011, Monaco was selected for the highly competitive and prestigious Information Assurance Scholarship Program by the U.S. Department of Defense. In 2013, Monaco was named one of Westchester's "Top Professionals under 30" for research

in keystroke biometrics at Pace University, and in 2014 he placed 1st in an international competition on identifying users based on eye movements. His primary area of interest is behavioral biometrics.



Charles C. Tappert is a Professor of Computer Science and Associate Program Director of the Doctor of Professional Studies in Computing Program at Pace University. He has a Ph.D. in Electrical Engineering from Cornell University and was a Fulbright Scholar. He worked on speech and handwriting recognition at IBM for 26 years, taught at the U.S. Military Academy at West Point for seven years, and has been a professor at Pace University since 2000.

He has over 100 publications and his research interests include pattern recognition, machine learning biometrics, speech recognition/voice applications, handwriting recognition/pen computing, human-computer interaction, artificial intelligence, and big data analytics.



Meikang Qiu is currently an Associate Professor of Computer Science at Pace University. He received the B.E. and M.E. degrees from Shanghai Jiao Tong University, China and obtained his Ph.D. in Computer Science and Engineering from University of Texas at Dallas in 2007. Dr. Qiu has worked at Chinese Helicopter RD Institute and IBM. He is an IEEE Senior member. In 2011, he won the Best Paper Award of the ACM Transactions on Design

Automation of Electronic Systems (TODAES). He has published more than 150 journal and conference papers and 3 books. Dr. Qiu's research interests include embedded systems, computer architecture, and high performance computing.