CHAPTER 1

A Survey on Keystroke Dynamics Biometrics: Approaches, Advances, and Evaluations

Yu Zhong and Yunbin Deng

In this review paper we present a comprehensive survey of research efforts in the past couple of decades on keystroke dynamics biometrics. We review the literature in light of various feature extraction, feature matching and classification methods for keystroke dynamics. We also discuss recent trends in keystroke dynamics research, including its use in mobile environments, as a soft biometrics, and its fusion with other biometric modalities. We further address the evaluation of keystroke biometric systems, including traditional and new performance metrics, and list publicly available keystroke datasets for performance benchmarks to promote synergy in the research community.

1.1 Introduction

With the ever increasing demand for more secure access control in many of today's security applications, traditional methods such as PINs, tokens, or passwords fail to

Yu Zhong and Yunbin Deng
BAE Systems
6 New England Executive Park, Burlington MA 01803, USA
e-mail: {Yu.Zhong, Yunbin.deng}@baesystems.com

Editors: Y. Zhong and Y. Deng, *Recent Advances in User Authentication Using Keystroke Dynamics Biometrics*

DOI: 10.15579/gcsr.vol2.ch1, GCSR Vol. 2, pp. 1-22, 2015

keep up with the challenges presented because they can be lost or stolen. On the other hand, biometrics [51, 52, 54, 74, 87, 104, 106, 109] based on "who" the person is or "how" the person behaves present a significant security advancement to meet these new challenges. Among them, keystroke dynamics [3, 11, 38, 78, 83, 84, 85, 96, 99, 105] provide a natural choice for secure "password-free" computer access. Keystroke dynamics refer to the habitual patterns or rhythms an individual exhibits while typing on a keyboard input device. These rhythms and patterns of tapping are idiosyncratic [29, 30], in the same way as a person's handwriting or signature, due to their similar governing neurophysiological mechanisms. In fact, as early as the 19th century, telegraph operators could recognize each other through their specific tapping styles [66]. This suggests that keystroke dynamics contain sufficient information to serve as a biometric identifier.

Compared to other biometric modalities, keystroke biometrics has more desirable properties due to being user-friendly and non-intrusive. Keystroke dynamics data can be collected without a user's cooperation or even awareness. Continuous authentication of a person is possible with keystroke dynamics just as a mere consequence of that person using a computer. Unlike many other biometrics, keystroke data can be collected using only software with no additional hardware. In summary, keystroke dynamics biometrics enables a cost effective, user friendly, and continuous user authentication mechanism.

Although keystroke dynamics is governed by a person's highly individualistic neurophysiological pathway, it can also be influenced by his or her psychological state. As a "behavioral" biometrics [110], keystroke dynamics exhibits instabilities due to transient factors such as emotion, stress, drowsiness, and etc. [12, 31]. It also depends on external factors, such as the input keyboard used, which can be further compounded possibly due to different key layouts. The keying times can be noisy with outliers. As keystroke biometrics exploits the habitual rhythms in a person's typing, keystrokes of frequently typed words or strings show more consistency and are better discerners [81, 83].

The use of keystroke dynamics for verification and identification purposes was first investigated back in the 1970's [34, 96]. Earlier research work [35, 105] on keystroke biometrics has mainly concerned the use of static text [6, 76, 88], when the keystroke dynamics of a specific pre-enrolled text, such as a password, is analyzed at a certain time such as at log-on. There has been a shift of focus toward the more challenging free text keystroke authentication [43], where a user is authenticated using unconstrained text [10, 77, 78, 79, 80, 81, 92, 112]. While static text keystroke dynamics biometrics are often used during the logon process to provide a onetime authentication, free text keystroke biometric systems enable continuously authentication of a user during the entire session for increased security [28]. A practical keystroke biometric system can use static text, free text, or a combination of both.

As keystroke dynamics biometrics has drawn intense research interest the past couple of decades, a number of survey papers have been published [2, 7, 11, 26, 60, 91, 98]. In our paper, we not only survey existing approaches but also look at possibilities and collect information necessary for future advances in keystroke dynamics biometrics. We start by looking at current keystroke feature extraction and classification methods (Section 1.2). We then review recent advances and new trends (Section 1.3). In

particular we focus on new development in keystroke biometrics for mobile devices, a rapidly growing and changing area of study. Finally, we see an urgent need for the research community to make a concerted effort in establishing large and representative keystroke biometrics benchmark datasets in order to assess the advances in new algorithms. To this end, we compile a list of publicly available keystroke dynamics datasets to promote the sharing of standard experimental and performance evaluation protocols that will lead to more effective and objective progress assessment (Section 1.4). We conclude the paper with a summary and discussion of future directions.

1.2 Keystroke Dynamics Algorithms

In this section we review the features that are used to characterize individual keystroke dynamics and the classification methods applied to interpret the extracted features.

1.2.1 Keystroke Dynamics Features

Keystroke dynamics features are usually extracted using the timing information of the key down/hold/up events. The hold time or dwell time of individual keys, and the latency between two keys, i.e., the time interval between the release of a key and the pressing of the next key are typically exploited. Digraphs, which are the time latencies between two successive keystrokes, are commonly used. Trigraphs, which are the time latencies between every three consecutive keys, and similarly, n-graphs, have been investigated as well. In their study on keystroke analysis using free text, Sim and Janakiraman [94] investigated the effectiveness of digraphs and more generally n-graphs for free text keystroke biometrics, and concluded that n-graphs are discriminative only when they are word-specific. As such, the digraph and n-graph features do depend on the word context they are computed in.

Gaines et al. [35] did a preliminary study on keystroke dynamics based authentication using the T-test on digraph features. Monrose and Rubin [81] later extracted keystroke features using the mean and variance of digraphs and trigraphs. Using the Euclidean distance metric with Bayesian-like classifiers, they reported a correct identification rate of 92% for their dataset containing 63 users.

Bergadano et al. [10] and later Gunetti and Picardi [43] proposed to use the relative order of duration times for different n-graphs to extract keystroke features that was found to be more robust to the intra-class variations than absolute timing. They demonstrated that the new relative feature, when combined with features using absolute timing, improved the authentication performance using free text.

Syed et al. [97] extracted features using variations in keying event, motivated by the observation that the same text string may be inputted using different key entry sequences. They found out that such variations in typing sequences contain distinguishing information for user authentication, while being independent of typing proficiency. In an interesting study, Roth et al. explored the use of keystroke acoustics for user identification [89]. They built a virtual vocabulary based on keystroke sound, and then extracted digraph latency features using the learned virtual keyboard. They were able to obtain an EER of 11% on a dataset of 50 subjects, indicating the promise

of keystroke acoustics for user authentication. Furthermore, the new approach does not require direct access of the computer as traditional methods do.

1.2.2 Keystroke Dynamics Classification

Over the years, keystroke biometrics research has utilized many existing machine learning and classification techniques. Different distance metrics, such as the Euclidean distance, the Mahalanobis distance, and the Manhattan distance, have been explored. Both classical and advanced classifiers have been used. These methods are discussed in more details in the following subsections.

1.2.2.1 Distance based classification

Once feature vectors are extracted to represent the typing characteristics, they are then classified for authentication and identification purposes. Early research mainly used the Nearest Neighbor classifier [16] with various distance functions that measure the similarities between keystroke features. Euclidean distance has been the default distance metric for its simplicity and geometric intuitiveness [11, 78]. Other distance functions have also been explored, and are reviewed as follows.

A. Mahalanobis Distance Despite its simplicity and intuitiveness, Euclidean distance has two major limitations:

1. It is highly sensitive to scale variations in the feature variables, and
2. It has no means to deal with the correlation between feature variables.

Mahalanobis distance, on the other hand, takes into account the covariance of data variables to correct for the heterogeneity and non-isotropy observed in most real data. Because it handles the correlated data well, it has been popularly used to match keystroke features [13, 16]. The squared Mahalanobis distance between two feature vectors x and y is defined as:

$$\|x - y\|^2 = (x - y)^T S^{-1} (x - y) \tag{1.1}$$

where S is the covariance matrix of the data. The Mahalanobis distance not only weights the distance calculation according to the statistical variation of each feature component, but also decouples the interactions between features based on their covariance matrix, providing a useful distance metric for feature comparisons in pattern analysis. In statistical literature, the Mahalanobis distance is related to the logarithmic likelihood under the assumption that the data follows a multivariate Gaussian distribution, which is a reasonable approximation for most practical data.

B. Manhattan Distance The Manhattan distance metric, also called L1 distance or city block distance, is defined as follows:

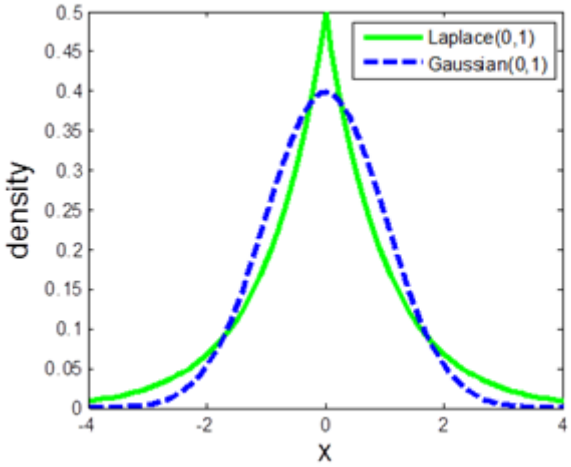

Figure 1.1: The probability density functions for univariate Laplace distribution and Gaussian distribution with mean 0 and variance 1. The Laplace distribution has fatter tails than the Gaussian distribution, and is therefore more tolerant to outliers..

$$\|x - y\|_1 = \sum_i |x_i - y_i| \tag{1.2}$$

The Manhattan distance has the advantages of simple computation and easy decomposition into contributions made by each variable. Most importantly, it is more robust to the influence of outliers when compared to higher order distance metrics including Euclidean distance and Mahalanobis distance. The Manhattan distance also has a statistical interpretation as the Mahalanobis distance does. It is related to the log likelihood of the multivariate Laplace distribution with an identity covariance matrix. The Laplace distribution is similar to the Gaussian distribution in that both are symmetric with one mode. However, the Laplace distribution has fatter tails than the Gaussian distribution (see Fig.(1.1)), and therefore it is more tolerant to outliers that significantly deviate from the mean. The Laplace distribution provides an attractive alternative to Gaussian distribution assumption for many real world data because of their heavy tails. Due to its robustness to outliers, the Manhattan distance has been used in keystroke biometrics research as well [1, 57].

A performance study of fourteen existing keystroke dynamics algorithms conducted by Killourhy and Maxion [62] indicates that the top performers are classifiers using scaled Manhattan distance [3], with an equal error rate (EER) of 0.096, and the nearest neighbor classifier using the Mahalanobis distance [16], with an EER of 0.10 on their keystroke dynamics benchmark dataset.

C. Decorrelated and Normalized Manhattan Distance The keystroke dynamics features consist of both dwell and latency timings, which exhibit large variations in individual components. The feature variables tend to interact with each other as well. These scale variations and feature correlations are handled well using the Mahalanobis distance metric. However, Mahalanobis distance is susceptible to the outliers in keystroke dynamics data caused by frequent pauses during typing. On the other hand, Manhattan distance is more robust to outliers, but it is not able to correct for the adverse interactions and redundancies between keystroke features. In summary, each of the two metrics, when used alone, has its respective advantages and limitations.

Zhong et al. have proposed a distance metric combining the benefits of both Mahalanobis distance and Manhattan distance [115]. First, the principle of Mahalanobis distance is applied to decorrelate and normalize the keystroke dynamics feature variables so that the covariance matrix of the transformed feature vectors becomes an identity matrix. This rectifying process is accomplished by applying the following linear transform to the keystroke dynamics input data:

$$x' = \Phi x \tag{1.3}$$

where $\Phi = S^{-1/2}$ is the inverse of the principle square root of the covariance matrix S such that $\Phi^T \Phi = S^{-1}$. With this transform, the data features become uncorrelated with identity variations in the feature variables. Once the data is normalized and decoupled, the Manhattan distance between two transformed data points x' and y' is then computed in a more standardized new feature space:

$$\|x - y\|' = \|x' - y'\|_1 = \left\| S^{-1/2} (x - y) \right\|_1 . \tag{1.4}$$

This distance metric not only ensures that the undesirable correlation and scale variations are accounted for, but also suppresses the influence of outliers for improved performance. As a result, the distance metric by Zhong et al. combines the benefits of both Mahalanobis and Manhattan distance metrics while overcoming their individual limitations. This distance metric was shown to outperform the other distance metrics on the CMU keystroke dataset [115].

1.2.2.2 Keystroke Dynamics Classification Using Statistical and Advanced Machine Learning Methods

Over the years, keystroke biometrics research has utilized many existing classification techniques. Both classical statistical methods [39, 46] and advanced machine learning approaches have been used, including K-Nearest Neighbor (KNN) classifiers [21, 112], K-means methods [59], Bayesian classifiers [81], Fuzzy logic [44], Boost learning [8], and Random Forests [8, 76], etc. Support vector machine (SVM) is a powerful machine learning method which computes decision boundaries by maximizing the margin in order to reduce the generalization error. With the kernel trick S is able to accommodate nonlinear decision boundaries for complex classification tasks. SVMs have been used to select effective features [111], and to classify feature patterns [75] in keystroke dynamics analysis.

Hidden Markov Model is a probabilistic model of a sequence of hidden variables with causal or conditional dependency. It is an appealing model for keystroke dynamics as it naturally represents the interconnections between consecutive features and noise in the data. Chen and Chang [20] and Jiang et al. [56] have respectively used Hidden Markov Models to learn the non-deterministic temporal dynamics in typing rhythms. A Gaussian mixture model approach has been used in [46] as well.

Neural networks provide a general learning paradigm for a variety of applications. They have been popularly used in keystroke dynamics in the past [16, 21, 44, 67, 68, 69, 72, 71]. However, traditional neural networks have their pitfalls that they may be trapped in local minimums during training to compromise the classification performance. Recent advances in deep learning have mitigated this risk in neural networks to allow well trained deep neural networks. Such deep learning approaches have achieved state of the art recognition performance for voice biometrics and many other object recognition applications. Deng and Zhong [27] applied the deep learning method to keystroke dynamics user authentication. Their study shows deep learning method significantly outperforms other algorithms on the CMU keystroke dynamics dataset [62].

1.3 Recent Trends

Keystroke dynamics biometrics is a rapidly growing field, driven by the high demand in secure access control for many applications, and adapting to fast evolving technology. In the following section, we review recent trends in keystroke dynamics biometrics research.

1.3.1 Keystroke Dynamics for Mobile Devices

These days, mobile devices are ubiquitous within our society. As they become an increasingly indispensable and important part of our everyday life, it is essential to ensure secure access of these devices as they store not only personal but oftentimes sensitive and even critical information [24, 86, 93]. Driven by this increasing demand, research work on keystroke dynamics biometrics on mobile devices has mushroomed in recent years [22, 42].

Compared to conventional keystroke dynamics on desktops or laptops, keystroke dynamics on mobile devices present many new challenges [14]. The keypads are much smaller on mobile devices due to their compact sizes. Many early mobile phones have used hardware keyboards with a reduced set of keys where each key is multiplexed for multiple characters. These hardware keyboards have since then been gradually replaced by virtual keyboards. These virtual keyboards have different layouts and responses to pressing from traditional keyboards. Keystroke dynamics on mobile devices is further complicated by new product features that were introduced to improve user experience, such as predictive text. All of the aforementioned factors can significantly affect a user's typing behavior compared to traditional desktop applications. For example, a user may change from a two-hand typing to single-hand or even single-finger typing. The limited computational resources within mobile devices impose additional

constraints on keystroke dynamics algorithms in those devices.

On the other hand, mobile devices are often embedded with a rich suite of advanced sensors. These sensors can capture additional data measurements during typing. A number of experiments have demonstrated that this additional data opens up abundant opportunities for boosting keystroke dynamics authentication accuracy in mobile devices.

Clarke and Furnell performed a feasibility study on keystroke-based user authentication on mobile phones. Key hold time and error rate (number of times pressing the backspace key) were used as features in their study. They achieved 12.8% EER using neural network classifiers on mobile handsets with a 12-key hardware keyboard [25, 23]. Maiorana et al. [73] also investigated the feasibility of using keystroke dynamics for user verification on mobile phones. They proposed a new statistical classifier which is computationally efficient for use in a mobile environment. They assessed the discriminative power of different subsets of keystroke timing features, and obtained an EER of 13.59%. Their study indicates that keystroke dynamics biometrics provides effective authentication in mobile devices, but needs supplementary features in order to facilitate a highly secure authentication scheme. Simple statistical methods were also employed by Campisi et al. [18] for keystroke dynamics biometrics on mobile phones. Buchoux and Clarke [17] studied various classifiers for keystroke analysis on smart phones, and their results suggest that statistical classifiers are the most effective given the trade-off between computational requirements and authentication accuracy. Zahid et al. investigated keystroke dynamics for mobile phones with numeric keyboards where each key is multiplexed for several characters [113]. They proposed four digraphs customized for these keyboards: a horizontal/vertical digraph which is the time to switch between keys horizontally/vertically, and a non-adjacent horizontal/vertical digraph which is the time to switch between non-adjacent keys horizontally/vertically. They demonstrated that these features, combined with the conventional key hold time and error correction rate, were capable of capturing user characteristics. Using particle swarm optimization (PSO) and genetic algorithm (GA) classifiers with these features, they obtained an average error rate of 2% FAR after the verification mode on a dataset containing 25 subjects. Hwang et al. [47] also concluded that keystroke dynamics can provide effective authentication for mobile devices. Furthermore, they proposed the use of artificial rhythms to improve the uniqueness and consistency of a user's keystroke signature, and therefore increase the authentication accuracy.

Trojahn and Ortmeier [101] compared keystroke dynamics performance using hardware keyboards and software keyboards on mobile phones during the login process. They found that despite small performance degradation, software keyboard input still provides feasible features for keystroke dynamics biometric authentication. Kambourakis et al. [58] proposed to enhance traditional keystroke dynamics features with additional data on the speed and distance of finger movement on smartphones with touchscreens. This upgrade resulted in an EER of 26% on a 10-digit PIN and an EER of 13.6% on short passphrases.

As the experiments have shown, although keystroke dynamics serves as an effective authentication scheme for mobile devices, it comes short of meeting strong security requirements when used as the sole criterion. However, there are many additional advanced sensors embedded in mobile devices which may be exploited for improved

authentication performance. These sensors can either facilitate more comprehensive keystroke characterization for augmented keystroke dynamics biometrics, or provide other biometric modalities that can be used in conjunction with keystroke dynamics.

Among these sensors are touchpad pressure sensors which measure finger pressure exerted on touchpad during typing events [65]. The pressure data were used for straightforward augmentation of keystroke dynamics biometrics for mobile devices with touch screens [1]. Saevanee and Bhatarakosol [90] explored the use of both finger pressures on touch pads and keystroke dynamics for user authentication. They found that finger pressure features are more discriminative than the conventional keying time features, and obtained an accuracy of 99% using finger pressure features with the PNN analytical method. Jain et al. [53, 55] also found that superior performance could be achieved by fusing touch screen features with conventional keystroke features. Chang et al. [19] proposed a graphical password interface with enlarged virtual keys for improved keystroke dynamics utility and authentication accuracy. They also examined using finger pressure features to enhance the authentication scheme. They demonstrated that by fusing pressure features with keystroke timing features, the EER of the keystroke dynamics based authentication system is reduced from 12.2% to 6.9% on a dataset containing 100 subjects and 20 imposters. Trojahn et al. [100] investigated combinations of keystroke time features for keystroke dynamics authentication in mobile devices. They also explored additional touch features such as finger pressure and the size of the key touch area for enhanced keystroke dynamics. They found that the additional touch features reduced more than 30% of the error in the timing feature-based keystroke authentication scheme using a dataset of 152 subjects.

Mobile devices are typically embedded with inertial sensors including accelerometers and gyroscopes which record the motion of the device. These motion characteristics have been exploited to improve the accuracy of keystroke dynamics biometrics for mobile devices [70]. Ho [45] explored the use of accelerometer statistics, key tap size, and key duration features to authenticate mobile device users during the login stage. The study showed that accelerometer statistics performed the best among the three feature types, while fusing the three feature types drastically improved the accuracy. Giuffrida et al. [42] used motion measurements from inertial sensors including accelerometers and gyroscopes to substantially boost keystroke dynamics authentication performance on mobile phones. In another study, [114] exploited four features extracted from sensors in touchscreen smartphones to fully characterize the keystroke dynamics: accelerations during key pressing, touching pressure, touching area, and key hold and inter key time. Experiments conducted using keystroke analysis on 4-digit and 8-digit PINs using a dataset containing more than 80 subjects yielded an EER of 3.65%. Trojahn and Ortmeier [102] fused keystroke dynamics biometrics with gait characteristics from gyroscopes for continuous authentication on mobile devices.

Keystroke dynamics biometrics for continuous mobile authentication has also been investigated. Feng et al. [33] studied mobile authentication for both the login and post login stages. They adopted text independent keystroke features comprising of keystroke time and tactile pressure from the capacitive touchscreen, with and without haptic feedback. Decision tree, random forest, and Bayes Net classification methods were used. Performance analysis on a dataset of 40 subjects indicated that adopting pressure information improved authentication accuracy, and typing with haptic feed-

back benefited the performance as well. Gascon et al. [36] performed a study on continuous authentication of mobile device users using typing motion behavior on a virtual keyboard. In addition to the typical keystroke time features, they also utilized data from the accelerometer, gyroscope, and orientation sensor to characterize the motion signature of the typing behavior. These features were collected for a predefined short text typed by 315 subjects. A 2376 dimensional feature vector encoded the statistics, and shape of motion measurements in both spatial and frequency domains were extracted to represent the typing motion behavior. SVM was then used to classify these high dimensional feature vectors.

1.3.2 Keystroke Dynamics as a Soft Biometrics

Soft biometric traits are "characteristics that provide some information about an individual but lack the distinctiveness and permanence to sufficiently differentiate any two individuals" [53], such as gender and race. Nonetheless, these traits could prove useful in improving the performance of person identification systems. There has been emerging research on keystroke dynamics as a soft biometrics for gender classification.

Fairhurst and Costa-Abreu [32] conducted experiments on keystroke dynamics for gender identification in social network environments in order to assess trust and reliability within online communications. Promising gender prediction results were achieved by directly adapting keystroke dynamics identification algorithms for gender prediction. They were able to obtain a 3% error rate for gender classification by fusing multiple classifiers. Giot and Rosenberger [41] investigated gender extraction using keystroke patterns and utilized the predicted gender to further enhance the accuracy of the keystroke dynamics authentication scheme. Using a support vector machine trained on keystroke time latency features with known gender labels, they were able to achieve a gender classification accuracy of 91% on the GREYC91 dataset (Section 1.4.2.2 [37]) of 7,555 samples from a total of 133 subjects. With a boost by the gender score computed from the keystroke latency features, they further improved the keystroke dynamics authentication accuracy from an EER of 10.65% to 8.45%, achieving a 20% error reduction. Most recently, keystroke dynamics has been studied to extract not only gender information, but also other soft biometric traits including age category, single or two-handed usage, and left or right-handedness [49, 48, 50]. Encouraging results using support vector machines have shown promise in keystroke biometrics as a soft biometrics when evaluated on the GREYC-NISLAB Keystroke Dynamics Soft Biometrics Dataset of 110 users with 100 sample passphrases per user (Section 1.4.2.11).

1.3.3 Keystroke Dynamics for Emotion Recognition

As a behavioral biometrics, keystroke dynamics is influenced by the emotion or mental state of the user. This has motivated research on stress detection and affective computing using keystroke dynamics analysis [5, 61, 63, 64, 82, 116]. Epp et al. also conducted analysis on emotional states using keystroke dynamics features such as digraphs [31]. The study performed by Tsihrintzis et al. [103] suggested that keystroke information can significantly boost emotion recognition using visual-facial

features. Bixler and D'Mello [12] explored combining keystroke timing statistics with task appraisals and stable traits to detect a user's affective state or emotion in order to enhance the human/computer interaction experience. By analyzing keystroke timing information and other linguistic features using advanced machine learning methods including decision trees, support vector machines, boost learning and neural networks, it is possible to achieve accurate recognition of both cognitive and physical stress conditions comparable to algorithms using other effective computing methods [107].

1.4 Evaluation of Keystroke Dynamics Biometric Systems

With the growing interest in keystroke dynamics research, it is fundamental to establish standardized test beds and performance metrics in order to compare various new algorithms and assess progress in the field. In this section, we review existing performance metrics and list publicly available keystroke dynamics datasets that can be used for continued performance evaluations and comparisons.

1.4.1 Performance Metrics

Two common metrics used to assess the performance of a keystroke biometric system are:

1. False acceptance rate (FAR), sometimes known as false match rate (FMR), is the probability that a system incorrectly classifies an imposter as a genuine user. It measures the percent of imposters that are incorrectly accepted as genuine users, i.e., how often an intruder is granted access.

2. False rejection rate (FRR), also known as false none match rate (FNMR), is the probability a system incorrectly classifies a genuine user as an imposter. It measures the percent of genuine users that are rejected as imposters.

For a secure system, both error rates need to be small.

Since different values in the operating threshold may result in varying values of FRR and FAR, the receiver operating characteristic (ROC) curve, i.e. the graphical plot of FAR against FRR for the whole range of threshold settings, is often used to illustrate the comprehensive performance of an algorithm. ROC curves from different algorithms are also plotted against each other for performance comparison. The ROC curve is sometimes known as a decision error trade-off (DET) curve. A special point on the ROC curve, where the FAR equals FRR, known as equal error rate (EER) or crossover error rate (CER), is typically used as a performance metric. In general, the system with the lowest EER is the most accurate. For certain applications which have zero tolerance for false rejection, the zero-miss false alarm rate (ZMFAR), which is the minimum false alarm rate when the miss rate is zero, is used. For other systems which suffer large losses for admitting intruders, the zero false acceptance rate, i.e., the FRR when FAR equals zero, is used as well.

Alternatively, accuracy and error rate have been used to measure authentication and identification performance. The accuracy is the percentage of correct classifications performed by the system.

In addition to these traditional performance metrics, new performance metrics have been proposed to address the challenges in continuous authentication, which demand accurate and timely detection of imposters along with minimum annoyances from false rejections of the genuine user. In an example of continuous authentication, Bours [15] proposed a trust model to continuously monitor a user's credibility, and only reject a user when the trust drops below a certain threshold from consistent and prolonged incorrect typing behavior. Two new performance metrics, called average number of impostor actions (ANIA) and average number of genuine actions (ANGA), have been used as performance indicators for continuous authentication. These metrics measure how much an impostor can type before he or she gets rejected and how much a genuine user can type before wrongfully locked out of the system.

1.4.2 Benchmark Keystroke Datasets for Performance Analysis and Comparison

Despite promising performances by keystroke dynamics authentication algorithms reported over the years, such results were often achieved on proprietary datasets that remain unavailable to the research community. As a result, it has not been possible to make a sound comparison of different algorithms because of the use of separate datasets and evaluation criteria across studies. To address this issue, keystroke dynamics databases, including benchmark results of popular keystroke biometrics algorithms, have been published in recent years to provide a standard experimental platform for assessing progress. These keystroke datasets, along with accompanying evaluation methodologies and performance studies of existing algorithms, provide multiple benchmarks to objectively gauge the progress of new keystroke biometrics algorithms.

1.4.2.1 CMU Static Keystroke Dynamics Benchmark Dataset [62]

The CMU keystroke dynamics benchmark dataset contains keystroke dynamics consisting of the dwell time for each key and the latencies between two successive keys for the static password string ".tie5Roanl". There are 51 subjects in the dataset. For each subject, there are eight data collection sessions with at least one day intervals between sessions. A total of 50 feature vectors were extracted in each session, resulting in a total of 400 feature vectors for each subject. The same publication includes the performances of fourteen existing keystroke dynamics algorithms on this dataset, including Neural Networks [21, 44], K-means [59], Fuzzy Logic [44], KNNs, Outlier Elimination [44], SVMs [111], etc. Various distance metrics, including Euclidean distance, Manhattan distance [13], and Mahalanobis distance [13] were used. This dataset has also been used to evaluate more recent algorithms [27, 115].

1.4.2.2 GREYC09 Static Keystroke [13] Dynamics Benchmark Dataset [37]

The GREYC Keystroke Benchmark [37] contains static typing rhythms for the fixed password "greyc laboratory" collected from 100 subjects over a duration of two months.

Two keyboards were used for data collection. The dataset includes 60 samples from each subject resulting in a total of 7555 captures with an average of 51 captures per subject.

1.4.2.3 GREYC12 Static Keystroke Dynamics Benchmark Dataset [39]

The GREYC'12 Keystroke Benchmark mimics the realistic scenario of keystroke dynamics based authentication in user logins. The data was collected in a web-based unconstrained environment. Unlike the dataset in Section 1.4.2.2 where all users typed a fixed passphrase, each user typed the passphrase of his/her own choice in addition to an imposed passphrase. The dataset contains 83 subjects with a total of 5,185 genuine samples, 5,754 imposter samples, and 5,439 imposed samples. Authentication performances determined by applying statistical analysis to multiple latency features were published along with the dataset as a baseline for performance comparison.

1.4.2.4 Queen Mary University Keystroke 100 benchmark dataset [72]

This dataset is a static dataset containing keystroke data from 100 users typing the password "try4-mbs". In addition to typing latency timing data, the set also includes the time series of typing pressure exerted, measured in volts. A total of 10 typing samples were collected from each subject. A number of keystroke authentication techniques including SVM and ARTMAP have been evaluated on this dataset [72, 71].

1.4.2.5 Si6 Labs Keystroke rhythm Dataset [9]

This dataset includes typing rhythm data from volunteer web users collected via a web application. It contains keystroke dynamics of a set of 20 long sentences from a total of 63 subjects typing during 66 sessions. The input language for this dataset is Spanish.

1.4.2.6 Beihang University Static Keystroke Dynamics Benchmark Dataset [68, 69]

The Beihang Keystroke Dynamics database boasts realistic keystroke dynamics collection using a commercialized system. This database includes 2057 test and training samples of user names and passwords from 117 subjects. This database contains two subsets: one collected in a cybercafé environment and the other collected online. The performances of three keystroke dynamics approaches, namely Nearest Neighbor method, Gaussian model, and OC-SVM, have been reported on this database.

1.4.2.7 University of Torino Free Text Keystroke Dataset [43]

This dataset contains free text samples from forty volunteers acting as genuine users, each with 15 typing samples, and from 165 volunteers acting as imposters, each with one typing sample. Each sample contains the time when a key is pressed along with the key's value. On average, the text samples contain between 700 and 900 characters. All subjects were native Italian speakers and all samples were written in Italian. The

data was collected during a period of six months. This dataset is available by sending a request to its authors.

1.4.2.8 Clarkson University Mixed Keystroke Dynamics Dataset [108]

The Clarkson University Keystroke Dynamics dataset contains keystroke data of short pass-phrases, fixed text (transcriptions of long proses), and free text. Two data collection sessions, each about one hour long, were conducted on two separate days with 39 subjects. Performance results using two existing algorithms [43, 66] were published together with the dataset. In addition to keystroke data, videos of facial expressions and hand movements of the users were also captured and included in the dataset.

1.4.2.9 TapDynamics Keystroke Dataset from Mobile Phones [45]

This dataset is collected using an android login application on mobile phones. It collects the following data during the login session when a user enters the PIN: the duration of each key tap, the latency between each key tap, the size of each key tap, and all accelerometer readings over the course of a login attempt. The dataset is obtained from 55 subjects, each with about 30 samples, resulting in 1704 data samples. The PIN code for each subject is randomly assigned from five prespecified codes. This dataset is accessible from https://github.com/grantho/TapDynamics/.

1.4.2.10 Graphical Password Keystroke+Pressure Mobile Dataset [19]

This dataset contains the keystroke dynamics data and touch pressure data for graphical passwords collected from 100 subjects using two touch screen mobile devices: a Motorola Desire HD with a 4.3 inch screen and a Viewsonic Viewpad with a 10.1 inch screen. The graphical password enlarges the virtual key size for improved keystroke dynamics utility. Each subject chose his or her password of choice during enrollment where five samples were collected for each subject to build the classifier. An additional five samples were collected over a following period of five weeks for testing. 10 people chosen as imposters were given the passwords for the 100 legitimate users and five samples were collected per password for each imposter, resulting in a total of 5,000 imposter samples. Performance results using the keystroke timing features, as well as using both finger pressure and keystroke timing features, were published along with the dataset. This dataset is accessible from `http://ty.ncue.edu.tw/N27/data.html`.

1.4.2.11 GREYC-NISLAB Keystroke Dynamics Soft Biometrics Dataset [48]

This dataset contains both keystroke dynamics data and additional categorical information on gender, age, left or right-handedness, and typing method (one-handed typing vs. two-handed typing). Keystroke timing information on five passphrases was collected from a total of 110 subjects. There were 10 repetitions per phrase per subject for each way of typing. Some performance analysis on the dataset was published in [49, 48]. This dataset can be accessed from the following link: `http://www.epaymentbiometrics.ensicaen.fr/images/pdf/greyc-nislab%20keystroke%20benchmark%20dataset.xls`.

1.5 Conclusions

Keystroke dynamics facilitates a natural and cost effective way for security and access protection of computers and mobile devices. It also allows for continuous authentication by monitoring a user's typing behavior during the entire login session without any interruption to the user's routine work. The use of keyboards for personal identification had been studied even before personal computers were introduced [96]. It has been attracting increasing attention and interests as our increasing dependency on computers and mobile devices to store private and sensitive information demands strong security protection. Despite decades of research, keystroke dynamics research is still evolving with many open challenges.

In this survey we review the literature of keystroke dynamic biometrics. We discuss recent advances and new trends in keystroke dynamics research. Echoing the sentiment on a lack of common evaluation framework [8, 60] in the field, we compile a list of publicly available keystroke datasets. We would like to note that despite the available datasets, we are still in need of large standard keystroke databases for the research community. The desirable databases should reflect data variations in realistic applications, including number and diversity of subjects, size of vocabulary, span of sessions across a prolonged period of time, etc.

Keystroke dynamics has unmatched usability and tremendous potential for cyber security applications. This research field faces the challenge common to all other biometric modalities such as fingerprints [74] and face recognition, that is, how to perform robustly in real world scenarios with the presence of various variations [54, 95]. New feature extraction and classification methods are still in demand. Fusion of keystroke biometrics with other biometric modalities will provide the ultimate comprehensive and secure authentication solution [4, 40].

References

[1] J.D. Allen. An analysis of pressure-based keystroke dynamics algorithms. Master's thesis, Southern Methodist University, Dallas, TX, U.S.A., 2010.

[2] A. Alsultan and K. Warwick. Keystroke dynamics authentication: a survey of free-text methods. *International Journal of Computer Science Issues*, 10(4):1–10, 2013.

[3] L.C.F. Araujo, L.H.R. Sucupira Jr., M.G. Lizarraga, L.L. Ling, and J.B.T. Yabu-Uti. User authentication through typing biometrics features. In *1st International Conference on Biometric Authentication (ICBA)*, volume 3072 of *LNCS*, pages 694–700, 2004.

[4] K.O. Bailey, J.S. Okolica, and G.L. Peterson. User identification and authentication using multi-modal behavioral biometrics. *Computers & Security*, 43:77–89, 2014.

[5] K. Bakhtiyari, M. Taghavi, and H. Husain. Implementation of emotional-aware computer systems using typical input devices. In *6th Asian Conference on Intelligent Information and Database Systems (ACIIDS)*, volume 8397 of *LNCS*, pages 364–374, 2014.

[6] K.S. Balagani, Vir V. Phoha, A.Ray, and S. Phoha. On the discriminability of keystroke feature vectors used in fixed text keystroke authentication. *Pattern Recognition Letters*, 32(7):1070–1080, 2011.

[7] S.P. Banerjee and D.L. Woodard. Biometric authentication and identification using keystroke dynamics: a survey. *Journal of Pattern Recognition Research*, 7(1):116–139, 2012.

[8] N. Bartlow and B. Cukic. Evaluating the reliability of credential hardening through keystroke dynamics. In *17th International Symposium on Software Reliability Engineering (ISSRE)*, pages 117–126, 2006.

[9] L. Bello, M. Bertacchini, C. Benitez, J. Pizzoni, and M. Cipriano. Collection and publication of a fixed text keystroke dynamics dataset. In *XVI Congreso Argentino de Ciencias de la Computación (CACIC)*, pages 822–831, 2010.

[10] F. Bergadano, D. Gunetti, and C. Picardi. User authentication through keystroke dynamics. *ACM Transactions on Information and System Security*, 5(4):367–397, 2002.

[11] S. Bhatt and T. Santhanam. Keystroke dynamics for biometric authentication - a survey. In *International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME)*, pages 17–23, 2013.

[12] R. Bixler and S. D'Mello. Detecting boredom and engagement during writing with keystroke analysis, task appraisals, and stable traits. In *International Conference on Intelligent User Interfaces (IUI)*, pages 225–234, 2013.

[13] S. Bleha, C. Slivinsky, and B. Hussien. Computer-access security systems using keystroke dynamics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(12):1217–1222, 1990.

[14] R. Botha, S.M. Furnell, and N.L. Clarke. From desktop to mobile: examining the security experience. *Computers & Security*, 28(3-4):130–137, 2009.

[15] P. Bours. Continuous keystroke dynamics: a different perspective towards biometric evaluation. *Information Security Technical Report*, 17(1-2):36–43, 2012.

[16] M. Brown and S.J. Rogers. User identification via keystroke characteristics of typed names using neural networks. *International Journal of Man-Machine Studies*, 30(6):999–1014, 1993.

[17] A. Buchoux and N.L. Clarke. Deployment of keystroke analysis on a smartphone. In *Australian Information Security Management Conference*, 2008. p.48.

[18] P. Campisi, E. Maiorana, M. Lo Bosco, and A. Neri. User authentication using keystroke dynamics for cellular phones. *IET Signal Processing*, 3(4):333–341, 2009.

[19] T. Chang, C. Tsai, and J. Lin. A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices. *Journal of Systems and Software*, 85(5):1157–1165, 2012.

[20] W. Chen and W. Chang. Applying hidden Markov models to keystroke pattern analysis for password verification. In *IEEE International Conference on Information Reuse and Integration (IRI)*, pages 467–474, 2004.

[21] S. Cho, C. Han, D.H. Han, and H.I. Kim. Web-based keystroke dynamics identity verification using neural network. *Journal of Organizational Computing and Electronic Commerce*, 10(4):295–307, 2000.

[22] N.L. Clarke and S.M. Furnell. Authentication of users on mobile telephones – a

survey of attitudes and practices. *Computers & Security*, 24(7):519–527, 2005.

[23] N.L. Clarke and S.M. Furnell. Authenticating mobile phone users using keystroke analysis. *International Journal of Information Security*, 6(1):1–14, 2006.

[24] N.L. Clarke, S.M. Furnell, B.M. Lines, and P.L. Reynolds. Keystroke dynamics on a mobile handset: a feasibility study. *Information Management & Computer Security*, 11(4):161–166, 2003.

[25] N.L. Clarke, S.M. Furnell, P.M. Rodwell, and P.L. Reynolds. Acceptance of subscriber authentication methods for mobile telephony devices. *Computers & Security*, 21(3):220–228, 2002.

[26] H. Crawford. Keystroke dynamics: characteristics and opportunities. In *8th Annual International Conference on Privacy Security and Trust (PST)*, pages 205–212, 2010.

[27] Y. Deng and Y. Zhong. Keystroke dynamics user authentication based on Gaussian mixture model and deep belief nets. *ISRN Signal Processing*, 2013, Article ID 565183, 7 pages, 2013.

[28] I. Deutschmann, P. Nordstrom, and L. Nilsson. Continuous authentication using behavioral biometrics. *IT Professional*, 15(4):12–15, 2013.

[29] S. Douhou and J.R. Magnus. The reliability of user authentication through keystroke dynamics. *Statistica Neerlandica*, 63(4):432–449, 2009.

[30] A. Dvorak, N. Merrick, W. Dealey, and G. Ford. *Typewriting Behavior*. American Book Company, New York, USA, 1936.

[31] C. Epp, M. Lippold, and R.L. Mandryk. Identifying emotional states using keystroke dynamics. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 715–724, 2011.

[32] M. Fairhurst and M. Da Costa-Abreu. Using keystroke dynamics for gender identification in social network environment. In *4th International Conference on Imaging for Crime Detection and Prevention (ICDP)*, pages 1–6, 2011.

[33] T. Feng, X. Zhao, B. Carbunar, and W. Shi. Continuous mobile authentication using virtual key typing biometrics. In *12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pages 1547–1552, 2013.

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

[35] R. Gaines, W. Lisowski, S. Press, and N. Shapiro. Authentication by keystroke timing: some preliminary results. Technical Report Rand Rep. R-2560-NSF, RAND Corporation, 1980.

[36] H. Gascon, S. Uellenbeck, C. Wolf, and K. Rieck. Continuous authentication on mobile devices by analysis of typing motion behavior. In *GI Conference Sicherheit (Sicherheit, Schutz und Verlasslichkeit)*, 2014.

[37] R. Giot, M. El-Abed, and C. Rosenberger. GREYC keystroke: a benchmark for keystroke dynamics biometric systems. In *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*, pages 1–6, 2009.

[38] R. Giot, M. El-Abed, and C. Rosenberger. *Biometrics*, chapter Keystroke Dynamics Authentication, pages 157–182. InTech, 2011.

[39] R. Giot, M. El-Abed, and C. Rosenberger. Web-based benchmark for keystroke

dynamics biometric systems: a statistical analysis. In *Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, pages 11–15, 2012.

[40] R. Giot, B. Hemery, and C. Rosenberger. Low cost and usable multimodal biometric system based on keystroke dynamics and 2D face recognition. In *20th International Conference on Pattern Recognition (ICPR)*, pages 1128–1131, 2010.

[41] R. Giot and C. Rosenberger. A new soft biometric approach for keystroke dynamics based on gender recognition. *International Journal of Information Technology and Management*, 11(1-2):35–49, 2012.

[42] C. Giuffrida, K. Majdanik, M. Conti, and H. Bos. I sensed it was you: authenticating mobile users with sensor-enhanced keystroke dynamics. In *Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA)*, volume 8550 of *Lecture Notes in Computer Science*, pages 92–111, 2014.

[43] D. Gunetti and C. Picardi. Keystroke analysis of free text. *ACM Transactions on Information and System Security*, 8(3):312–347, 2005.

[44] S. Haider, A. Abbas, and A.K. Zaidi. A multi-technique approach for user identification through keystroke dynamics. In *IEEE International Conference on Systems, Man, and Cybernetics (ICSMC)*, volume 2, pages 1336–1341, 2000.

[45] G. Ho. Tapdynamics: strengthening user authentication on mobile phones with keystroke dynamics. Technical report, Stanford University, 2014.

[46] D. Hosseinzadeh and S. Krishnan. Gaussian mixture modeling of keystroke patterns for biometric applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 38(6):816–826, 2008.

[47] S. Hwang, S. Cho, and S. Park. Keystroke dynamics-based authentication for mobile devices. *Computers & Security*, 28(1-2):85–93, 2009.

[48] S.Z.S. Idrus, E. Cherrier, C. Rosenberger, and P. Bours. Soft biometrics database: a benchmark for keystroke dynamics biometric systems. In *International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–8, 2013.

[49] S.Z.S. Idrus, E. Cherrier, C. Rosenberger, and P. Bours. Soft biometrics for keystroke dynamics. In *10th International Conference on Image Analysis and Recognition (ICIAR)*, volume 7950 of *LNCS*, pages 11–18, 2013.

[50] S.Z.S. Idrus, E. Cherrier, C. Rosenberger, and P. Bours. Soft biometrics for keystroke dynamics: profiling individuals while typing passwords. *Computers & Security*, 45:147–155, 2014.

[51] A. K. Jain, R. Bolle, and S. Pankanti. *Biometrics: Personal Identification in Networked Society*. Kluwer Academic Publishers, 1999.

[52] A. K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

[53] A.K. Jain, S.C. Dass, and K. Nandakumar. Soft biometric traits for personal recognition systems. In *1st International Conference on Biometric Authentication (ICB)*, volume 3072 of *LNCS*, pages 731–738, 2004.

[54] A.K. Jain, S. Pankanti, S. Prabhakar, L. Hong, and A. Ross. Biometrics: a grand challenge. In *17th International Conference on Pattern Recognition (ICPR)*,

volume 2, pages 935–942, 2004.

[55] L. Jain, J.V. Monaco, M.J. Coakley, and C.C. Tappert. Passcode keystroke biometric performance on smartphone touchscreens is superior to that on hardware keyboards. *International Journal of Research in Computer Applications & Information Technology*, 2(4):29–33, 2014.

[56] C. Jiang, S. Shieh, and J. Liu. Keystroke statistical learning model for web authentication. In *2nd ACM Symposium on Information, Computer and Communications Security (ASIACCS)*, pages 359–361, 2007.

[57] R. Joyce and G. Gupta. Identity authentication based on keystroke latencies. *Communications of the ACM*, 33(2):168–176, 1990.

[58] G. Kambourakis, D. Damopoulos, D. Papamartzivanos, and E. Pavlidakis. Introducing touchstroke: keystroke-based authentication system for smartphones. *Security and Communication Networks*, 2014. in press.

[59] P. Kang, S. Hwang, and S. Cho. Continual retraining of keystroke dynamics based authenticator. In *International Conference on Advances in Biometrics (ICB)*, volume 4642 of *LNCS*, pages 1203–1211, 2007.

[60] M. Karnan, M. Akila, and N. Krishnaraj. Biometric personal authentication using keystroke dynamics: a review. *Applied Soft Computing*, 11(2):1565–1573, 2011.

[61] P. Khanna and M. Sasikumar. Recognising emotions from keyboard stroke pattern. *International Journal of Computer Applications*, 11(9):1–5, 2010.

[62] K.S. Killourhy and R.A. Maxion. Comparing anomaly detectors for keystroke dynamics. In *39th Annual International Conference on Dependable Systems and Networks*, 2009.

[63] A. Kirke, M. Bonnot, and E. Miranda. Towards using expressive performance algorithms for typist emotion detection. Ann Arbor, MI: MPublishing, University of Michigan Library, 2011.

[64] A. Kolakowska. A review of emotion recognition methods based on keystroke dynamics and mouse movements. In *6th International Conference on Human System Interaction (HSI)*, pages 548–555, 2013.

[65] K. Kotani and K. Horii. Evaluation on a keystroke authentication system by keying force incorporated with temporal characteristics of keystroke dynamics. *Behaviour & Information Technology*, 24(4):289–302, 2005.

[66] J. Leggett and G. Williams. Verifying identity via keystroke characterstics. *International Journal of Man-Machine Studies*, 28(1):67–76, 1988.

[67] J. Leggett, G. Williams, M. Usinick, and M. Longnecker. Dynamic identity verification via keystroke characteristics. *International Journal of Man-Machine Studies*, 35(6):859–870, 1991.

[68] D.T. Li. Computer-access authentication with neural network based keystroke identity verification. In *International Conference on Neural Networks*, volume 1, pages 174–178, 1997.

[69] Y. Li, B. Zhang, Y. Cao, S. Zhao, Y. Gao, and J. Liu. Study on the BeiHang keystroke dynamics database. In *International Joint Conference on Biometrics (IJCB)*, pages 1–5, 2011.

[70] M. Lopatka and M.H. Peetz. Vibration sensitive keystroke analysis. In *18th Annual Belgian-Dutch Conference on Machine Learning*, pages 75–80, 2009.

[71] C.C. Loy, W.K. Lai, and C.P. Lim. Keystroke patterns classification using the

ARTMAP-FD neural network. In *3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP)*, volume 1, pages 61–64, 2007.

[72] C.C. Loy, C.P. Lim, and W.K. Lai. Pressure-based typing biometrics user authentication using the fuzzy ARTMAP neural network. In *International Conference on Neural Information Processing (ICONIP)*, 2005.

[73] E. Maiorana, P. Campisi, N. Gonzalez-Carballo, and A. Neri. Keystroke dynamics authentication for mobile phones. In *ACM Symposium on Applied Computing (SAC)*, pages 21–26, 2011.

[74] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer, NY, 2003.

[75] W. Martono, H. Ali, and M.J.E. Salami. Keystroke pressure-based typing biometrics authentication system using support vector machines. In *International Conference on Computational Science and Its Applications (ICCSA)*, volume 4706 of *LNCS*, pages 85–93, 2007.

[76] R.A. Maxion and K.S. Killourhy. Keystroke biometrics with number-pad input. In *IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 201–210, 2010.

[77] D. El Menshawy, H.M.O. Mokhtar, and O. Hegazy. A keystroke dynamics based approach for continuous authentication. In *10th International Conference on Beyond Databases, Architectures, and Structures (BDAS)*, volume 424 of *CCIS*, pages 415–424, 2014.

[78] A. Messerman, T. Mustafic, S.A. Camtepe, and S. Albayrak. Continuous and non-intrusive identity verification in real-time environments based on free-text keystroke dynamics. In *International Joint Conference on Biometrics (IJCB)*, pages 1–8, 2011.

[79] J.V. Monaco, N. Bakelman, S.H. Cha, and C.C. Tappert. Developing a keystroke biometric system for continual authentication of computer users. In *European Intelligence and Security Informatics Conference (EISIC)*, pages 210–216, 2012.

[80] J.V. Monaco, N. Bakelman, S.H. Cha, and C.C. Tappert. Recent advances in the development of a long-text-input keystroke biometric authentication system for arbitrary text input. In *European Intelligence and Security Informatics Conference (EISIC)*, pages 60–66, 2013.

[81] F. Monrose and A.D. Rubin. Keystroke dynamics as a biometric for authentication. *Future Generation Computing Systems*, 16(4):351–359, 2000.

[82] A.F.M.N.H. Nahin, J.M. Alam, H. Mahmud, and K. Hasan. Identifying emotion by keystroke dynamics and text pattern analysis. *Behaviour & Information Technology*, 33(9):987–996, 2014.

[83] B. Ngugi, B.K. Kahn, and M. Tremaine. Typing biometrics: impact of human learning on performance quality. *Journal of Data and Information Quality*, 2(2):Article No. 11, 2011.

[84] M.S. Obaidat and B. Sadoun. Verification of computer users using keystroke dynamics. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 27(2):261–269, 1997.

[85] A. Peacock, X. Ke, and M. Wilkerson. Typing patterns: a key to user identification. *IEEE Security and Privacy*, 2(5):40–47, 2004.

[86] M. La Polla, F. Martinelli, and D. Sgandurra. A survey on security for mobile devices. *IEEE Communications Surveys & Tutorials*, 15(1):446–471, 2013.

[87] S. Prabhakar, S. Pankanti, and A.K. Jain. Biometric recognition: security and privacy concerns. *IEEE Security & Privacy*, 1(2):33–42, 2003.

[88] J.A. Robinson, V.W. Liang, J.A.M. Chambers, and C.L. MacKenzie. Computer user verification using login string keystroke dynamics. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 28(2):236–241, 1998.

[89] J. Roth, X. Liu, A. Ross, and D. Metaxas. Investigating the discriminative power of keystroke sound. *IEEE Transactions on Information Forensics and Security*. in press.

[90] H. Saevanee and P. Bhatarakosol. User authentication using combination of behavioral biometrics over the touchpad acting like touch screen of mobile device. In *International Conference on Computer and Electrical Engineering (ICCEE)*, pages 82–86, 2008.

[91] D. Shanmugapriya and G. Padmavathi. A survey of biometric keystroke dynamics: approaches, security and challenges. *International Journal of Computer Science and Information Security*, 5(1):115–119, 2009.

[92] S.J. Shepherd. Continuous authentication by analysis of keyboard typing characteristics. In *European Convention on Security and Detection*, pages 111–114, 1995.

[93] E. Shi, Y. Niu, M. Jakobsson, and R. Chow. Implicit authentication through learning user behavior. In *13th International Conference on Information Security (ISC)*, volume 6531 of *LNCS*, pages 99–113, 2010.

[94] T. Sim and R. Janakiraman. Are digraphs good for free-text keystroke dynamics? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–6, 2007.

[95] E. Al Solami, C. Boyd, A. Clark, and A.K. Islam. Continuous biometric authentication: can it be more practical? In *12th IEEE International Conference on High Performance Computing and Communications (HPCC)*, pages 647–652, 2010.

[96] R. Spillane. Keyboard apparatus for personal identification. *IBM Technical Disclosure Bulletin*, 17(11), 1975.

[97] Z. Syed, S. Banerjee, and B. Cukic. Leveraging variations in event sequences in keystroke-dynamics authentication systems. In *15th International Symposium on High-Assurance Systems Engineering (HASE)*, pages 9–16, 2014.

[98] P.S. Teh, A.B.J. Teoh, and S. Yue. A survey of keystroke dynamics biometrics. *The Scientific World Journal*, Article ID 408280, 2013.

[99] C.M. Tey, P. Gupta, K. Muralidharan, and D. Gao. Keystroke biometrics: the user perspective. In *4th ACM Conference on Data and Application Security and Privacy (CODASPY)*, pages 289–296, 2014.

[100] M. Trojahn, F. Arndt, and F. Ortmeier. Authentication with keystroke dynamics on touchscreen keypads-effect of different N-Graph combinations. In *3rd International Conference on Mobile Services, Resources, and Users (MOBILITY)*, pages 114–119, 2013.

[101] M. Trojahn and F. Ortmeier. Biometric authentication through a virtual key-

board for smartphones. *International Journal of Computer Science & Information Technology*, 4(5):1–12, 2012.

[102] M. Trojahn and F. Ortmeier. Keygait: framework for continuously biometric authentication during usage of a smartphone. In *3rd International Conference on Mobile Services, Resources, and Users (MOBILITY)*, pages 30–33, 2013.

[103] G.A. Tsihrintzis, M. Virvou, E. Alepis, and I.O. Stathopoulou. Towards improving visual-facial emotion recognition through use of complementary keyboard-stroke pattern information. In *5th International Conference on Information Technology: New Generations (ITNG)*, pages 32–37, 2008.

[104] U. Uludag, S. Pankanti, S. Prabhakar, and A.K. Jain. Biometric cryptosystems: issues and challenges. *Proceedings of the IEEE*, 92(6):948–960, 2004.

[105] D. Umphress and G. Williams. Identity verification through keyboard characteristics. *International Journal of Man-Machine Studies*, 23(3):263–273, 1985.

[106] J.A. Unar, W.C. Seng, and A. Abbasi. A review of biometric technology along with trends and prospects. *Pattern Recognition*, 47(8):2673–2688, 2014.

[107] L.M. Vizer, L. Zhou, and A. Sears. Automated stress detection using keystroke and linguistic features: an exploratory study. *International Journal of Human-Computer Studies*, 67(10):870–886, 2009.

[108] E. Vural, J. Huang, D. Hou, and S. Schuckers. Shared research dataset to support development of keystroke authentication. In *International Joint Conference on Biometrics (IJCB)*, 2014.

[109] J.D. Woodward, N.M. Orlans, and P.T. Higgins. *Biometrics: Identity Assurance in the Information Age*. McGraw-Hill, New York, USA, 2002.

[110] R.V. Yampolskiy and V. Govindaraju. Behavioral biometrics: a survey and classification. *International Journal of Biometrics*, 1(1):81–113, 2008.

[111] E. Yu and S. Cho. GA-SVM wrapper approach for feature subset selection in keystroke dynamics identity verification. In *International Joint Conference on Neural Networks (IJCNN)*, volume 3, pages 2253–2257, 2003.

[112] R.S. Zack, C.C. Tappert, and S.H. Cha. Performance of a long-text-input keystroke biometric authentication system using an improved k-nearest-neighbor classification method. In *4th IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pages 1–6, 2010.

[113] S. Zahid, M. Shahzad, S.A. Khayam, and M. Farooq. Keystroke-based user identification on smart phones. In *12th International Symposium RAID*, volume 5758 of *LNCS*, pages 224–243, 2009.

[114] N. Zheng, K. Bai, H. Huang, and H. Wang. You are how you touch: user verification on smartphones via tapping behaviour. Technical report, College of William & Mary, 2012.

[115] Y. Zhong, Y. Deng, and A.K. Jain. Keystroke dynamics for user authentication. In *IEEE Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 117–123, 2012.

[116] P. Zimmermann, S. Guttormsen, B. Danuser, and P. Gomez. Affective computing–a rationale for measuring mood with mouse and keyboard. *International Journal of Occupational Safety and Ergonomics*, 9(4):539–551, 2003.