Multi-modal Behavioural Biometric Authentication for Mobile Devices

Hataichanok Saevanee[1], Nathan L. Clarke[1,2], and Steven M. Furnell[1,2]

[1] Centre for Security, Communications and Network Research, University of Plymouth, Plymouth, United Kingdom
[2] School of Computer and Information Science, Edith Cowan University, Perth, Western Australia
info@cscan.org

Abstract. The potential advantages of behavioural biometrics are that they can be utilised in a transparent (non-intrusive) and continuous authentication system. However, individual biometric techniques are not suited to all users and scenarios. One way to increase the reliability of transparent and continuous authentication systems is create a multi-modal behavioural biometric authentication system. This research investigated three behavioural biometric techniques based on SMS texting activities and messages, looking to apply these techniques as a multi-modal biometric authentication method for mobile devices. The results showed that behaviour profiling, keystroke dynamics and linguistic profiling can be used to discriminate users with overall error rates 20%, 20% and 22% respectively. To study the feasibility of multi-modal behaviour biometric authentication system, matching-level fusion methods were applied. Two fusion methods were utilised: simple sum and weight average. The results showed clearly that matching-level fusion can improve the classification performance with an overall EER 8%.

Keywords: Behavioural Biometrics, Authentication, Mobile Devices, Behavioural Profiling, Keystroke Dynamics, Linguistic Profiling.

1 Introduction

Mobile devices, such as cellular phones and Personal Digital Assistants (PDAs) are rapidly evolving technologies capable of providing many services through a wide range of applications over multiple networks such as the Internet (e.g. e-mail's and online banking), entertainment (e.g. photos and video games) and the sharing of data (via Bluetooth, laptop/computer). The plethora of functionalities offered by mobile devices enables users to store increasing amounts of wide ranging types of information from business to personal and sensitive data. With this in mind, previous research [1] highlights mobile users concerns of their devices being lost or stolen.

Many authentication mechanisms have been developed for mobile devices with the aim of providing a greater level of security for the end user. Biometric authentication is commonly acknowledged as a reliable solution which provides enhanced authentication over the traditional password ("something you know") and token

D. Gritzalis, S. Furnell, and M. Theoharidou (Eds.): SEC 2012, IFIP AICT 376, pp. 465–474, 2012.
© IFIP International Federation for Information Processing 2012

("something you have") approaches. Biometric characteristics are uniquely individual ("something you are"), non-transferable to others, impossible to forget or lose, difficult to reproduce, usable with or without the knowledge/consent of the individual and difficult to change or hide. However, current approaches are still focused upon point-of-entry authentication (e.g. PIN/passwords, fingerprint), which has a number of weaknesses. In the case that a user chooses not to use authentication in the first place or once the identity of the user has been verified at login, the mobile device is typically accessible to the user until they specifically exit the system. This can be lead to a high risk environment in which an imposter targets a post authenticated session.

To increase the level of authentication beyond the standard point-of-entry technique, Clarke and Furnell [2] proposed using a combination of secret based knowledge and behavioural biometric techniques to provide transparent, non-intrusive continuous authentication. To this end, research suggests that no single biometric approach is ideally suited to all scenarios and several studies show that multi-modal biometric approaches are superior to one single biometric approach [3-6]

The popularity of the Short Messaging Service (SMS) is one of the most widely recognised and embraced functionalities of mobile communications with over 6.1 trillion messages sent in 2010; close to 200,000 messages sent every second [7]. This provides a unique opportunity to authenticate and discriminate between users based on their individual linguistic morphology.

This paper investigates three individual behavioural biometric techniques: behavioural profiling, keystroke dynamics and linguistic profiling. The performance of each of the aforementioned techniques is discussed together with the development of a multi-modal behavioural biometric approach in which the three individual techniques are combined.

Section 2 provides an overview of biometric authentication. Section 3 describes the methodology and Section 4shows the results. Section 5discusses the implications of the results. Finally, section 6 presents the conclusions and recommendations for future work.

2 An Overview of Biometric Authentication

The International Biometrics Group (IBG) defines biometrics simply as "the automated use of physiological or behavioural characteristics to determine or verify identity" [8]. Physiological biometrics perform authentication based on bodily characteristics such as their fingerprint or their face. By contrast, behavioural biometrics perform authentication based on the way people do things, such as their typing rhythm, their voice or their signature. Physical features are likely to stay more constant over time and under different conditions, and tend to be more unified within a large population [9]. Physiological biometrics therefore tends to be used for identification-based system because they are more trustable approaches. However, some behavioural biometrics have very good accuracy for verification but the identification accuracy of most behavioural biometrics is considerably lower as the number of users in the database becomes larger [10]. This is because users act differently depending on mood, illness, stress, previous events, environment, to name a few. For this reason, behavioural biometrics tends to be only used for authentication-based systems.

Behavioural biometrics provides a number of advantages; they can be collected without the knowledge of the user (non-intrusive) and continuously. Collection of behavioural data often does not require any special hardware and is therefore more cost effective. Based upon a typical mobile device, considering biometric approaches that do not require additional hardware to enable collection, the following biometrics could be utilised; facial recognition, voice verification, keystroke dynamics, behavioural profiling, handwriting recognition and linguistic profiling. Of those that are of interest in this paper: keystroke feature information can be captured through the keyboard interface when users type text messages or mobile phone numbers; linguistic profiling can analyse inputted text messages during SMS compilation; and behaviour profiling can capture users' behaviour continuously during their interaction with the mobile phone. It is hypothesised that each of the three behavioural biometric techniques described can be used to authenticate users. However, more interestingly, it is hypothesised that these three techniques combined together offer the opportunity to improve the underlying performance significantly.

2.1 Behaviour Profiling

Based on mobile devices, behaviour profiling aims to identify patterns of usage based upon characteristics of a user's behaviour. Research in mobile behavioural-based can be divided into two categories: network and host based mechanisms. The former will focus upon user calling and migration behaviour over the service provider network based upon the hypothesis that people have a predictable travelling pattern [11, 12]. A host-based mechanism is founded upon the hypothesis that mobile users utilise their applications differently in different time periods and at different locations. This approach would for example monitor user's calling features (e.g. the day of calling, start time of call, duration of call, dialled telephone number and the location), device usage and Bluetooth scanning [13,14], thereby providing a richer set of potential features than network-based approaches.

2.2 Keystroke Dynamics

Keystroke dynamics is a behavioural biometric which is based on each person's individual typing style on a keyboard. This behavioural biometric is not expected to be unique to each person but it offers sufficient discrimination information to permit identity authentication [15]. Considerable research has been undertaken on the keystroke dynamics and two main characteristics were identified: inter-key latency and hold time. The inter-key time is the duration or interval between two successive keys. Hold-time represents the duration between the press down and releasing of a single key. Several studies have concluded that keystroke dynamics provided valuable discriminative information [16, 17]. Since no additional hardware is required, this has been a favoured technique, with much research on the subject since the 1980's [18].

2.3 Linguistic Profiling

Linguistic profiling is a behavioural biometric that attempts to identify and discriminate between users based on linguistic morphology [19]. Linguistic profiling was used for determination of the language variety or genre of a text, or a classification for

document routing or information retrieval. Linguistic features such as specific aspects of the text often based upon frequency counts of functions words in linguistics and of content words in language engineering are used as a text profile, which can then be compared to average profiles for groups of texts. Considerable research has been undertaken on this technique and many types of linguistic features can be profiled such as lexical patterns, syntax, semantics, information content or item distribution throughout a string of text. Many researchers concluded that structural and stylometric features are valuable tools for author identification and verification [19-21].

The weakness of individual biometric approaches is that no single biometric is ideally suited to all scenarios. For example, linguistic profiling is only practical in scenarios with sufficient word messages. One of the key aims of transparent authentication is to provide a system that enables a variety of biometric techniques to be utilised in order to solve the problem. A multi-modal biometric approach offers the advantage of relaxing the assumption of universality, collectability, acceptability and integrity [22]. Multi-modal biometric approaches require a combination of biometric data. The combination or fusion method can occur effectively at any point within the biometric system: feature-level, matching-level or decision-level [23]. The feature-level method is achieved by combining the variety of feature vectors derived from different biometric techniques. Matching-level fusion takes the output of resulting matching classifications and combines the results (raw score) prior to presenting them to the decision process. At the end of the biometric system, decision-level fusion can occur when each individual biometric system has provided an independent decision. The decision results in a Boolean value which lacks the richness of information for fusion. Amongst the literature, match-level fusion has been shown to be the best performing of the fusion approaches [22, 23].

3 Experiment Procedure

In this paper, three behavioural biometric techniques were investigated: behavioural profiling, keystroke dynamics and linguistic profiling. After studying the performance of each single biometric, the final experiment built upon these findings through the fusion of the three individual techniques. For single modal biometric technique, the general biometric authentication system is illustrated in Fig.1.

Fig. 1. A generic biometric system

3.1 Behaviour Profiling

The experiment based on behavioural profiling described in [13] has been used. For this study, a total of 30 participant's text messaging activities were recorded from a database provided by the MIT Reality Mining project [24]. As not all participants started or finished the experiment at the same time, each user has a varied number of

logs. This dataset contains 1470 logs and 274 unique texting numbers. The maximum number of logs for a user is 149 and the minimum number is 8 logs. For each text log, the following features: receiver's telephone number and location of texting were extracted to create user behavioural profiling. In the analytical process, neural network (Feed-Forward Multilayer Perception Neural Network) was used in the classification.

3.2 Keystroke Dynamics

The dataset of this experiment was provided by [16]. A total 30 participants were obtained with a total of 900 text messages. In this experiment, two main traditional characteristic features were utilised. To create the hold time dataset, the five letters ('e', 't', 'a' , 'o' and 'n') were used in the classification. For the inter-key time dataset, the latency between five pair of letters: 't' – 'g', 'e' – 'p' , 'e' – 'm' , 'h' – 'd' and 'a' – 'm' were calculated. The database contains 3510 hold-time data, 1080 inter-key time data and outliers were removed (a standard procedure for keystroke analysis studies). Analyses were undertaken using Feed Forward Multilayer Perception Neural Network (FF-MLP) as it had demonstrated better performance in previous studies over other techniques [17].

3.3 Linguistic Profiling

In this experiment, the SMS dataset provided by previous research [25] was utilised. A total of 30 participants were required to send at least 15 messages to each other using a non-predictive text input method. The frequency distribution of abbreviations emotional words were used to create user profiles, including every possible type of feature. For each message, a total of 64 discriminating characteristics were extracted for example, average word length (number of characters), total number of sentences, total number of symbols etc. To create a user profile, t-test ranking measure were apply to rank input features according to its discriminative capability. According to the ranking list, features with p value less than 0.05 ($p < 0.05$) were selected for input vectors to reduce the unnecessary features in classification. Therefore, the number of linguistic features required for discrimination significantly differs between users. To analyse individual user's performance, a number of analyses were undertaken, using the Radial Basis function (RBF) neural network algorithm. Different network configurations were tested, looking for the optimum performance.

For each individual biometric technique, the dataset was divided into two groups: 171 data samples were used for the testing set and the rest were used for training. The pattern classification test was performed with one user acting as the valid user, while all others are acting as impostors. The Equal Error Rate (EER) was calculated to evaluate the system. The EER is the value where False Acceptance Rate (FAR) is crosses the False Rejection Rate (FRR), and is typically used as a comparative measure within the biometric industry [26].

3.4 Fusions

In light of the foregoing exploration, the multi-modal biometric study was conducted using a novel combination of behaviour profiling, keystroke dynamics and linguistic profiling. Of all the fusion approaches, matching-level fusion is the most widely used.

However, invariably the use of different classifiers results in different outputs being produced. The range of output result values might vary. In this study, to solve this problem, score normalisation was applied. The equation provides a mechanism to ensure all outputs are bounded between 0 and 1 is shown below:

$$Score\ normalization(X) = \frac{(x_i - \text{Min}\ (X))}{(Max(X) - Min(X))} \quad (1)$$

Where: x_i = the raw score of input i
X = the set of raw score of individual biometric system
Max(X) = the maximum value of raw score vector
Min(X) = the minimum value of raw score vector

After applying score normalization into the raw score results, two fusion approaches were utilised: simple sum and weight average. To evaluate the experiment by simple sum technique, the raw scores of each individual biometric system were simply added and rescaled into [0, 1] as below:

$$Simple\ Sum = normalization\left(\sum_{i=1}^{N}\sum_{j=1}^{M} X_{ij}\right) \quad (2)$$

Where: X_{ij} = the raw score of input i from biometric system j
N = the total number of multi-modal biometric input score
M = the total number of biometric system

For the average weight technique, Weights are assigned to the individual matchers based on their EER and the weights are inversely proportional to the corresponding errors; the weights for less EER are higher than those of high EER.

$$Weight\ average = \frac{\sum_{i=1}^{N}(1 - EER_i)}{\sum_{i=1}^{N} EER_i} \quad (3)$$

Where: i = the number of biometric system
N = the total number of biometric system

4 Results

4.1 Behaviour Profiling

The results of using behaviour profiling to classify user is shown in Table1. The results illustrate that user text messaging application has significant potential to discriminate some users with the overall performance EER 20%. The best case individual user was achieving an EER 1%. Moreover, more than half of participants achieved EER less than 20%. However, the result of worst case individual performance showed fairly high EER 49%. This may be caused by the number of samples assigned to the training of the classification was too small (and a limitation of dataset). Interestingly, only two features: receiver's telephone number and location of texting can achieve the good performance. This is caused by these features having a good level of unique information.

Table 1. Best and worst case results for behavioural profiling

Classifier	EER Worst Case	EER Best Case	EER Average
SMS texting Profile	49%	1%	20%

4.2 Keystroke Dynamics

The main three biometric measurements were investigated: the hold-time, inter-key time and the combination of the hold time and the inter-key time using different network configurations. Table 2 shows the EER of all biometric measurement.

Table 2. Best and worst case results of individual and combination of keystroke characteristics

Classifier	EER Worst Case	EER Best Case	EER Average
Inter-Key Time	46%	7%	31%
Hold – Time	49%	5%	20%
Combination	50%	8%	28%

As illustrated in Table 2, considering the two traditionally keystroke characteristics, the results show that the hold-time gave the lowest average EER 20% with the best individual result EER 5%. These findings illustrate that using hold-time as the key to identify users is the most effective measurement. In contrary to the hold-time investigation, the inter-key characteristic provide fairly high EER 31%, there was the best case of user achieving an EER 7%, showing the ability to classify some users. Therefore, these two main traditional keystroke characteristics provide the valuable discriminative information to classify users. This study experience good performance that has been found in previous studies [16, 17]. In order to further assess the performance of keystroke dynamics, the combination of the hold time and the inter-key time was utilised. The results show that using combination can improve the overall EER of inter-key time by 3%. This is because increasing the component of features can increase the uniqueness of users.

4.3 Linguistic Profiling

The findings from this experiment are illustrated in Table 3. Using linguistic profiling to discriminate users showed positive results. The best individual result achieved the lowest EER 0.00 %. The overall EER also showed promising result with EER 22%. The positive results clearly illustrate that linguistic characteristics could be used successfully to discriminate some users. However, some users generated a fairly high EER. This may caused by selection of keywords and effective features process can result in classification performance.

Table 3. Best and worst case results for linguistic profiling

Classifier	EER Worst Case	EER Best Case	EER Average
Linguistic Profile	49%	0%	22 %

4.4 Fusion

To enhance the overall performance of multi-modal biometric, matching-level fusion of the aforementioned behavioural biometric techniques was investigated. Behavioural profiling results, hold-time results from keystroke dynamics and linguistic profiling were combined. In this experiment, two different fusion methods were utilised: simple sum and weight average. The results show below in Table 4.

Table 4. Best and worst case results of fusion experiments

Classifier	EER Worst Case	EER Best Case	EER Average
Fusion by sum	40%	0%	10%
Fusion by weight average	37%	0%	8%

As shown in Table 4, the results showed that both fusion methods can reduce the overall error rate thus increasing the overall performance. Fusion by weighted average produced better overall results with an EER of 8%, which improves upon the overall performance when compared against a single biometric 14% (based on the worst EER). Fusion method by sum is also efficient because the overall EER is 10%. In both studies of fusion experiments, the performance was improved for every participant. Therefore, using fusion method can improve the performance with low EER for every participant. Additionally, 90% of participants achieved EER less than 20%.

5 Discussion

The results have shown that behavioural biometric techniques based on user texting activities (behavioural profiling), user typing message rhythm (keystroke dynamics) and word messages (linguistic profiling) has significant potential to authenticate users. However, there are some users that have fairly high error rate for each technique. To improve the performance of classification, multi-modal biometric were investigated. In the fusion experiment, two fusion methods were applied: simple sum and weight average. The results demonstrate the utility of using multimodal biometric systems for achieving better matching performance than single modal system. The user achieved the optimum performance by utilising different fusion methods. This also indicates that the method chosen for fusion has a significant impact on the resulting performance. An additional advantage of fusion at this level is that a common fusion method can be utilised to create the reliable system and existing biometric systems do not need to be modified.

In biometric systems, implementers are forced to make a trade-off between usability and security. However it might not all techniques are available to fusion. For example, some biometric technique might have insufficient biometric data to classify. Therefore, a dynamic system needs to be developed. The framework requirements to drive the selection of tolerable error rates and in both single modal and multimodal biometric systems.

6 Conclusions

Behavioural biometric authentication tends to be used for authentication-based systems. This is because users act differently depending on mood, illness, stress, previous events, environment etc. The potential advantages of behavioural biometrics are that they can be utilised transparent and continuous authentication system. Additionally, the collection of behavioural data often does not require any special hardware and is so very cost effective. However, individual biometric techniques are not suited to all users. One way to increase the reliability of transparent and continuous authentication system is create a multi-modal behavioural biometric authentication system.

This research investigated three behavioural biometric techniques, behaviour profiling keystroke dynamics and linguistic profiling based on texting SMS activities and messages, looking to apply these techniques as a multi-modal biometric authentication method for mobile devices. The results showed that individual biometric technique can be used to discriminate users with low error rates. Moreover, the overall EER of multi-modal biometric also showed clearly can be successfully used to authenticate user.

The next step in this research is to further implement dynamic authentication system. The proposed framework also should be flexible and scalable in that it can adopt other biometric techniques. Moreover, the system can integrate new techniques or new biometric techniques without having to change the overall system design.

References

1. Edison, http://www.mformation.com/mformation-news/press-releases/mformation-sponsored-survey-reveals-mobile-users-worried-about-loss-and-mobile-fraud
2. Clarke, N., Furnell, S.M.: Advanced user authentication for mobile devices. Computer and Security 26, 109–119 (2007)
3. Brunelli, R., Falavigna, D.: Personal Identification using Multiple Cues. IEEE Transaction on Pattern Analysis and Machine Intelligence, 955–966 (1995)
4. Kittler, J., Matas, J., Jonsson, K., Ramos Sanchez, M.U.: Combining Evidence in Personal Identity Verification Systems. Pattern Recognition Letters 18, 845–852 (1997)
5. Poh, N., Korczak, J.J.: Hybrid Biometric Person Authentication Using Face and Voice Features. In: Bigun, J., Smeraldi, F. (eds.) AVBPA 2001. LNCS, vol. 2091, pp. 348–353. Springer, Heidelberg (2001)
6. Ross, A., Jain, A.K., Qian, J.-Z.: Information Fusion in Biometrics. In: Bigun, J., Smeraldi, F. (eds.) AVBPA 2001. LNCS, vol. 2091, pp. 354–359. Springer, Heidelberg (2001)

7. International Telecommunication Union, http://www.itu.int/ITUD/ict/material/FactsFigures2010.pdf
8. International Biometric Group, http://www.biometricgroup.com/reports/public/reports/best_biometric.html
9. Woodward, J.D., Orlans, N., Higgins, P.: Identity Assurance in the Information Age. McGraw-Hill/Osborne, Berkeley, California (2003)
10. Yamploskiy, R., Govindaraju, V.: Chapter 1 Taxonomy of Behavioural Biometrics. Behavioural Biometrics for Human Identification: Intelligent Applications (2010)
11. Gosset, P.: ASPeCT: Fraud Detection Concepts: Final Report (1998)
12. Hall, J., Barbeau, M., Kranakis, E.: Anomaly-based intrusion detection using mobility profiles of public transportation users. In: The Proceeding of IEEE International Conference on Wireless And Mobile Computing, Networking And Communications, vol. 2, pp. 17–24 (2005)
13. Li, F., Clarke, N., Papadaki, M., Dowland, P.: Behaviour profiling on mobile devices. In: International Conference on Emerging Security Technologies, UK, pp. 77–82 (2010)
14. Li, F., Clarke, N., Papadaki, M., Dowland, P.: Behaviour profiling for Transparent Authentication for Mobile Devices. In: 10th European Conference on Information Warfare and Security, Estonia, pp. 307–314 (2011)
15. Obaidat, M.S., Sadom, B.: Verification of Computer Users Using Keystroke Dynamics. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 27, 261–269 (1997)
16. Clarke, N., Furnell, S.: Authenticating Mobile Phone Users Using Keystroke Analysis. Information Security, 1–4 (2006)
17. Karatzouni, S., Clarke, N.: Keystroke Analysis for thumb-based Keyboards on Mobile Devices. In: 22nd IFIP International Information Security Conference, pp. 253–263. Springer, Heidelberg (2007)
18. Gaines, R., Lisowski, W., Press, S., Shapiro, N.: Authentication by keystroke timing: some preliminary results. Rand Report R-2560-NSF, Rand Corporation California (1980)
19. Halteren, H.: Linguistic Profiling for Author Recognition and Verification. In: 42nd Annual Meeting on Association for Computational Linguistics, NJ, pp. 199–206 (2004)
20. Argamon, S., Saric, M., Stein, S.: Style Mining of Electronic Messages for Multiple Authorship Discrimination: First Results. In: 9th ACM SIGDD International Conference on Knowledge Discovery and Data Mining, Washington (2003)
21. Goodman, R., Hahn, M., Marella, M., Ojar, C., Westcott, S.: The use of stylometry for email author identification: a feasibility study. In: Student/Faculty Research Day, CSIS, Pace University (2007)
22. Poh, N., Bengio, S., Korczak, J.: A multi-sample multi-source model for biometric authentication. In: IEEE International Workshop on Neural Networks for Signal Processing, pp.375-384 (2002)
23. Clarke, N.: Transparent User Authentication, Springer, p. 229 (2011)
24. Eagle, N., Pentland, A., Lazer, D.: Inferring Social Network Structure using Mobile Phone Data. In: International Conference on Security & Management, pp.207-212, Las Vegas (2006)
25. Saevanee, H., Clarke, N., Furnell, S.: SMS Linguistic Profiling Authentication on Mobile devices. In:5th International Conference on Network and System Security, pp. 224-229 (2011)
26. Ashbourne, J.: Biometric, Advanced identity verification. The complete guide. Springer (2000)