CONTINUOUS USER AUTHENTICATION USING MULTI-MODAL BIOMETRICS

by

HATAICHANOK SAEVANEE

A thesis submitted to the Plymouth University in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Computing and Mathematics

May 2014

Copyright Statement

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author's prior consent.

Abstract

Continuous User Authentication Using Multimodal Biometrics

Hataichanok Saevanee

It is commonly acknowledged that mobile devices now form an integral part of an individual's everyday life. The modern mobile handheld devices are capable to provide a wide range of services and applications over multiple networks. With the increasing capability and accessibility, they introduce additional demands in term of security.

This thesis explores the need for authentication on mobile devices and proposes a novel mechanism to improve the current techniques. The research begins with an intensive review of mobile technologies and the current security challenges that mobile devices experience to illustrate the imperative of authentication on mobile devices. The research then highlights the existing authentication mechanism and a wide range of weakness. To this end, biometric approaches are identified as an appropriate solution an opportunity for security to be maintained beyond point-of-entry. Indeed, by utilising behaviour biometric techniques, the authentication mechanism can be performed in a continuous and transparent fashion.

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 linguistic profiling; keystroke dynamics and behaviour profiling can be used to discriminate users with overall Equal Error Rates (EER) 12.8%, 20.8% and 9.2% respectively. By using a combination of biometrics, the results showed clearly that the classification performance is better than using single biometric technique achieving EER 3.3%. Based on these findings, a novel architecture of multi-modal biometric authentication on mobile devices is proposed. The framework is able to provide a robust, continuous and transparent authentication in standalone and server-client modes regardless of mobile hardware configuration. The framework is able to continuously maintain the security status of the devices. With a high level of security status, users are permitted to access sensitive services and data. On the other hand, with the low level of security, users are required to re-authenticate before accessing sensitive service or data.

Contents

List of Figures

List of Tables

Acknowledgement

First and foremost, I am deeply indebted to my Director of Studies, Associate Professor Nathan L Clarke for being a wonderful supervisor, for his generous support and continuous encouragement. Thanks go to him for his immense help in guiding me toward the successful completion of my PhD studies and spending a significant amount of time proof reading research papers and my thesis. His valuable comments and sincere suggestion have significantly enhanced my work. Without him I would not be able to make this achievement. Also I would like to take this opportunity to express my gratitude toward my co-supervisor Professor Steven Furnell, who provided invaluable advice, guidance and support throughout of my PhD journey.

I would like to thank all my colleges in CSCAN group and Plymouth University and all friends in Thailand, UK and other parts of the World for their support, assistance and entertainment. In particular, I would like to thank Mr Valerio Biscione for his help, support and encouragement.

Last, but not least, I sincerely thank my parents Sanit and Sansanee Saevanee and my family. Without their unreserved love, support and understanding, I could not have gone so far. Also, I wish to thank Sanders family, who cares and supports me during studying in the UK. This thesis is dedicated to them.

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

Relevant seminars and conferences were regularly attended at which work was often presented and were published in the course of this research project.

Word count of main body of thesis: 53440 words

Signed……

Date……

1 Introduction and Overview

This research presents an enhanced authentication approach for mobile devices, provides transparent (non-intrusive) and continuous identify verification of the user. In order to protect against an imposter using the mobile phone 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 (point of entry authentication). This authentication technique requires nothing more than the user interacting with the mobile device in a normal manner.

1.1 Introduction

Whilst the term mobile devices can refer to a variety of devices, such as mobile phones, laptops and games consoles, the mobile phone and its variants form the largest market segment. In this thesis, the phrase 'mobile device' describes two kinds of mobile computing devices: the standard mobile phone and the Smartphone. After almost 40 years of development, the mobile devices 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-mails and online banking), entertainment (e.g. photos and video games) and the sharing of data (via Bluetooth, laptop or computer). Currently, mobile cellular networks are available for 90% of the world population including people living in rural areas. With the rapid evolution and widespread of mobile network, a number of mobile subscriptions worldwide continues to rapidly increase with 6.2 billion over the world (Ericsson, 2012a).

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. The increasing accessibility, capability and storage of sensitive information on mobile devices pose an increasing threat to the owners when a device is misused such as being lost, stolen or infected by malware. According to the National Mobile Phone Crime Unit (2013) website, more than 250,000 – 300,000 mobile devices are stolen and reported to police each year. The motivations for phone theft were investigated and some of the results showed that it is easier to steal a phone than buy a new one and it is ease and value of selling stolen mobile phones. The industry fact also showed that 42% of all incidences in which a phone was stolen, was not in the owner's possession at the time it was stolen. Therefore, malicious users can abuse the mobile devices in many ways by using information and services. For example, review the personal information and eventually steal the owner's identity or buy goods online at the owner's expense (Helium, 2008). If a stolen device belongs to a company, malicious users could sell the information stored on the device to its competitors; also, by using the device as a gateway, malicious users could harvest more information from the company internal network. Therefore, it is mission critical to protect the mobile device from being misused.

Many authentication mechanisms have been developed for mobile devices with the aim of providing a greater level of security for the end user. However, current approaches are still focused upon PIN/passwords or fingerprint authentication. This technique requires users to provide a correct PIN or fingerprint to the fingerprint scanner before accessing to the mobile device regardless of their legitimacy. Although this approach is mostly available on mobile devices, a number of users (40%) failed to utilise this simple point-of-entry mechanism according to a survey conducted by Credant (Credant, 2009). Even though all mobile users utilised this technique on their

devices, misuse could still occur if they did not use it properly in practice, for instance, never changing the PIN, writing it down on a paper or sharing with others (Clarke and Furnell, 2005; McAfee, 2013). In addition, in case of user chose to not using authentication at the first place or once the identity of the user has been verified at login, the mobile system are typically made available to the user until they exit the system. This can be lead to high risk environment which imposter targets a post authenticated session and result in financial, private and sensitive information loss to the user.

The aim of this research is to create an advanced authentication system that increase security required for mobile devices and the level of authentication beyond point of entry techniques to ensure the identity of the user on a continual basis. In order to achieve this continuous authentication, it is important that the technique should provide transparent or non-intrusive authentication to minimise user inconvenience and increase user acceptance. By being able to authenticate a user without their knowledge, the system can be automatically monitored and maintained without the user's explicit interaction until such time that the system detects an imposter accessing the mobile devices.

1.2 Aims & Objectives

This research investigates current state-of-the-art mobile authentication approaches and proposes to define, design and develop enhanced authentication system on mobile devices. The framework is capable of fulfilling the increased security requirements and providing continuous protection to ensure the legitimacy of the current user. The proposed framework encompasses the following objectives:

- To evaluate the performance current authentication techniques and asset the requirement of additional authentication.
- To investigate feasibility of behavioural biometric authentication techniques for deployment on mobile devices.
- To design and evaluate a new multi-modal behavioural biometric technique that provides reliability, transparent and continuous authentication for mobile devices.
- To design a multi-tier authentication architecture that flexible and scalable in that it can be applied to other biometric techniques along with transparent and continuous authentication of users.
- To implement and evaluate an authentication system to study its practical performance.

1.3 Thesis Structure

The thesis addresses the aforementioned objectives in order and is consisted of the following chapters.

Chapter 2 begins by reviewing the popularity of mobile devices along with current mobile services and the increasing reliance upon them establishing the importance of the security of the mobile device. By presenting the security risks and current security approaches, the need for a new security mechanism which can provide continuous and transparent protection from misuse for the mobile devices is identified. The chapter concludes by highlighting the need for an enhanced authentication system and suggesting possible solutions.

Having establishing the need for an advanced authentication technique for mobile devices, Chapter 3 presents and discusses the feasibility of utilising biometric verification techniques as a solution to deal with this issue. The chapter starts by presenting a generic biometric system, the performance measurement and requirement factors. The chapter proceeds to describe an overview of existing biometric techniques based upon their physiological and behavioural characteristics. By comparing all applicable biometric techniques for the mobile device, a novel linguistic profiling technique, keystroke dynamic and behaviour profiling were chosen due to their various advantages that can provide continuous and transparent protection security.

In order to understand the current state of the art of selected behaviour biometrics, Chapter 4 presents a comprehensive literature review on the practical use of linguistic profiling technique, keystroke dynamic and behaviour profiling. Based on the knowledge of the literature review, the practical use of multi-biometric technique is discussed.

Chapter 5 presents a feasibility study of using linguistic profiling to authenticate a user. The studies have been examining text-based features that can be extracted from SMS messages to classify the user. By analysing and testing linguistic features using statistical and a pattern classification method based upon artificial intelligence algorithm in preliminary studies, a number of useful features that can provide discriminative information toward success verification have been identified. The chapter finishes with evaluating the linguistic profiling technique based upon SMS text messages by using the dynamic profiling technique.

Chapter 6 builds on the success of chapter 5, a number of experimental studies into the feasibility of linguistic profiling on a mobile device. The chapter compares a number of pattern classification methods based upon artificial intelligence algorithms and the most appropriate classifier has been identified. The chapter also present the feasibility study based upon using keystroke dynamic and behaviour profiling techniques to identify users. Finally, the study based on the use of multi-modal biometric technique is demonstrated.

Chapter 7 presents a novel mechanism for composite authentication. By utilising composite authentication techniques, the ability to correctly verify the identity of a user becomes stronger as the weaknesses of one technique are overcomes by the strengths of others. Furthermore, the use of more than one authentication approaches permits a more transparent means of authentication. By using different techniques in varying handset scenarios it is possible to non-intrusively authenticate the user during a wider range of handset interaction. The proposed framework consisted of linguistic profiling, keystroke dynamics and behaviour profiling authentication techniques. However, the architecture has been designed in a flexible and scalable manner in that it can be applied to other authentication techniques. Moreover, the system can integrate new techniques or new biometric techniques without having to change the system design and also can be applied with any given mobile environment.

Finally, Chapter 8 presents the main conclusions from the research, highlighting the key achievements and limitations. The chapter also discuss on the future research and development.

2 The Impact of Mobile Devices in Society

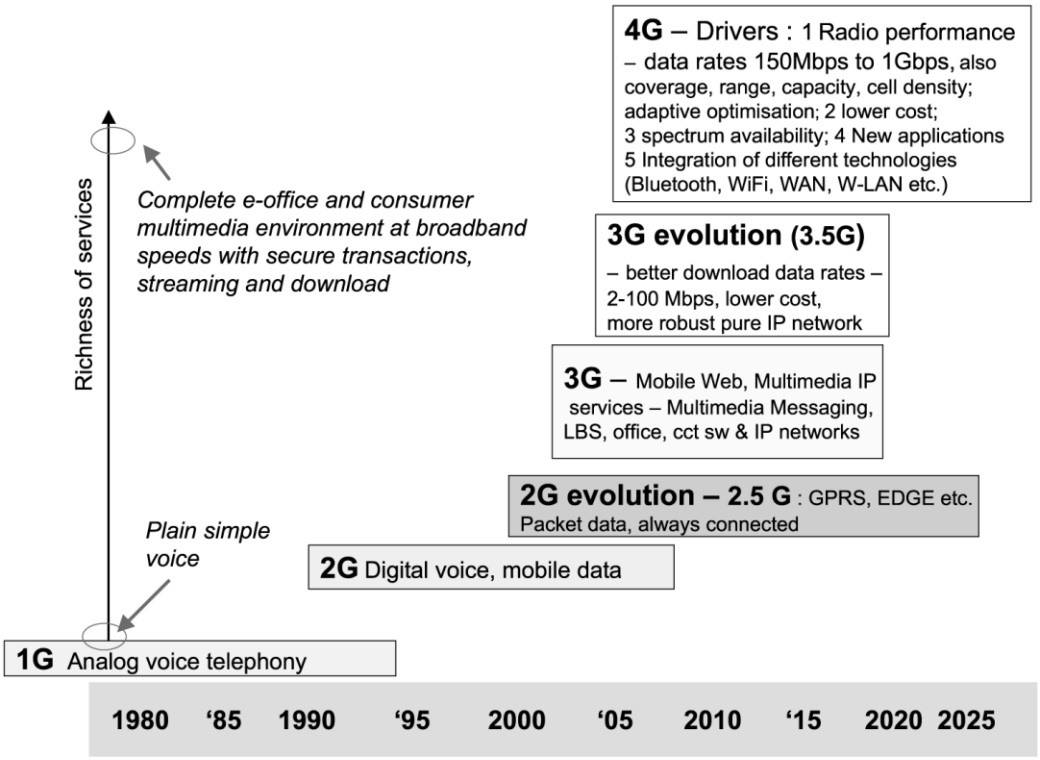
2.1 Introduction

This chapter investigates and discusses the need for user authentication on mobile phones by highlighting the popularity of mobile devices, the increasing reliance upon them, and the security risks. The chapter beings with a discussion on the prevalence of mobile devices and proceeds to evaluate the impact of they have upon society. The security risks of mobile devices are investigated focussing upon the current security challenges that mobile devices experience. Consequently, the need for more protection on mobile handsets is discussed. Current authentication approaches available on mobile devices are explored with in-depth analysis to identify the weaknesses and vulnerabilities of these approaches. Finally, additional techniques were discussed which aim to increase user convenience and stronger authentication.

2.2 The prevalence of mobile devices

The vast majority of countries throughout the world have access to mobile cellular networks with 90% of the world population are covered by 2G networks and 45% of the world population are covered by 3G networks (ITU, 2011). 2G networks (Second-generation wireless telephone technology) provide digital voice services and data transmission, including sending and receiving text and multimedia messages via Short Message Services (SMS) and Multimedia Message Services (MMS) respectively. Furthermore, the 2G networks also offer Internet communication services; users are able to access Internet for sending or receiving email and surfing websites. The early adoption of 2G data services, particularly SMS, provided an insight into the popularity for mobile data services. Several network technologies, including 3G and GSM to

name but two have all sought to increase the bandwidth available to handsets. Figure 2.1 illustrates the development of mobile cellular communication system.

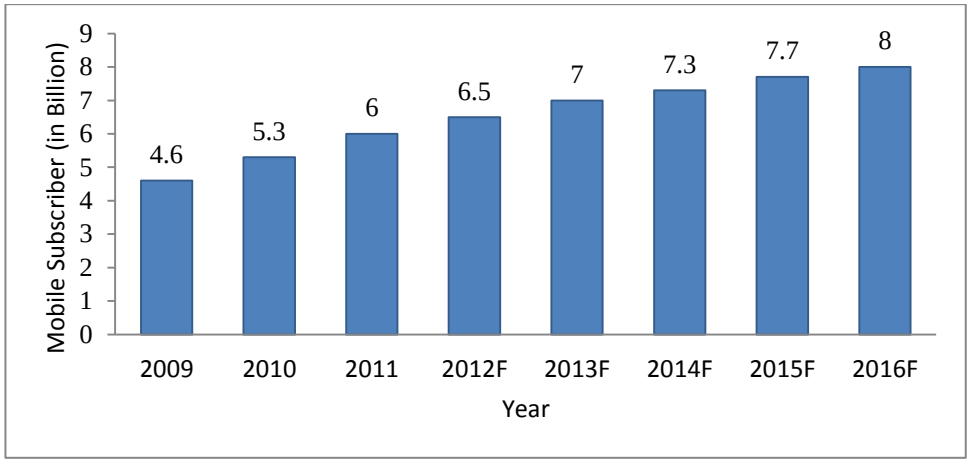


Source: Forge, 2004

Figure 2.1: Evolution of mobile cellular communication system

The 3G networks were developed to provide high-speed data transmission (2 Mbit/s), a great network capacity and more advanced network services. This allows users quick and easy access to all online multimedia and Internet tool. With these benefits, over 159 Countries including the UK have launched 3G networks, which covered almost half (45%) of the world populations. A number of western countries such as Sweden, Norway, Ukraine and United States are already moving to 4G (ITU, 2011). The 4G networks offer even higher data transmission (up to 1 Gbit/s) together with Quality of Services (QoS) for multimedia purpose such as real time audio, high-definition mobile TV and 3D television.

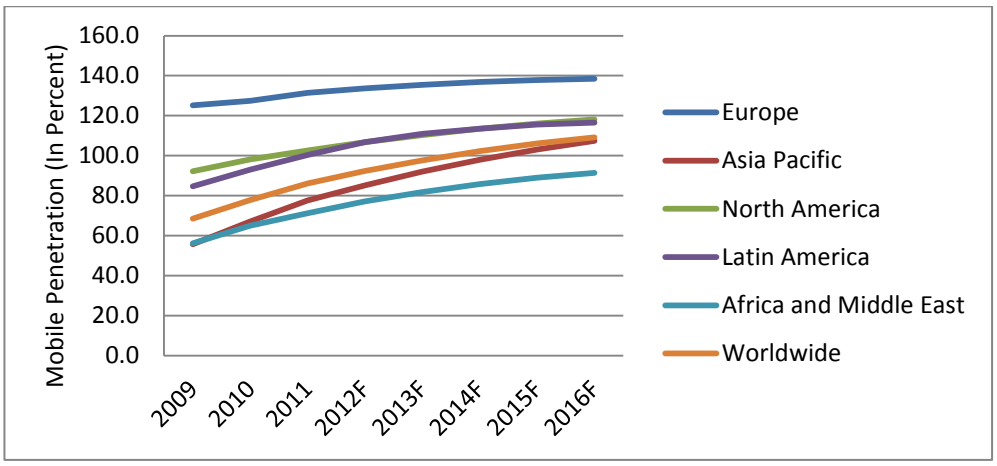
With the rapid evolution and widespread of mobile networks, the number of mobile subscriptions worldwide continues to rapidly increase with 6.2 billion over the world (Ericsson, 2012a). However, it should be noted that the actual number of mobile subscribers is only around 4.2 billion as many subscribers have several subscriptions (e.g. a person who has two mobile phones for personal and business uses). Figure 2.2 highlights the growth in worldwide mobile subscriber during 2009-2016. At the end of 2016, the number of mobile subscriptions worldwide is predicted to reach 8 billion according to report by Portio Research (2012).



F - Forecasted
Source: Portio Research Ltd., 2012

Figure 2.2: Mobile subscriber base-worldwide

Figure 2.3 illustrates a constant and continued growth in mobile subscription. When comparing number of the world populations and mobile subscription worldwide; the global mobile penetration rate shows 91% at the end of 2012(Ericsson, 2012b).



Source: Portio Research Ltd., 2012

Figure 2.3: Mobile penetration by regional between 2009 and 2016

In some regions the number of mobile subscriptions is already more than number of people in those areas as shown in Figure 2.3, Europe is by far the highest mobile penetration rate, followed by North America and Latin America where penetration rates exceed 100%. By the end of 2016, the number of mobile subscriptions will exceed the number of world populations with the global mobile penetration at 110% (Portio Research, 2012). This indicates that every person will have at least one mobile device within a few years.

2.3 The impact of mobile devices in Society

Along with the rapid development of mobile network technology and the increasing popularity of mobile devices, the functionality of these devices has also dramatically improved. Modern mobile devices come equipped with a number of network interfaces, powerful computing capabilities and high capacity data storage. Currently, mobile devices can do much more than make calls or send text messages. As an illustration, the latest features that come with latest mobile devices are shown in Table 2.1.

Applications/Services	IPhone 5 (2012)	IPhone 5s (2013)
Design		
Weight	140 grams	112 grams
Size	115x59x9.4 (mm)	123.8x58.6x7.6 (mm)
Movability	✓	✓
Hardware		
Data Capacity	64 GB (internal)	64 GB (internal)
Processing power	Dual core A5	Dual core 1.3GHz
Microphone (built-in)	✓	✓
Connectivity		
Internet	3G, GPRS EDGE, GSM	4G, GPRS EDGE, GSM
Wi-Fi	✓	✓
Bluetooth	✓	✓
Navigator		
GPS	✓	✓
Communication Services		
Voice calls	✓	✓
Data transmissions	✓	✓
Personal Manager		
Contact book	✓	✓
Calendar	✓	✓
Reminder	✓	✓
To-do list	✓	✓
Memo/note	✓	✓
Email	✓	✓
Web browser	✓	✓
Multimedia Service		
Camera (built-in)	✓	✓
Recorder	✓	✓
Music player	✓	✓
Video player	✓	✓
Game player	✓	✓
Photo viewer	✓	✓

Table 2.1: Mobile device specification

Table 2.1 illustrates that mobile devices can provide a wide variety of services and application similar to that offered by a personal computer. The dual core processor makes mobile devices even more responsive. Data storage supports a wide range type of information and is almost infinite with 64GB built in memory. This massive data

storage allows user stored a large number of information. Modern mobile devices also come equipped with following functionalities:

- **Internet access**

  The majority (85%) of mobile devices are now capable of accessing Internet throughout mobile network (Gartner, 2010). Mobile devices with Internet access allow users to access webpage, send/receive emails and chat online while users are on the move. Because of the mobility, the number of people accessing Internet through their mobile devices has dramatically increased across the world and is forecasted to overtake the number of people accessing Internet through their PCs for the first time in 2014 (Gartner, 2012). This indicates that a mobile device has become a primary Internet access point.

- **Wi-Fi technology**

  Wi-Fi is a short for “wireless fidelity” and is a limited-rang wireless networking. This technology enables mobile device to connect to public and private networks for Internet access with a bandwidth up to 150 Mbps. In general, Wi-Fi offers more reliable and faster data transmission speeds than many cellular networks. Users can exchange or share data via FTP or similar and access Internet-based services such as check email and access website by using a Wi-Fi connection. Currently, many network operators provide the Wi-Fi services around the world. BT alone offers over 4 million public Wi-Fi hotspots across the UK and 2 million more around the world (BTopenzone, 2011). The wide availability of these hotspots has encouraged 4.9 million UK mobile users connected to wireless hotspots (ONS, 2011).

- **Bluetooth technology**

  Bluetooth can be defined as a wireless form of communication enables mobile devices to connect to other compatible devices at speed up to 3Mbps with a maximum range of approximately 30-60 feet. This technology is a global initial setup function by many manufacturers such as Nokia, Ericsson, Toshiba and IBM. Within Bluetooth network, known as Personal Area Network (PAN), a user can directly exchange data with each other such as text messages, multimedia message and data files. At the end of 2010, a total of 5 billion Bluetooth enabled mobile devices, which mainly mobile phones were shipped and predicted to reach over 20 billion by 2017 (ABI Research, 2012).

- **Near Field Communication technology**

  NFC is a wireless radio standard which allows two devices to communicate between each other by touching or bringing them into close range (i.e. no more few inches). This technique can be used in many areas such as access control, data exchange and contactless payment. Frost & Sullivan (2011) suggests that NFC-enabled mobile phones will reach 863 million units in 2015. Furthermore, they also noted that NFC will clearly be the most-used solution for mobile payment. In addition, mobile payment via NFC is expected to reach €111.19 billion in 2015.

- **Productivity tool**

  By default, a number of common applications are preinstalled on mobile devices by their manufacturers such as clock, calendar, contact book, to-do list and notepad. However, users can download and install a plethora of applications on the devices according to individual preference, including office

applications and PDF readers. Therefore, users can use their mobile devices to complete various tasks such as storing and modifying documents, preparing and delivering presentations.

- **Global Position System**

  Global Position System (GPS) is a system that provides positioning, navigation and timing information services. With the embedded GPS, users can use their mobile devices to get real-time position of where they are, guide directions, information on traffic conditions with suggesting alternative directions and information on nearby interesting facilities such as restaurants, theatres and hotels. GPS system has become a standard features on mobile phones and were considered to be a must-have feature. In 2009, a total 150 million GPS-enabled mobile handsets were shipped and is expected to grow to 770 million units in 2014 (Berg Insight, 2010).

- **Entertainment**

  Most mobile devices, even inexpensive ones have a built-in digital camera. Users can use these devices to capture photos or video regardless of time or location. Mobile Internet access enables users to post directly to Facebook, Twitter and YouTube, send via email or share multimedia files with other devices quickly and efficiently. Additionally, mobile devices also can be used for video conferencing and play song, movie and game.

- **Mobile Applications**

  Mobile devices have the ability to run third-party applications that users download or purchase from application stores. These mobile applications can

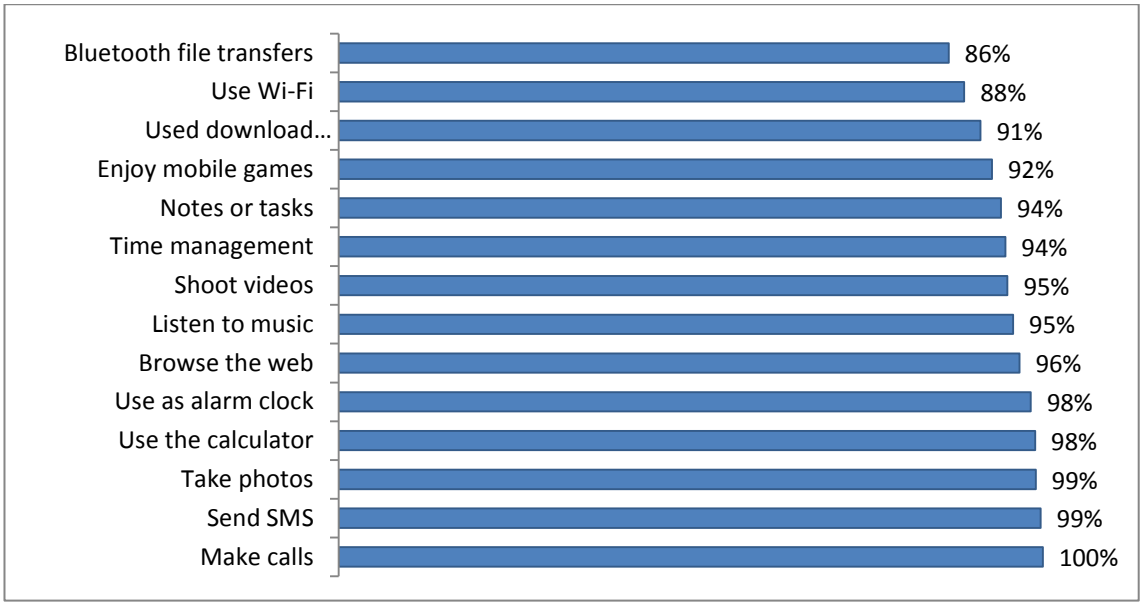
provide amazing functionality for such as communications, games, multimedia, productivity, travel and utility purposes. The emergence and growth of mobile application software stores have created a huge extended functionality of mobile devices. Currently, more than one million applications are available for people to download from across different mobile platforms. In addition, almost 15,000 new mobile applications become available to download every month (Distimo, 2010). The mobile application global market is expected to grow triple from $10 billion in 2009 to $32 billion in 2015 (Juniper Research, 2010).

- **Mobile payments**

  Instead of paying by cash or credit card, mobile users can use mobile phone to pay for a wide range of services, digitals and goods. By using Wi-Fi or internet access, users can use internet banking to transfer money and pay utility bills. Also, by using the mobile NFC service, people can pay for goods at supermarket checkouts or vending machines. For instance, Orange and Barclaycard launched a mobile payment service which enables the users to pay for goods up to £15 (Aol tech, 2011). According to a research report of Berg Insight, there were 133 million mobile money users who made a total of $25 billion of transactions in 2010. Furthermore, they also predicted that with an annual increase rate of 40%, there will be 709 million mobile money users in 2015 with a total of $215 billion transactions (Berg Insight, 2011).

As illustrated above, modern mobile devices are multifunctional, providing the ability to perform a wide range of functions beyond voice communication. This functionality greatly enhances people effectiveness and efficiency for both personal and business purposes. In terms of general usage, people now use their mobile devices to facilitate

the most of fundamental personal daily activities. Figure 2.4 shows the most essential features used on mobile devices by a large number of users.



Source: TNS mobile life, 2012

Figure 2.4: Consumer mobile activities 2012

According to Figure 2.4, a majority (99%) of mobile users use their device for sending text messaging in addition to using mobile devices as a phone. The average length of sending a text message is only 30-seconds however; user spends an average of 10 minutes using their mobile devices for text messaging (O2, 2012). This has created a huge number of messages sent worldwide with an average 1,670 million messages per day (ITU, 2010). Moreover, mobile devices are being used as a multipurpose handheld device replacing traditional portable devices such as cameras, calculators, watches and multimedia players. Furthermore, the majority of users also use their Internet connected mobile devices to browse the web and download applications. In total, 29 billion applications were downloaded in 2011 and mobile applications were projected to generate more than $15 billion in apps store revenue (Gartner, 2011). Additionally, the majority of users also enable Wi-Fi and Bluetooth connections on their devices for

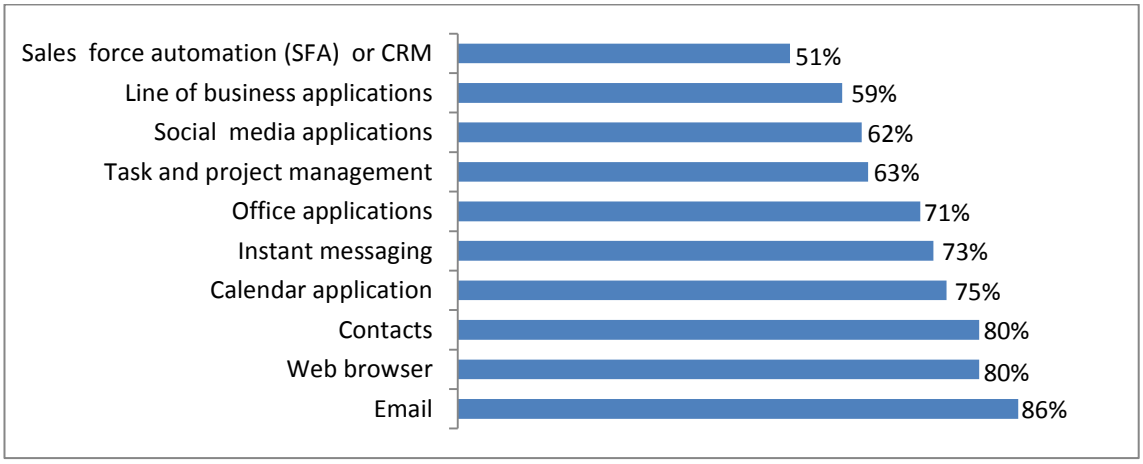
accessing Internet based services and sharing files. Mobile devices have become a part of individual day live. People are now reliant on their mobile devices because they currently use them for multiple purposes as such the amounts of confidential personal sensitive information will be stored on them.

Mobile device have not just changed user's personal lives, they are changing their work lives too. Many organizations are now allowing employees to "bring their own mobile devices" (BYOD) into work. Currently, more than one in five organizations allow use of any smartphone platform in the workplace and almost half of businesses (49%) allow employees to purchase their own mobile devices (MacAfee, 2011). A majority of businesses believe that BYOD can help improve efficiency, increase workplace effectiveness and reduce time to accomplish job tasks (Symantec, 2012) as shows in Figure 2.5 Cisco IBSG estimates that the annual benefits form BYOD range from $300 to $1300 depend on the employee's job role and work requirements (Cisco IBSG, 2012).

Source: Symantec, 2012

Figure 2.5: BYOD benefits

Many organizations allow employees to access the Internet and download mobile applications freely using their mobile devices, and more than a third of businesses allow mobile devices users to connect to corporate network with their mobile devices (McAfee, 2011). Consequently, mobile device have become increasingly used in workplace and continue to replace desktop computer. According to survey report showed that about three-quarters (71%) of business are developing business and customer applications and making plan to create corporate application stores for employees to download software and then store on their mobile devices (Symantec, 2012). Employees are now adapting to BYOD because they want more control of their work experience, thus they can improve productivity and job satisfaction. On average, employees use mobile devices for a work purposes between two and 4.5 hour a day (McAfee, 2011). This indicated that both employees and businesses are heavily reliant on the use of mobile devices. Employees use their mobile devices for a wide variety of tasks ranking from email to CRM functions as demonstrates in Figure 2.6.



Source: McAfee, 2011

Figure 2.6: Employee mobile activities

Since mobile devices have been used for work and business purposes, a significant amount of sensitive corporate information were stored on them. According to various studies (Kaspersky Lab, 2011; McAfee, 2011; Dimensional Research, 2012) showed that

corporate email, business contacts, customer data, network login credentials and corporate information made available through business applications were stored on mobile devices.

As shown in this section, mobile devices are now capable of providing a wide range of services and applications. Currently, people use these devices to complete various tasks and store private information or critical business data in their daily life. Therefore, it is mission critical to ensure the legitimacy of a mobile user throughout every single session of usage. Otherwise misuse would occur on both the services and information provided by the mobile devices.

2.4 Mobile devices security threats and controls

Although the modern mobile devices provide a number of advantages from wide range of communication technologies and store variety of information they can also propose a number of security concerns.

- **Mobile malware and viruses**

  As the modern mobile device has become much like a normal computer in terms of information processing, data storage and network abilities, it is possible to face similar security issues which the traditional computers experience such as malware. Malicious software (malware) is designed to harm a computing system, gather information or launch other types of attacks. The most well-known malware are viruses, worms, Trojans and spyware. Malware can cause a loss of personal or confidential data, send premium Short Message Service (SMS) text messages, make phone calls in the background or make the device unusable. Mobile malware can be spread in many ways, such as through

a wireless connection, via a multimedia message to another device or a company's intranet. Some malware also embedded in a mobile application. Once the application is downloaded and installed, the malware is also installed but without the owner's awareness. For instance in 2010, up to 4.6 million Android users downloaded a suspicious application that secretly collected and transmitted users' information to a website in China (ComputerWeekly, 2010).

- **Loss and theft of the device**

  Mobile devices can be easily lost or stolen due to their small size and portability. According to the metropolitan police website, there are around 10,000 mobile devices lost or stolen in London every month (Metropolitan Police Service, 2011). When mobile devices are misplaced, unauthorized access, calls and services could be occurred. More seriously, information on mobile devices could be at risk of theft or leakage. According to the security study, Symantec showed that almost (96%) of lost smartphones were accessed for both business and personal related applications and information by the attacker (Symantec, 2012). However, it is not only data on mobile devices that is at risk but also critical business data stored on corporate networks. Due to many employees using their mobile devices to access corporate networks and data, the confidential data could be accessed. The consequences of data breach are significant. An attacker can use sensitive information for identity theft, selling business technical or financial information to competitors, abusing customer's confidential data and misusing a corporate name or product brands. All of these misuses could result in, for example, damage to a company's reputation;

reduce customer confidence and damage relationships as well as negative financial impact on the organization (Cisco, 2009).

- **Mobile service fraud**

  A service fraud occurs when a user uses services such as voice calling and text messaging provided by the services providers without paying a charge. For example, when a mobile device is lost or stolen, an unauthorised person could access the mobile services until the owner of the device reports the incident to their service provider. In the worst case scenario, a criminal could use more sophisticated attacks such as a Subscriber Identity Module (SIM) card cloning attack. By exploring the coding flaws within a cellular network authentication process, criminals could clone a victim's SIM card and abuse the services at the victim's expense (Rao *et al*, 2002). In this way, it is possible that a mobile owner would not notice that the abuse has occurred unless they check their billing account. According to the Global Fraud Loss Survey 2009 of the Communications Fraud Control Association (CFCA), service fraud is estimated to cost telecommunication service providers $72-80 billion every year around the world (CFCA, 2009). In addition, the number of fraud attacks has increased significantly (by 74%) between 2008 and 2009 (BBC news, 2009).

- **Bluetooth and Wi-Fi**

  Bluetooth and Wi-Fi effectively increase the connectivity of mobile devices within a certain range, but they can be exploited to infect a mobile device with malware or compromise transmitted data. It is possible that a mobile device may accept a Bluetooth connection request from a malicious device. In a "man-in-the-middle" attack, when mobile devices connect, attackers can steal and

download confidential data stored in the phonebook and calendar, they can retrieve pictures and text messages stored on memory without alerting the owner of the target device.

In order to address the highlighted mobile security threats, various security projects have been proposed and developed in many ways such as:

- **Mobile anti-malware or virus solutions**

  In order to prevent infection by viruses, worms and Trojan, several antivirus companies have developed mobile antivirus software for mobile device operating systems, such as Kaspersky Mobile Security®, Trend Mobile Security® and Norton Mobile Security®. Unfortunately, a stand-alone antivirus solution will be difficult to manage, as each OS will likely need its own solution. Alternatively, anti-malware technology is built into a variety of endpoint security and mobile device management products that are designed to protect tablet PCs and smartphones in a centralized implementation. However, it is imperative to ensure that the existing OS is covered.

- **Mobile Encryption**

  Encryption techniques transform information into a format that cannot be read or understood unless a key for the secured data is provided. As mobile devices have the ability to store large amounts of private and sensitive corporate information, anyone obtaining the device could access the information if the information is not encrypted. A number of encryption methods have been proposed in order to protect the information stored on the mobile devices. For instance, Microsoft Corporation offers encryption on Windows based mobile

devices and RIM provides email encryption for the email system of Blackberry. With the increasing BYOD trend, 40% of organisations are planning to deploy data encryption methods on their employee mobile devices (Goode Intelligence, 2011). However, encryption methods require a certain level of education for the average mobile device user before they can implement this technique.

- **Mobile firewall products**

  In order to protect a mobile device from network-related attacks, mobile firewall software can be a solution. The software continuously monitors and controls the network traffic and only allows legitimate service (e.g. web browsing) to go through a mobile device. Firewalls can be implemented on the network or on the mobile device. Nokia Corporation proposed a network based firewall which can be implemented on the mobile service provider's networks for blocking malicious data going into a mobile device (Newscientist, 2007). Unfortunately, the service provider's firewall is unable to protect mobile devices from other networks such as Wi-Fi or Bluetooth. On the other hand, a host based firewall has the ability to monitor all networks that the mobile handset connects with. Several companies offer the services of firewall protection such as ProtecStar Mobile Firewall 1.0 and Trend Micro™ Mobile Security 5.0 (ProtectStar, 2009; Trend Micro, 2009).

Although antivirus software can detect the presence of malware, firewall applications can block malicious network traffic, battery based mobile Intrusion Detection System can sense both malware and network related attacks and encryption can translate information to an unreadable format unless a secret key is provided, these security countermeasures are basically useless if the security software has any software

vulnerability. A hacker can take advantage of the weakness to obtain access to the system. Once the compromise is successful, the hacker is able to modify the firewall access control policy or antivirus to allow for further attacks. If the system is setup for remote access, the hackers needs to only compromise the authentication credential to obtain access to the system. From a physical attack perspective, the only control preventing access to the system is authentication – assuming they have successfully bypassed the physical protection (if present). It should be clear now that authentication appears across the spectrum of technical control (firewall, encryption, antivirus, etc.). It is the first barrier in ensuring the effective and secure operation of the system, applications and security controls. The next section presents the state of the art of authentication on mobile phones.

2.5 Current authentication mechanisms

To protect mobile devices from unauthorized user access, an authentication mechanism is unavoidable utilized. Authentication is the process of confirming a user to ensure that access can only be given to an authorized person. Currently, most mobile devices commonly use the following security features:

- **Personal Identification Number (PIN) Authentication**

  PIN authentication can be deployed to prevent an unauthorized person accessing a mobile device and the Subscriber Identification Module (SIM). Normally, a mobile PIN code contains between four and eight digits. A user is required to enter the correct PIN code before accessing the mobile device and most of the time the user will not be required to re-enter the PIN until the next reboots. However, for additional more layers of authentication, a PIN can be

set to be requested again after a certain period of time dependent upon the user's preference. By using a PIN code to prevent unauthorized persons accessing the SIM card, a PIN code is request at "switch on" or when a SIM card is inserted into a compatible mobile device. Without a correct PIN, the mobile device would not start and the SIM card would not authenticate with the cellular network. In the case of a user entering SIM PIN code incorrectly three times; the SIM card will be blocked; consequently a user is unable to access the mobile network. The user has to ask for Personal Unblocking Key (PUK) to unlock from network operators.

- **Password Authentication**

  Password authentication protects mobile devices from unauthorized user access. A user is required to enter the correct password before accessing the mobile device. Passwords can contain a string of letters, characters and numbers, which can provide a large number of set of passwords in comparison to PIN. The length of a password is dependent on the security policy of the particular application. For example, Blackberry devices with software 4.6 to 6.0 support security passwords between four and thirty-two characters in length with password rules applied such as passwords cannot contain a natural sequence (e.g. 1234). However, it could be difficult to type long password on small keypads.

- **Recognition-based passwords Authentication**

  These passwords are not based on character or digit input. Instead, this method is based on the way a user draws an unlock sequence. For example, in the Android password pattern, a user is required to draw a link between nine (3x3)

dots to create a pattern as their password to access mobile devices as shows in Figure 2.7. The length of a pattern is between four and nine. However, the limitation of this method is that a dot cannot be used more than one time. As a result, this technique provides less number of password patterns than traditional PIN and password.

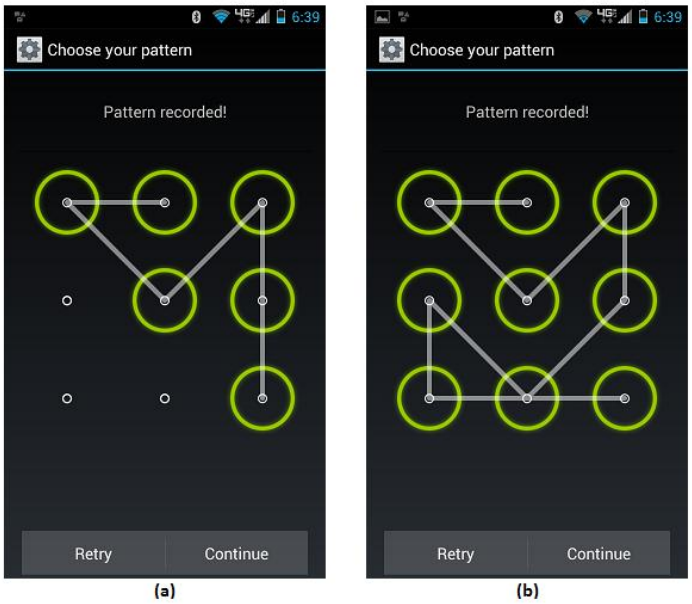


Figure 2.7: An example of Android password pattern

Although these techniques are available on mobile handsets, in practice, may users do not utilize a PIN or password to protect their devices (McAfee, 2013). The lack of use of PIN or password authentication has been stated that because users do not confident with the protection of it provided and considered it is inconvenient (Clarke and Furnell, 2005). In addition, many users who utilize authentication do not use it properly. According to the survey (McAfee, 2013), it is reported that majority of respondents use the default password that is given after initially purchasing the handset. More than one in ten respondents uses the same PIN across multiple devices and accounts, over half of users shared the same passwords with others and 15% save password

information on their phone. These make the PIN/password based authentication technique inadequate as a protection for mobile devices (Clarke and Furnell, 2005; Kurkovsky and Syta, 2010).

One observation regarding PIN/password authentication is that these approaches are based on point-of-entry mechanisms that require the user to actively do something such as enter a PIN number or password in order to be authenticated at the beginning of a session. Once authenticated, the mobile device is able to remain on and no further verification of the user is undertaken until the device is switched off or the session has expired. During the session, all services, application and information on the mobile device is accessible to the user. The user is able to do everything for example play a game, send a text message, make an international phone call, access personal/corporate email and remotely access corporate networks to copy a company's customer database. Another limitation of the point-of-entry authentication techniques is that they require the user to actively do something in order to be authenticated. This can be considered inconvenient as it is intrusive or interrupts the user activities (Rodwell *et al*, 2007).

Beyond secret-knowledge authentication, two other authentication techniques are available, namely token and biometrics. Token based approach is considered impractical in the sense that the user would need to carry them around with the mobile device thus increasing the risk of one of the necessary items being lost or forgotten. Another issue is that if the authentication process requires the token to be placed into the device then it is highly probable that many users would leave it in mobile device. This can be illustrated by the use of a SIM card on current mobile devices. When the users do not want to use the mobile, users could remove it when

the device is not in use. However, removing the SIM card would be inconvenient. By utilising contactless technology (i.e. Bluetooth or RFID) it is possible to develop tokens that can be integrated within the item that users would always expect to have with them such as rings or wristwatches. Although this technique is feasible and also can increase user convenience over the secret-knowledge approach as no interaction is required. However, this approach still requires the user to remember to wear the token.

The last approach to authentication is biometrics. This technique authenticates a user based on unique characteristics of the user such as fingerprint, hand and face. The biometric method does not require users to remember anything just being themselves. With the evolution of mobile devices and network technology, some mobile devices come equipped with fingerprint-reader or face recognition technology which can provide more secure authentication mechanism. Examples of available devices are several models e.g. iPhone5 by Apple with a built-in fingerprint-reader on the home button of mobile handset or Samsung Galaxy Nexus by Samsung with a face recognition technology. However, although fingerprint and face recognition technologies increased the level of security and convenient, these techniques remain point-of-entry-authenticate and intrusive to the user.

As described above it is clear that the point-of-entry authentication approach has been developed to determine permission over access to the device itself and provides no further protection during the usage session. However, in reality, the need for security will vary depending upon what the user is doing and different services and data should require different levels of security. Since each service carries a certain risk of misuse, this must be a factor in deciding the appropriate level of security. If the level of

security is appropriately assigned to each service so that each service or function can independently require a certain level of authentication in order to grant access to the specific service. In this way, more critical operations could be assigned a greater protection and leave less risky operations to a low level of trust. As a result, an advanced authentication approach which is capable of continuously monitoring and authenticating a user based upon the users' legitimacy is desperately needed. This can be best achieved by using non-intrusive or transparent technique; so that users would not be aware that authentication is taking place, avoiding the user having to stop operation in order to re-enter a PIN for instance.

2.6 Conclusions

The current state-of-the-art mobile phones are now as powerful as a personal computer and have become an integral part of an individual's everyday life. People use their devices for a variety of tasks and also utilise sensitive information such as storing personal and business data, accessing corporate emails and networks. As the services and sensitivity of information stored on a mobile device increases, the need for effective security increases. In order to ensure the right user access to sensitive and valuable data, an enhanced authentication system is required.

This chapter has explored the current authentication approaches utilized on mobile devices and found that a knowledge-based approach is not enough to safeguard mobile device and data access through them because PIN/passwords have never been completely protected by the owners; sharing passwords with friends or any other systems are common problems. Token based approaches that still rely on the user to remember the password or PIN, or to remember to pick up the token along with the handset. The third approach to authentication, "something the user is", known

generally as biometrics, does not rely on the user to remember anything, just on being themselves.

On the basis of these findings, it is clear shows that stronger authentication is required for mobile devices. This advance authentication technique should increase the level of security beyond point-of-entry authentication by providing non-intrusive, convenient (to user) and continued security throughout the duration of session. In addition, the majority of mobile device manufacturing companies (69%) believe that device-integrated security is the most effective and efficient way to protect devices, carriers and users at the same time (McAfee, 2009). This means protection should be part of the device, so it is secure without additional user action required. The possible foundations to achieve these are considered in the next chapter.

3 Biometric Authentication

This chapter presents a background of biometric authentication systems, how to evaluate the performance of the system and the properties of biometric characteristics that can be used in biometric systems. An overview of biometric techniques is presented, focusing on a number of approaches that are suitable deployed on mobile devices. Finally, an analysis of the specific biometrics that has the ability to achieve transparent authentication is discussed.

3.1 Introduction

Compared to traditional knowledge-based and token-based security systems, biometrics techniques tend to be more complex to implement. However, they are more difficult to hack, break and circumvent by an attacker to compromise. Furthermore, biometric techniques do offer a user to be authenticated without any reliance upon them to having to remember, by merely utilising a natural consequence of our being. Therefore, the use of biometrics is the best solution to provide more robust and reliable user authentication.

Unfortunately, biometrics also suffers from the problem of usability as they are typically implemented in an intrusive fashion. However, Transparent Authentication Systems (TAS) have been considered for use in mobile devices due to their ability to continuously monitor and authenticate users non-intrusively (Clarke, 2011). Non-intrusive authentication negates the need for explicit user interaction whilst continuously verifying the user throughout the session. This helps to minimise user inconvenience through not explicitly requiring the user to authenticate and also improves security through continuous evaluation of the user's identity. Given these

advantages, this chapter will focus upon discussing biometric techniques that lend themselves to being applied in a transparent fashion.

3.2 An introduction of the biometric system

3.2.1 A generic biometric system

Biometrics as defined by the International Biometrics Group (IBG) is *"The automated use of physiological or behavioural characteristics to determine or verify identity"* (IBG, 2010). Biometric technology could operate in two modes: a verification (authentication) system or an identification system; both of which are defined below:

- **Authentication:** (Does this identity belong to you?)

  In authentication mode, the system will verify a personal identity by comparing a submitted biometric sample with the biometric reference template of a user whose identity is being claimed. In this system, the user needs to claim an identity and the system will conduct a one-to-one comparison to determine whether the claimed identity matches with the user template. An example of verification task is a fingerprint enabled computer login system. When a user attempts to access the computer, the user needs to type a username and then present their finger. Their fingerprint sample is then compared with the fingerprint reference associated with the given username. If the samples match with each other the user will be granted access, otherwise they will be refused.

- **Identification**: (Whose identity is this?)

  In identification mode, the user does not need to claim an identity. The system will identify the user by searching a matched identity from all of the biometric reference templates stored in database. The system conducts one-to-many

comparison to determine the identity whose template was matched. This system is seeking to find an identity rather than verify a claimed identity. Therefore, the unique characteristics used in discriminating people for an identification need to be more unique or distinct than have used in a verification system. An example of Identification task is claiming benefit. Before a person claims social services benefit, their biometric information is checked on the system database, which contains sample from people who have already claimed the benefit enabling the detection of fraudulent claims.

The capability and effectiveness of biometric systems vary depending on the operational mode of the application required. Typically, the cost to implement, the complexity, and processing time of identification system are higher than verification system. However, the selection and implementation of biometric for particular application will mainly depend on purpose of the application. Within the application of mobile devices, being highly personal devices, authentication is performed in a verification mode, with the device only supporting a single user. Therefore, the remainder of the thesis will focus upon the application of biometrics within that context rather than identification.

A biometric-based system consists of the following five main components: data collection, feature extraction, storage, classification and decision. Descriptions of each component are explained below:

- **Capture component**

  This component captures the biometric sample from a user through using a sensor. For example, a camera can be used for capturing images of a face for

face recognition or a fingerprint sensor can be used to image the rides and valley pattern on the surface of a user's finger for fingerprint recognition.

- **Feature Extraction component**

  The extraction module extracts a set of unique biometric data (referred to a feature vector) from the captured sample or image. For example, a feature vector of the position and direction of minutiae point of fingerprint image are extracted in the extraction module of fingerprint-based biometric systems.

- **Storage component**

  The storage is used by the biometric system to store the feature vector or biometric template for future comparison during authentication. The templates in the database can be updated over time. The storage component of an identification system is far more sophisticated than a verification system because it needs to provide fast indexing and searching of the database.

- **Classification component**

  In the classification process, the biometric sample template and the stored template are compared during authentication and the level of similarity or a matching score is generated. Typically, the higher the score value representing a higher probability that the two biometric measurements come from the same person.

- **Decision component**

  The final component of a biometric system is the decision process. This process uses a matching score that is generated from a classification process to make a system decision. The decision threshold is selected to determine whether the claimed identity is confirmed (verification) or a user's identity is established (identification) or not.

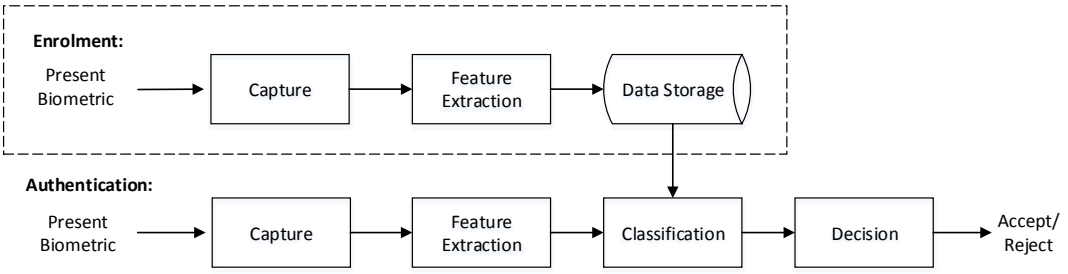


Figure 3.1: A generic biometric authentication system

In general, a biometric-based system consists of two stages as shown in Figure 3.1: the enrolment process and the authentication process.

- **Enrolment process**

  This is the user's registration process within a biometric system. The biometric sample of a user will be collected and used to create a reference template. The reference template is the essential key for the success of the biometric system. As such it is imperative that only the authorised user provides a sample at this stage and that the sample if of good quality. The quality of the sample is evaluated to ensure it is sufficient to be used for further processing. The extraction process then generates the reference template by extracting features that as the system requires from the sample. If the quality of the sample is poor, the authorized user is requested to present the biometric data again. The reference template is then stored in a database to be used for the authentication process. The reference template can be extracted from a single sample or more generally is generated based upon multiple samples.

- **Authentication or verification process**

  This is the process that determines whether the claimed identity matches with the reference template. At the time that the user requests access to a system, a new sample is acquired from the sensor. The quality of the sample is evaluated

to ensure it is sufficient. If it is, the features are extracted from the sample by a feature extraction process to create the sample template. The sample template is subsequently compared to the reference template(s) (one-to-one for verification, one-to-many for identification). However, biometric data of the same person taken at different times (two biometric data of user's left index finger) are not exactly the same due to many factors such as the quality of images (sensor noise), change in user's physiological or behaviour characteristics (cut on the finger), environment conditions (light and temperature) and user's interaction with the sensor (finger placement). Therefore, the comparison process generates a match score or a probability of similarity between the sample template and reference template. The higher score value represents the more similarity of sample and reference template come from the same person. Finally, the system decision threshold of match or no-match is selected and used to validate a claimed identity. In a verification system, the decision policy can be formally illustrated as follows:

$$(I, X_G) = \begin{cases} True\,, & if\ S(X_G, X_T) \geq t, \\ False, & otherwise \end{cases}$$

Where

(I, X_G) determines if a sample template X_G and a claimed identity I is true or false

X_G is a sample template which is a feature vector extracted from the sample biometric data

X_T is a reference template which is a feature vector stored in database

S is the function that measures the similarity between sample template X_G and X_T

$S(X_G, X_T)$ is a match score value between a sample template X_G and the reference template X_T

t is a predefined decision threshold

From the above decision policy, it can be said that the claim identity will be true (a genuine user) when a match score between biometric sample template and reference template is equal or more than the decision threshold (t) otherwise the claimed identity will be false (an imposter). Therefore, the result of verification is based on the variables: sample template (X_G), reference template (X_T), claimed identity (I), a predefined decision threshold (t), and similarity function (S). It should be noted that selecting a decision threshold is an important issue because a poor selected threshold could affect the security level of the system significantly.

3.2.2 Biometric system performance

To evaluate the performance of a particular biometric system, three important error rates are represented: the False (unauthorised) Acceptance Rate (FAR), the False (authorised) Reject Rate (FRR) and the Equal Error Rate (EER).

- False Rejection Rate (FRR)

  A statistic that represents the percentage of times the system produces a false rejection, which occurs when a biometric sample of an authorised user is not matched with the stored template and then rejected by the system. This error rate is defined as follows:

$$FRR = \frac{Number\ of\ genuine\ Rejects}{Number\ of\ Genuine\ Attempts}$$

- False Acceptance Rate (FAR)

  A statistic that represents the percentage of times a system produces a false accept, which occurs when a biometric sample of imposter is matched to a stored biometric template and accepted by the system. This error rate is defined as follow:

$$FAR = \frac{Number\ of\ imposter\ Accepts}{Number\ of\ imposter\ Attempts}$$

There is a trade-off between FAR and FRR, if one error rate decreases the others will increase. Figure 3.2 illustrates trade-off between FAR and FRR. The security level of a biometrics system can be controlled by adjusting the decision threshold setting. By setting a high threshold, the FAR error rates can be close to zero, and similarly by setting a low threshold, the FRR rate can be close to zero. To provide a high security, a system can be achieved by setting a high threshold, which means allowing few unauthorised individuals to be verified by the system. While a slack system can be achieved by setting a low threshold, which means allowing more unauthorised individuals to be verified by the system.

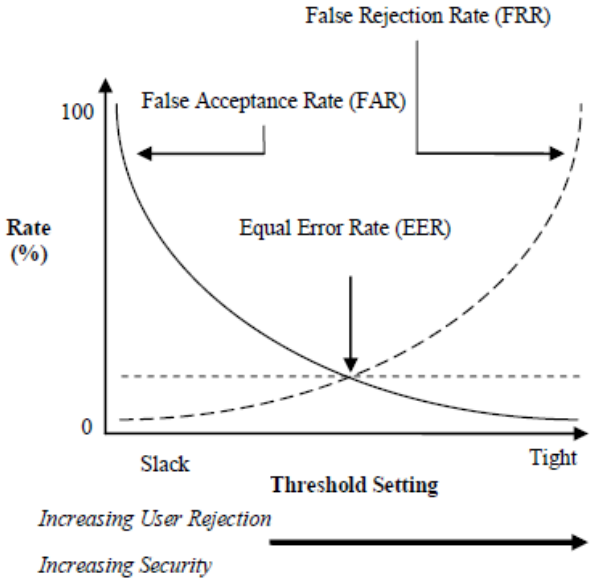


Figure 3.2: FAR/FRR Performance Rate

To compare the performance of different biometric techniques, a universal measurement parameter such as Equal Error Rate (EER) can be employed. The EER is the location on a curve where the false accept rate and false reject rate intersect, as illustrated in Figure 3.2. A system with low EER is more accurate and better performance system.

3.2.3 Biometric requirements

People have a large number of biometric characteristics such as face, fingerprint and gait. However, biometric characteristics that can be used in a biometric application exhibit varying levels of the following properties (Jain *et al*, 2004):

- **Universality**

  Every person using the technique should have the characteristics. For example, people need to have fingers for the fingerprint method to be used.

- **Uniqueness**

  The characteristics should be sufficiently different for an individual amongst the population such that they can be discriminated from one another. For example, the fingerprint of a person is more unique than eye color however, less unique than an iris or retina.

- **Permanence**

  The characteristics should not change over time. The fingerprint of a person tends not to change over time; however, the way a person walks tends to change depending on footwear, ground, tiredness and injury.

- **Collectability**

  The characteristics should be easy to capture using a suitable sensor. For example, capturing a face image can be achieved using a normal camera in a few seconds while capturing an iris image requires a user to position the eye very close to the scanner lens and stay still at that time.

In a practical biometric system, there are a number of other factors that should be considered including:

- **Acceptability**

  People should accept the system and be willing to present their biometric characteristic to the system in their daily lives. For example, people would prefer a normal camera than a retina scans, as some people believe that the latter technique could damage their eyes.

- **Performance**

  The system should meet the specified recognition accuracy, speed and resources requirements.

- **Circumvention**

  The system should be sufficiently robust to various fraudulent techniques such as sample forgery.

As a result, the ideal of a biometric system should fulfill all the above requirements however, currently no existing biometrics do.

3.3 Biometric characteristics

Based upon their primitive characteristics, biometric systems can be subdivided into two categories: physiological and behavioural biometrics. Physiological characteristics are related to some physical attribute of human such as their fingerprint, their face or hand. In contrast, behavioural characteristics are the ways in which the human behaves such as how they write their signature (i.e. hand writing), how they walk (i.e. gait) from one location to another and how they speak (i.e. voice).

Within this research, biometric techniques that led themselves to mobile devices and transparency are of particular interest. Many mobile devices come equipped with a number of hardware components that are able to be used for capturing a variety of biometric characteristics. As a result, a number of biometric techniques can be deployed on them as shown in Figure 3.3.

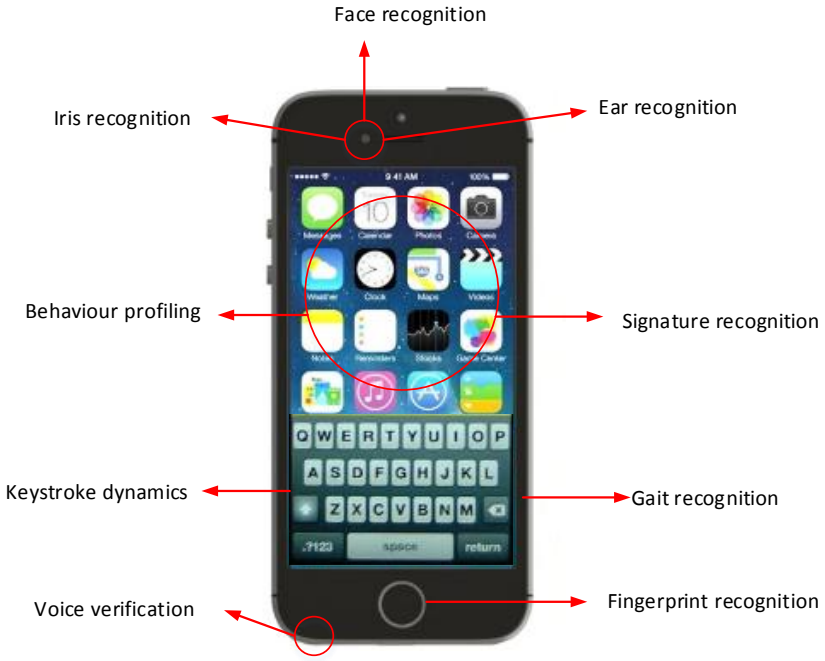


Figure 3.3: Biometric techniques on mobile phone

The built-in camera that can be used for taking a picture, shooting videos and making video conference calls creates the opportunity to use facial recognition and ear recognition on the mobile device. Given the high quality camera, iris and retina recognition could also be utilised. The microphone for making telephone calls would provide the potential for voice verification and the keypad would allow a keystroke dynamics technique to be applied. For a handset without a keypad, a touch sensitive screen could be used for utilising signature recognition. An overview of a number of biometrics of interest is described below:

3.3.1 Physiological Biometrics

The physiological biometric based systems utilise the characteristics of a human body part to identify an individual. Physical features are more likely to stay more constant over time and under different conditions. For example, a fingerprint or retina patterns of people will not be affected by mood; age (to a certain extent) or weather conditions. In addition, physiological biometrics typically contain high levels of discriminative

information and thus exhibit high levels of recognition performance. For example, iris patterns of individuals are so unique that even identical twins have different iris patterns. As a result, physical-based biometrics can be utilised in both verification and identification systems.

- **Fingerprint Recognition**

  Fingerprint recognition is one of the most widely used biometric technologies as it is a mature and proven technique to identify users because of their uniqueness and permanence (Maltoni *et al* 2003). It is has been determined that these patterns are so unique; even on each finger of the same person and the patterns of identical twins and tend not to change over time (Jain *et al*, 2002). Therefore, many sources (ten fingers) are available for collection. This technique utilises the unique characteristics based upon the minutiae or pattern of ridges and valleys on the surface of the finger (Lee and Gaensslen, 1991). A number of fingerprint recognition algorithms have been proposed in previous studies such as pattern matching or ridge feature based technique, minutiae based algorithm and correlation based approach (Lindoso *et al*, 2005; Cha, 2009; Maddala *et al*, 2011). The accuracy of the fingerprint technology has been shown to be very high so that it is sufficient for both verification and identification (Wilson *et al*, 2004; Jain *et al*, 2007). So far, this technique has been developed in many applications for the purpose of verification, such as access control (e.g. computer and mobile phone login system), and identification, such as immigration and law enforcement systems. Unfortunately, the system performance could be reduced when the finger is covered by dirt or has suffered a small cut (Jain *et al*, 2004). Since this

technique requires users to physically swipe their finger across a sensor to capture the image, people find this technique rather intrusive.

- **Face Recognition**

  Face recognition is a technique that identifies and verifies people using their face as face is the most common biometric feature that is used by humans to recognise one person from another. This technique measures the location and shape of facial attributes such as the eyes, lips or the overall analysis of the facial image that represents a face as a weighted combination of number of canonical faces in order to discriminate people (Jain *et al*, 2007). Several methods have been proposed to perform the classification, for instance, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Elastic Bunch Graph Matching (EBGM) (Turk and Pentland, 1991; Wiskoot *et al*, 1997; Jing *et al*, 2004; Yang *et al*, 2007; Wright *et al*, 2009). The facial recognition technique has a variety of potential applications in information technology security, law enforcement and surveillance, smart cards, access control, among others (Zhao *et al*, 2003; Face-rec, 2011; AxxonSoft, 2011;) and (Toshiba America Information Systems, 2011). The facial recognition technique is non-intrusive because a face image can be taken from a distance without any user interaction. Therefore, authentication can be performed without knowledge to the user. Unfortunately, the external factors such as lighting, expression, facial accessories (e.g. glasses, hats and facial hair) and pose can affect the performance of the system (Abate *et al*, 2007).

- **Iris recognition**

Iris recognition identifies a person using unique iris patterns, which has quite a complex structure that comprises features such as arching ligaments, furrows, ridges, crypts, rings, corona, freckles, and a zigzag collarette (Daugman, 1993; Wildes, 1997; Basit and Javed, 2007). The original iris patterns are generated after three months of birth and are unique characteristics of an individual (Daugman, 1993) even the irises of identical twins are different and everyone's left and right iris is also different (Wildes, 1997). The iris pattern remains stable over a person's life but could be affected by several diseases. A major approach for iris recognition is to generate feature vectors corresponding to individual iris images and to perform iris matching based on some distance metrics such as hamming distance, weighted Euclidean distance and the most famous algorithm using iriscode (Daugman, 1993; Sun *et al*, 2006; Miyazawa *et al*, 2008; Hollingsworth *et al*, 2009; Kekre *et al*, 2011). The iris recognition is less intrusive as the iris image can be taken from the distance up to 3 metres (Du, 2006). Since there is no physical contact with the camera, the iris scan can be performed safely and hygienically. However, the system does require users to align their eyes with the camera which may cause a certain level of inconvenience. Iris recognition performs very fast computing and accurate results (Wildes, 1997), but may fail in non-ideal conditions, such as a very constrained environment (Daugman, 1993; 2007). As the iris has a very stable, unique and non-intrusive nature, it is now considered as the most secure and reliable biometric method and it is therefore used for many secure and reliable security systems such as criminal investigation and citizen identification

agencies, financial services, access to PC and network based system, to name a few (Jain *et al*, 2007; UKBA, 2011).

- **Retina Recognition**

  Retina recognition is a method of biometric authentication that uses the unique pattern-recognition techniques based on the pattern of blood vessels at the back of the eye (Bhattacharyya *et al*, 2009). It is claimed to be the most secure biometrics since it is not easy to change or replicate the retina vasculature (Jain *et al*, 1999a; Delac and Grgic, 2004). The pattern is so unique that retina scanning based identification systems can achieve an error rate of 1 in 10 million which is 10 times better than the performance of iris recognition based systems (Wisegeek, 2011). Unfortunately, this technique suffers from the problem of user inconvenience and intrusiveness because when capturing retinal blood vessels, a user must position the eye very close to the scanner lens and stay still at that time. In addition, some people also believe that the retina scan damages the eye, which is another factor affecting the public acceptance of retina biometrics. All of these factors also affect the public acceptance of retina biometric (Delac and Grgic, 2004; Farzin *et al*, 2008). Due to the high cost of implementation, the technique is mainly deployed for maximum security requirement areas, such as governments, banks and the military (Biometric Newsportal, 2011).

- **Ear Shape recognition**

  Ear recognition utilises the shape of the outer human ear to identify a user. Ear recognition is a reliable technique to discriminate people through 2D and 3D images (Yuan and Mu, 2007; Yan and Bowyer, 2007; Bustard and Nixon, 2008,

2010). The ear shape is unique even for biological twins and it is a stable structure that does not change much with age and it does not change its shape with facial expressions (Iannarelli, 1989; Chen and Bhanu, 2007). There are many features that can be extracted from the ear including the outer rim(helix) and ridges(antihelix) parallel to the helix, the lobe, the concha (hollow part of ear), and the tragus (the small prominence of cartilage over the meatus) (Chen and Bhanu, 2007;). Based upon individual system preferences, these features are classified by using one of the following methods: set of point matching, edge shape matching and area matching (Bustard and Nixon, 2008; Pflug and Busch, 2012). Ear recognition is considered as user convenient because the ear image can be taken from a distance. Unfortunately, external factors such as lighting, accessory and pose can affect the performance of the system. Research by Burge and Burger (2000) suggested that by employing an expensive infrared camera, a certain level of identification can still be obtained by capturing a thermogram ear image. Currently, there are no commercial ear recognition systems available and authentication of individual identity based on ear recognition is still a research topic (Ross and Abaza, 2011)

Apart from the aforementioned physiological biometric techniques, other physiological characteristics such as DNA, body odour, fingernail bed and hand vein have been proposed and investigated as alternative methods to discriminate individuals in the future (Bhattacharyya *et al*, 2009).

3.3.2 Behavioural biometrics

Behavioural biometrics identify a person based on their unique behaviour, such as the way they write. Human behaviour is likely to change over time because users act

differently depending on mood, illness, stress, previous events and environment, to name a few. As a result, the discriminatory characteristics also tend to change, affecting the performance of the system. Nonetheless, the impact can be minimized if the template is regularly examined and updated. Compared with physical biometrics, behavioural-based techniques are less unique however they are more flexible and user friendly. Although most behavioural biometrics is not unique enough to provide a reliable identification system, they are sufficient to provide a good accuracy for a verification system (Yampolskiy and Govindaraju, 2008). Moreover, behavioural approaches tend to lend themselves to being applied in a transparent mode more than their physiological countermeasures.

- **Signature Recognition**

  As the name implies, this behavioural biometric system utilises unique characteristics based on a user's signatures. Signature recognition systems can be performed in static and/or dynamic modes. Static signature authentication involves the use of static or geometry of the signature, whereas dynamic authentication also uses dynamic or behavioural characteristics including the number of signature strokes and direction, the pressure applied by pen, the acceleration and the overall size of the signature (Das, 2007; Maiorana *et al*; 2010). The latter approaches are more robust due to behavioural characteristics that are almost impossible to replicate. Signatures are changed over period of time and are influenced by physical and emotional conditions of an individual (Delac and Grgic, 2004). Since signatures have been used for decades as a verification method, the signature recognition system is considered to be non-intrusive. As a result, the technology could be widely

accepted by the users. The performance of signature recognition systems can be affected when behavioural characteristics of a signature are inconsistent. In addition, this technology involves the use of a pen and special tablet that users may have difficulties getting used to. Nevertheless, several organizations have adopted signature recognition to verification or authorisation including Chase Manhattan Bank (the first bank to do so), the Internal Revenue Service, which uses this technology to verify tax returns filed online, and Charles Schwab & Company, which uses it for new client applications (Das, 2007).

- **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 (Robinson *et al*, 1993). The distinctive characteristics used to discriminate between users include the cumulative typing speed, inter-key (the duration of interval between two successive keys), hold-time (the duration of interval between the pressing and releasing of a single key), the frequency of the keys are used and the sequence used to type a capital letter (Das, 2007). Keystroke dynamics does not require any additional or special hardware as the technique can be embedded within the keyboard system. Keystroke dynamics can be performed in static (text-dependent) and/or dynamic (text-independent) modes. In the static approach, a user's typing pattern is examined when certain keys are pressed (e.g. when entering a password). In the dynamic approach, a user is verified based upon their overall typing pattern (e.g. the typing rhythm speed). Keystroke rhythm

authentication has some drawbacks such as alcohol consumption, drug use and fatigue and broken or damaged hands can change the speed of keystroke entry (Bartholomew, 2008). Keystroke dynamics (static mode) has been used as an additional layer of protection to existing password based access control mechanisms. For example, BioPassword, an existing commercial product, requires users to type their user names and passwords in a precise way to be logged into a system (Times Newspapers, 2007).

- **Voice Verification**

  Voice verification is also known as speaker recognition. The technique utilises a combination of physiological and behavioural characteristics based on a person's voice to discriminate between speakers. These features include the shape of the vocal tract, the movement, manner, and pronunciation of speech (Newman, 2009). Similar to keystroke dynamics, voice verification can be performed in static (text-dependent) and dynamics (text-independent) modes, again the former approaches are a simpler task than the latter. A static voice verification system requires a user to speak a predefined phase, which is also known as the "pass phrase". Dynamic voice verification does not require any pass phrases to identify a user. Instead, systems continuously monitor their speech behavioural characteristics (e.g. rhythm). Compared to a static approach, the dynamic approach provides more convenient because the user can speak freely to the system. However, the performance of text-independent tends to be worse that text-dependent system (Doddington *et al*, 2000). Voice verification can be used for both identification and verification purposes (Campbell, 1997). The advantages of this biometric technology are that easy to

use, no special hardware required as mobile devices come equipped with a microphone and no special training needed for the user. However, factors such as background noise, mood, medication and physical change of the vocal tracts can affect the performance of the system. Recently, speaker recognition was used by Barclays Wealth to verify the identity of telephone customers within 30 seconds of normal conversation (Business Wire, 2013).

- **Linguistic Profiling**

  Linguistic profiling is a behavioural biometric that attempts to identify and discriminate users based on linguistic morphology (Halteren, 2004). In the linguistic profiling technique, a large number of counts of linguistic features 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, and information content or item distribution throughout a string of text (Holmes, 1994; Stamatatos *et al*, 2001; Zheng *et al*, 2006). A number of studies utilised the techniques for both author identification and verification purposes (Stamatatos *et al*, 2001; Keselj *et al*, 2003; Zheng *et al*, 2006).The advantages of linguistic profiling are that it does not require any additional or special hardware and can be implemented in non-intrusive and continuous verification mode. The accuracy of current techniques depends mainly on the number of candidate authors, the size of texts and the amount of training texts. However, this technology is not reliable enough to meet the court standards in forensic cases (Stamatatos, 2009).

- **Gait Recognition**

  Gait recognition attempts to identify an individual based on the way they walk. Currently, three approaches have been developed for gait recognition: Machine Vision (MV) based gait recognition, in this case the walking behaviour is captured on video and video processing techniques are used for analysis; Floor Sensor (FS) based gait recognition by placing sensors in the floor that can measure force and using this information for analysis; Wearable Sensor (WS) based gait recognition, in which scenario the user wear a device that measures the way of walking and recognizes the pattern recognition for recognition purpose (Bours and Shrestha, 2010). Gait recognition is a non-intrusive approach as the gait image could be captured from a distance without the user knowing and any cooperation from the user in order to capture samples unlike fingerprint recognition. However, a user' gait could change over a long period of time due to their age, body weight or injury. In addition, factors such as footwear, ground conditions and personal emotion can also affect the way of a person walks (Boyd and Little, 2005). As a result, the performance of gait recognition can vary. Gait recognition applications could potentially be utilized on a mobile phone, such as an iPhone, to verify the user (Tanviruzzaman *et al*, 2009).

- **Behavioural Profiling**

  Behaviour profiling aims to identify patterns of usage based upon the way a person interacts with applications or/and services (Furnell *et al*, 2001). For example, in a PC environment, behaviour profiling would include the monitoring of the usage of applications with other factors such as which

application a person used, when and for how long in addition to utilising other factors (Aupy and Clarke, 2005). Based upon mobile phones this would include the monitoring of user's calling features (e.g. the day of calling, start time of a call, duration of a call, dialled telephone number and the location), device usage and Bluetooth scanning (Li *et al*, 2010). This technique is not expected to be unique and distinct enough to use for an identification system however, it is a non-intrusive approach and can be used to continuously monitor the identity of users while they work on their computer or use a mobile device. Currently, there are no commercial authentication systems utilising this technique. Several companies are utilising behavioural profiling for fraud protection on credit card and mobile telephony systems (Gosset, 1998; Stolfo *et al*, 2000)

In general, physiological biometric techniques tend to perform better on uniqueness, permanence and performance. However, these methods are considered to be intrusive, as they require some level of physical contact with users. In contrast, behavioral characteristics tend to change over time but they are able to perform in verification mode. In addition, behavioral biometric techniques tend to have better levels of acceptability, user friendliness and have less intrusive.

Based upon the biometric system requirements mentioned in Section 3.2.3, Jain *et al* (2004) a comparison of all the aforementioned biometric approaches is presented in Table 3.1 (H, M and L represent High, Medium and Low respectively). Table 3.1 shows that none of the biometric approaches outperforms any of the other approaches based upon all seven requirements. For example, iris recognition is one of the most unique biometric approaches, stays permanent and is extremely difficult to reproduce but people may find it hard to accept this technology due to the difficulty in capturing

a good quality biometric sample. In comparison, voice verification tends to have a very high acceptability because it is easy to acquire, however, its permanence is much poorer as a user's behaviour is likely to change over time.

	Biometrics approaches	Universality	Uniqueness	Permanence	Collectability	Performance	Acceptability	Circumvention
Physiological	Ear recognition	M	M	H	M	M	H	M
	Face recognition	H	L	M	H	L	H	H
	Fingerprint recognition	M	H	H	M	H	M	M
	Iris recognition	H	H	H	M	H	L	L
	Retina recognition	H	H	M	L	H	L	L
Behavioral	Signature recognition	L	L	L	H	L	H	H
	Voice verification	M	L	L	M	L	H	H
	Gait recognition	M	L	L	H	L	H	M
	Keystroke dynamics	L	L	L	M	L	M	M
	Behavioral profiling[1]	M	L	L	H	L	H	H
	Linguistic profiling[1]	L	L	L	M	L	M	M

Source: Jain *et al*, 2004

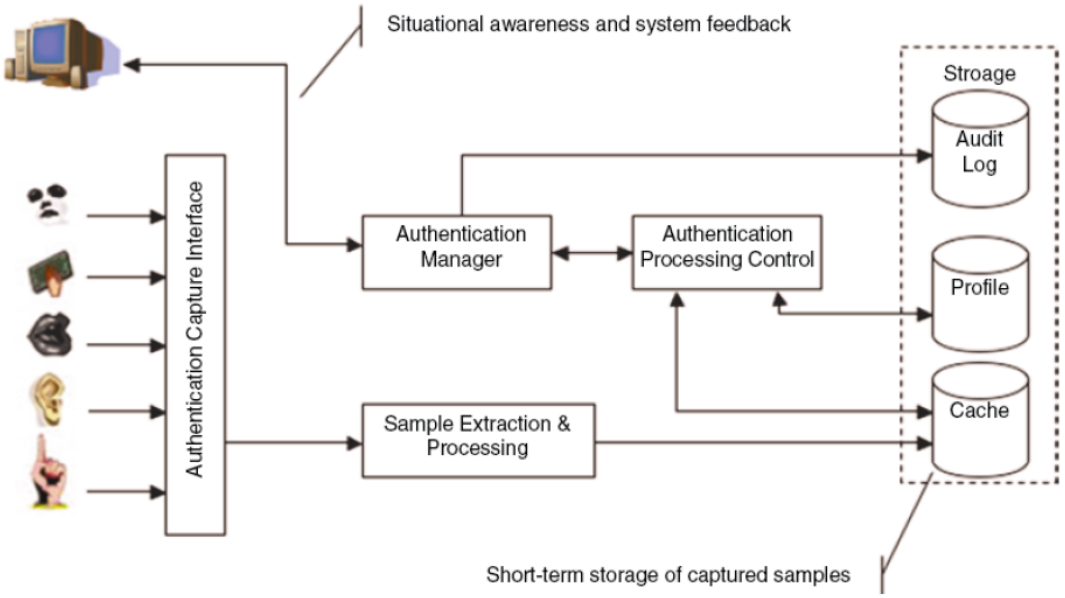
Table 3.1: A brief comparison of biometrics approaches

3.4 Biometric techniques for transparent authentication on mobile devices

Point-of-entry authentication is a useful approach in scenarios where one-off intrusive verification of the user is necessary. However, the approach is considered to be unsuitable to be used on mobile devices because it is unable to provide ongoing persistent verification of the user. As mentioned above, mobile devices are now capable of utilizing a number of biometric techniques so that it provides the opportunity for a Transparent Authentication System (TAS) to be implemented. According to Clarke (2011), a TAS is a framework for providing transparent, continuous, risk-aware, user-convenient and robust authentication. The TAS non-intrusively authenticates users and

[1] Inserted by the author

provides a continuous confidence measure in with respect to the identity of the user. When the confidence level is high, the user is given open access to all services and resources. However when the confidence level is low, access to sensitive services and information can be restricted until reassurance of a user's identity has been confirmed. One example for TAS is the Non-Intrusive and Continuous Authentication (NICA) (Karatzouni *et al*, 2007).



Source: Clarke, 2011

Figure 3.4: A generic TAS framework

As shown in Figure 3.4, the TAS is capable of utilizing a variety of biometric techniques. The availability of biometric capture would entirely depend on the available hardware incorporated into the mobile device (e.g. camera, keyboard). When the user interacts with the handset, it is feasible to suggest that certain hardware components can be utilized to autonomously capture biometric data. For example, a built-in front facing camera can be used for capturing photographs of the user whilst video conferencing and subsequently using facial recognition. The following biometric techniques have been proposed as viable non-intrusive, authentication mechanisms:

- **Keystroke Dynamics**

  While a user is typing text messages, writing a document or entering a password, their keystroke activities can be used for keystroke dynamics. Several researchers such as Clarke and Furnell (2006), Buchoux and Clarke (2008) and Maiorana *et al* (2011) have investigated the ability to authenticate users by using keystroke dynamics on mobile devices. The researchers found that keystroke dynamics could be deployed in practice on a mobile phone to verify a user with an acceptable average EER of 13%.

- **Face recognition**

  Face recognition can verify users when they make a video or conference. Research by Weinstein *et al* (2002) studied the ability to authenticate a user using facial recognition on mobile devices. A user's facial image was captured using a built-in mobile phone camera and then sent to a computational server for further processing (i.e. feature extraction and comparison process). In line with previous studies, Clarke *et al* (2008) and Dave *et al* (2011), all showed the ability to deploy face recognition on mobile device with the results of accuracy are between 75%-95.6% depending upon individual approach and dataset.

- **Handwriting recognition**

  A user's identity can be verified when they perform their signature (static) or while they write a message (dynamic) by using a stylus. Research by Clarke and Mekala (2007) investigated the feasibility of applying the signature recognition to handwriting verification where the signature input is placed by a series of commonly used words to provide transparent authentication. The best

performance was better than other behaviour techniques with a low EER of 1%, highlighting the possibility to deploy the technique on mobile devices.

- **Behaviour profiling**

  The technique monitors the interaction of the user with device based, for instance, on application use, frequency and timing of use. Research by Li *et al* (2010) investigated the ability to classify users the way in which they utilise services and applications on mobile devices. The result showed that behaviour profiling of voice calling, device usage and text messaging do have the ability to correctly classify users in an acceptable rates with EER of 5.4%, 13.5% and 2.2% respectively, highlighting the potential of developing the approach on mobile devices.

- **Voice verification**

  A user's identity can be authenticated during a conversation by using voice verification. Woo *et al* (2006) and Hazen *et al* (2007) investigated the feasibility of using voice verification on mobile devices. A user's voice samples were captured using a mobile handset and then transmitted to a network server for further processing. By using a speaker-dependent speech recognition approach, the results showed that a mobile user's identity could be verified by their voice with an acceptable EER of 7.8 %.

As shown above, a number of biometric techniques have direct application within a transparent authentication mechanism, with a number of research studies specifically focusing upon the applicability of various biometric techniques as the concept of non-intrusive operation. The implementation of a particular approach has a significant effect over its applicability in transparency. For example, for mobile devices without a front-

facing camera, face recognition would not be able to provide transparency as a user would be required to turn the handset around. In the same manner, some authentication approaches are currently not workable for transparent authentication; however, future implementation could well change this. For example, fingerprint recognition is an intrusive approach that requires a user to swipe their finger on the surface of the sensor installed at specific places such as at the back or on the top of handset. However, if the sensor were installed in a location on the handset where a user's fingerprint would more naturally reside when being held. This could provide a mechanism for capturing fingerprint samples without requiring the user to intrusively provide the sample.

3.5 Conclusion

Biometrics offer the most secure and convenient approach compared to knowledge based and token based techniques as it does not rely upon users remembering anything, just being themselves. The physiological based biometric approaches provide stronger protection as they are highly unique between users and difficult to reproduce. The behavioural based biometric techniques tend to offer more transparent and continuous protection as they can be performed during user normal interaction with a device. To date many biometric techniques such as fingerprint recognition and face recognition have been implemented for the purposes of identification and authentication.

A number of biometric techniques are applicable for development on a mobile device to provide transparent and continuous protection such as keystroke dynamics, behaviour profiling and voice verification. However, for transparency to work effectively, a TAS requires a more comprehensive set of transparent biometric

techniques in order to ensure verification can be continuously performed. In that basis the techniques to utilise should be based on the regular use of the device so that no explicit interaction is required. In addition implementing the techniques based on integrated hardware in current and future devices can be very cost effective. Therefore, the next chapter will explore behavioural biometric techniques based on the most popular method for user interaction with a mobile device without additional hardware requirement. Furthermore, the potential technique will be introduced in order to integrate transparent biometric mechanisms to provide authentication for a wide range of user interaction.

4 A Review of Text-Based Behavioural Biometric Authentication Techniques

Chapter 3 has identified a number of biometric techniques that are appropriate for deployment on mobile devices to provide transparent authentication. Behavioural biometrics provides a number of advantages over traditional biometrics technologies. They can be collected non-intrusively or even without the knowledge of the user. Developing security techniques that can operate transparently without further hardware requirements can be cost effective. By utilising the techniques based on the regular used of the devices, the authentication can be performed continuously and no explicit interaction is required. Since texting is one of the most popular applications and services that user uses on a daily basis. Therefore, attention is given to three behavioural biometric techniques: linguistic profiling, keystroke dynamics and behaviour profiling. To understand the current state of the art of these three techniques, the chapter presents a comprehensive literature review and evaluation of these techniques to date. The chapter then proceeds to give further literature review focus on the use of multi-modal biometric techniques.

4.1 Introduction

According to Ofcom (2012), text based services are currently the most used method of communication on mobile devices with 68% of UK mobile subscribers using services such as SMS messaging, social networking applications, emails and instant messaging tools on a daily basis; when compared to voice based services which is used by 63% of subscribers. Indeed, text messaging is the most regularly used method for daily communications with friends and family as the average UK consumer now sends 50

texts per week or approximately 7 messages per day. The aforementioned evidence demonstrates that there has been a shift change in communication preference with people preferring to text as opposed to talking. Therefore, an authentication mechanism that can be performed based on text-based communication can provide robust security that will be able to ensure the identity of the user on a continual basis. In order to achieve this continuous authentication, it is necessary to utilise transparent authentication techniques which can authenticate users in the background in order to minimise user inconvenience and increase user acceptance. By being able to authenticate a user without their knowledge, the integrity of the system can be automatically maintained and monitored without the user's explicit interaction, until such time as the system deems an imposter is accessing the system. Furthermore, by utilising a variety of transparent authentication techniques, authentication can continue to be performed during a wider range of user interaction. As a result, an advance authentication system is capable of providing a robust, dynamic, transparent and continuous.

As outlined in Chapter 3, a number of biometric techniques can be identified as appropriate for deployment on mobile devices to provide transparent authentication including: behaviour profiling, facial recognition, keystroke dynamics, linguistic profiling, handwriting recognition and voice verification. However, the selection of biometric techniques for use in transparent authentication is dependent upon a number of aspects including the user's activity, hardware, computation, storage, acceptability, and data collection. The selected approach should be utilized based on the regular use of a mobile device so that no explicit interaction is required. In that basis, the techniques to utilize should be based on integrated hardware in current and future devices, which is

used during normal usage of the device. This can also significantly reduce the cost of additional hardware. Other factors such as the computational and storage requirements of the techniques are not considered because mobile devices today come equipped with high computing power and storage capacity comparable to basic desktop computers (and this is increasing on almost yearly basis). Therefore, mobile devices will not have problems processing the data required for enrolment and authentication. Furthermore, wireless-networking technologies on mobile devices would allow the use of a client-server topology for authentication – where the server will provide the responsibility for the computationally intensive task and storage of biometric templates.

Considering their potential for use within mobile devices under a texting scenario, facial recognition can be used on mobile devices, via a front-facing camera, to provide transparent authentication. However, the limitation of this technique is that it can only be employed on mobile devices with a built-in front facing digital camera. Handwriting recognition can be employed when the user is entering words using the transcribe function of a PDA. Unfortunately, the majority of mobile devices are unable to support handwriting recognition. The remaining applicable techniques: linguistic profiling, keystroke dynamics and behaviour profiling represent the most favourable technique due to the transparent nature of authentication. More interestingly, it is hypothesized that these three techniques combined together offer the opportunity to provide continued authentication that can be performed during a wider range of user typing interaction. The next section presents a brief overview of the research performed in linguistic profiling, keystroke dynamics and behaviour profiling before introducing the domain of multi-modal biometrics and fusion.

4.2 A Review of Linguistic Profiling Techniques

The studies into the authorship of documents have been conducted with the assumption that people have a characteristic pattern of language use, a sort of "authorial fingerprint" that can be detected in their writings. The first attempts to quantify writing style go back to 19th century with the pioneering study of Mendenhall (1887) on the plays of Shakespeare. The most seminal work in this field was conducted by Mosteller and Wallace (1964). Their research studied on the authorship of The Federalist Papers. They used a method based on Bayesian statistical analysis of the frequencies of a small set of common words (i.e. "and", "to") for discrimination between the candidate authors. The results showed that 12 disputed papers were written by Madison and this conclusion was accepted by historical scholars and subsequently became a milestone in this research field. During the last decade, research in this area has advanced, taking advantage of the state-of-the-art research in areas such as machine learning, information retrieval, and natural language processing. With the plethora of available electronic texts (e.g. e-mail messages, online forum messages and blog), this technique has been widely studied. A summary of literature and results of linguistic profiling in text documents during the last decade is illustrated in Table 4.1 Please note that the performance of linguistic profiling presented in the table is represented by the classification correct rate.

Study	Document types	Authors /#Texts	Linguistic Metrics	Approach	Classification Correct Rate (%)
De Vel *et al* (2001)	Email	3/156	Lexical	Neural N.	88.2
			Lexical +Syntactic		91.6
Koppel & Schler (2003)	Email	11/480	Lexical	Neural N.	47.9
			POS tags		46.2
			Idiosyncratic usage		50.0
			All features		60.0
			Lexical	Statistical	38.0
			POS tags		40.4
			Idiosyncratic usage		67.6
			All features		72.0
Halteren (2004)	Essay	8/72	Lexical	Statistical	93.0
			Syntactic		86.0
			All features		97.0
Gamon (2004)	E-Text	3/1441	Function word	Neural N.	89.0
			POS trigram		94.0
			Syntactic		90.0
			Semantic		91.0
			All features		97.5
			No semantic		97.0
			Shallow features		96.2
Zheng *et al* (2006)	E-Text	English 20/960	Lexical	Neural N.	86.1-89.4
			Lexical +Syntactic		88.7-90.0
			Lexical +Syntactic +Structural		93.1-94.7
			All features		96.7-97.7
			Lexical	Statistical	78.1
			Lexical +Syntactic		79.1
			Lexical +Syntactic +Structural		88.7
			All features		93.4
		Chinese 20/740	Lexical	Neural N.	55.6-57.8
			Lexical +Syntactic		68.0-69.2
			Lexical +Syntactic +Structural		80.3-82.8
			All features		83.1-88.3
			Lexical	Statistical	52.2
			Lexical +Syntactic		66.2
			Lexical +Syntactic +Structural		72.5
			All features		74.2

Table 4.1: A summary of literature and results of linguistic profiling

Study	Document types	Authors /#Texts	Linguistic Metrics	Approach	Classification Correct Rate (%)
Goodman *et al* (2007)	Email	134/596	Lexical	Neural N.	80.0
Mohan *et al* (2010)	SMS	28/1400	Lexical	N gram	72.0
Iqbal *et al* (2010)	Email	5/200	Lexical	Neural N.	78.0
			Syntactic		76.0
			Structural		78.0
			Content-specific		73.0
			All features		88.0
Castro *et al (2011)*	Online-test	15/120	Keystroke +Lexical	Neural N.	84.0-94.0
Ouamour &Sayoud (2012)	Book	10/30	Character-n-gram +Word-n-gram Rare words	Neural N.	80.0

Table 4.1: A summary of literature and results of linguistic profiling (Cont.)

Research by De Vel *et al* (2001) studied the ability to discriminate between authors for the case of both aggregated email topics as well as across different email topics. The corpus of email documents used in the experiment contained a total of 156 documents sourced from three native English language authors, with each author contributing e-mails on three topics (approx. 12,000 words per authors for all topics). The topics chosen were movies, food and travel. The body of each email document was parsed based on an e-mail grammar proposed by authors, and the relevant e-mail body features were extracted. The body was pre-processed to remove any salutations, reply text and signatures. Attachments are excluded, though the e-mail body itself is used. A total of 170 style marker attributes (i.e. average sentence length, average word length, number of blank lines) and 21 structural features (i.e. has a greeting acknowledgement, number of attachments, contain signature text) were extracted from each e-mail document and subsequently analysed using a Support Vector Machines (SVM) classifier. The results indicated that an SVM classifier combined with the style marker and structural attributes is able to discriminate between the authors without consideration to the topic as well as when multiple topic categories are used. Indeed,

they observed that the style markers are the dominant features that contribute to the classification performance compared to the structural features. The performance of function word features was also investigated. In this experiment, a total of 320 function words were created and the set of these features were split into two categories: parts-of-speech (POS) words and others. Example of POS words included adverbs, auxiliary verbs, determiners, preposition and an others category included numbers and ordinal numbers (i.e. “first”). The results of this experiment showed no improvement in classification performance when including word collocation and even a reduction in performance when the function word dimensionality was increased.

Koppel and Schler (2003) proposed the use of syntactic information based on syntactic error and assess the usefulness of such features, both in and of themselves and in conjunction with other types of features, for authorship attribution. The corpus used in this experiment consisted of 480 emails written by 11 different authors during a period of about a year. Three classes of features, including lexical (i.e. function words: “the”, “and”, “that”), Part-of-Speech (POS) Tags (i.e. noun, verb) and idiosyncratic usage (i.e. syntactic, formatting and spelling usage) were considered. A standard set of 480 function words was used to create a lexical feature set. The appeal of function words could be used as a marker of writing style and may be an indicator of authorship. Function words are words such as propositions, articles that have little semantic content of their own and usually indicate a grammatical relationship or generic property. The POS tagger was applied to the entire corpus to tag each word with one of 59 POS tags and then used the frequencies of all POS bi-grams which appeared at least three times in the corpus as POS feature set. They proposed various writing error measures to capture the idiosyncrasies of an author’s style. To this end, a set of 99

spelling errors (e.g. letter omissions and insertions) and formatting errors (e.g., unbroken sequences of multiple question marks and other punctuation) was defined by using a commercial spell checker. Individual and various combinations of feature types were analysed using decision trees (C4.5) classification algorithms. The experimental results showed that the combination of three feature types is better than any one of them alone but of the individual feature types stylistic idiosyncrasies constitute the most effective type.

Gamon (2004) demonstrated that a combination of features based on shallow linguistic analysis (function word frequencies, part of speech trigram) and a set of deep linguistic analysis features (context free grammar production frequencies and features derived from semantic graphs) yields very high accuracy in a short random text. The experiment conducted was based on a total of 1441 text documents from three authors. Various types of linguistic information including function word frequencies, part of speech trigrams, n-gram frequency, syntactic information captured in syntactic rewrite rules and semantic information from semantic graphs were extracted from the documents. As a result, over 10,000 observed features were created. The author employed a simple frequency cut-off (where the frequency of a pattern that occurs less than *n* times is not included as a feature) in order to eliminate irrelevant features and used an SVM classifier for the classification stage. The experimental results showed that semantic features which constitute the most abstract and linguistically sophisticated class, combined with lexical and syntactic information improved the classification accuracy. Indeed, the author observed that the true stylistic and "content-independent" features produce a classification accuracy that outperforms the use of *n*-gram features by a wide margin.

Research by Halteren (2004) examined linguistic profiling to distinguish between texts written by different authors. The experiment employed the Dutch Authorship Benchmark Corpus (ABC-NL1), which consists of 72 Dutch text essays by 8 students. Each author was asked to write nine texts of about a page and a half and the topics for the nine texts were fixed. For each text, a lexical feature set that consists of 100k features and a syntactic feature set that consists of 900k active syntactic features were extracted to create a text profile. The deviation of a text profile was calculated and then compared to average of profiles for groups of texts. The experiment showed that lexical features perform much better than syntactic features with the ability to select the correct author 93% and 86% respectively. In addition, a profile system using combination of lexical and syntactic features can select the correct author with 97% accuracy. It is also able to perform the verification task in the way that it rejects no texts that should be accepted, and accepting only 8.1% of the texts that should be rejected. As a result, the study highlighted linguistic profiling can be used for both authorship recognition and verification purposes.

Zheng *et al* (2006) proposed a framework of authorship identification on online messages. In this framework, four types of writing-style features (lexical, syntactic, structural, and content-specific features) were extracted and an inductive learning algorithm was used to build feature-based classification models to identify authorship of online messages. The author conducted experiments on English and Chinese online-newsgroup messages. For the English message experiment, the database consisted of 20 authors and the number of messages collected for each author was varied from 30 to 92. The following features: 87 lexical features, 158 syntactic features, 14 structural features and 11 content-specific features were extracted. For the Chinese message

experiment, the database contained 20 authors and each author has 30 to 40 messages. Due to the language differences, some English features do not exist in Chinese. 117 features from English features were selected including 16 lexical features, 77 syntactic features, 14 structural features and 10 content-specific features. The discriminating power of the four types of features and of three classification techniques: decision trees (C4.5), back-propagation neural networks and support vector machines were compared. The experimental results showed that the three classifiers achieved accuracy of 90% to 97% and 72% to 88% for English and Chinese datasets respectively, when all features were used, highlighting the ability of writing-style features to identify authors of online messages with satisfactory accuracy for both English and Chinese dataset. Indeed, the structural features and content-specific features showed particular discriminating capabilities for authorship identification on online messages. Additionally, SVM and neural networks outperformed C4.5 and neural networks significantly for the author-identification task. However, different parameter settings (i.e. number of authors, number of training and testing data) of authorship identification had an impact on the performance.

Research by Goodman *et al* (2007) studied the use of Stylometry for Email author Identification. Stylometry is the study of the unique linguistic styles and writing behaviours of individuals in order to determine authorship. The experiment conducted was based on 596 emails messages from 134 users. For each email message, a total of 66 stylometric features were extracted such as the number of sentences per paragraph, average word length, number of words, paragraphs and average number of words per paragraph. For the classification process, the data was divided into two sets for training and testing. A Nearest Neighbour classifier using Euclidean distance was used

to compare the features of the test data set against those of the samples in the training data set. The top 10 training data sets with the smallest Euclidean distance to the test data set are identified. The unknown author name for each of test data set was declared using a majority voting approach. Based on twenty test emails, the experimental results showed that 80% of the emails were correctly identified the author using linguistic styles and writing behaviours of individuals.

Research by Mohan *et al* (2010) investigated an N-grams based approach for determining the authorship of SMS messages. The study employed a public SMS corpus provided by National University of Singapore. The final SMS dataset contained a total of 1400 messages from 28 authors. For the experiment, a total of 1750 data sets were created with different sizes of training sets. However, each user contributed the same number of messages to the training set and the testing set. The varying of gram sizes (between 1 and 5) was investigated. A number of classification techniques such as Euclidean distance, block distance, cosine similarity and matching coefficient were utilised. Authorship attribution of SMS messages using N-grams approach revealed positive results, achieving the best accuracy 72%.

Iqbal *et al* (2010) studied email authorship verification by using four types of linguistic features: lexical, syntactic, structural and domain-specific features. A total of 419 different features were extracted from each email. The study employed 5 authors from the Enron email dataset. For each author, 40 emails were selected. The experiment is evaluated by using three clustering algorithms: Expectation Maximization, k-means and bisecting k-means. The results showed a good correct classification performance with accuracy ranging from 77.6 % to 82.9%

Castro *et al* (2011) investigated the use a combination of keystroke dynamics and stylistic features for the purpose of authenticating online test takers. The study utilised timing features for keystroke and 55 lexical (49 from character based features and 6 from word based features) features extracted from exam documents. The experiment employed a dataset containing 15 students and each student answered 8 questions. The sizes of sample document are between 1,710 and 70,300 characters. The experiment evaluation was performed by using a K-Nearest Neighbour (KNN) classifier. The results showed that using a combination of keystroke and linguistic features achieved a good level of performance within 85%-94% accuracy. This study provides an indication of the usefulness of linguistic features for identity of online test takers.

Ouamour and Sayoud (2012) investigated the task of authorship attribution on Arabic texts. Several features such as character, characters-*n*-grams, words, word-*n*-grams and rare words were used as input of Sequential Minimal Optimization based Support Vector Machine. The experiment employed a database containing 10 ancient Arabic travellers and each author provided 3 different texts. The sizes of texts are between 209 words to a maximum of 800 words in the text. The results showed a good accuracy of 80% by using one of the following three features: the character bigram, character trigram and rare words. In addition, this study also showed that linguistic profiling for authorship attribution can also be applied in Arabic language by using the same rule of English language.

All of the studies have illustrated the potential use of linguistic profiling for authorship identification and verification. The best performance of systems all used a machine learning approach, but different choices of techniques and features used. The most common features used were lexical and syntactic features. By using a combination of

linguistic features including deep analysis features can improve the performance of accuracy. However, deep analysis features required more complex additional tools for their extraction.

The accuracy of current authorship identification technology depends mainly on the number of candidate authors, the size of texts, and the amount of training texts. Since the majority of previous studies tend to focus on long texts per author or multiple short texts. Usually, 10,000 words per author are regarded to be "a reliable minimum for an authorial set" (Burrows, 2007). However, there are many researches have focused on using a large number of short texts such as poems or students essay per author for training (Coyotl-Morales *et al*, 2006; Halteren, 2004). Some studies have shown promising results with short texts but the minimum requirements for a text have not been set. (Sanderson and Guenter, 2006; Hirst and Feiguina, 2007)

A number of studies focus on data size for authorship attribution. Research conducted by Hirst and Feiguina (2007) studied on an authorship attribution of short texts in works by Anne and Charlotte Bronte. They found that using multiple short texts (more than 2000) overcomes part of the obstacle of having only short texts, even when "short" means only 200 words per author. Stamatatos (2007) investigated the class imbalance problem and conducted several test experiments for compensation of imbalanced data sets. He concluded that the best method uses many short text samples for minority classes and less but longer ones for the majority classes. Sanderson and Guenter (2006) observed that the amount of training material has more influence on performance than the amount of test material. They also found that a minimum of 5000 words in training are required to obtain a relatively good performance.

Although a number of literature studies were available on using linguistic profiling for author verification based on short texts, a significant amount of research has proved that by using a pattern classification method an author can be classified by the way of their writing style. However, a large number of training texts are required to achieve good performance. In practice it can be seen that a large number of short texts with minimum 200 words might not be available in mobile-based communication and, furthermore, long emails are not likely to happen very frequently. Therefore a comprehensive study on linguistic profiling using mobile texts is needed.

Compared with formal text documents such as student essays, one challenge of author verification of SMS messages is the limited length of text messages, which is typically limited to 128 characters. The short length of SMS messages may cause some identifying features in normal texts to be ineffective. For example, since SMS messages do not conform to a fixed syntactical structure and are relatively unstable (vary from user to user), measures such as syntactic features may be not as effective as in previous studies and may also reduce the relevant features set for classification. Therefore, how to correctly verify the author of these short SMS messages with appropriate features becomes a challenge.

On the other hand, SMS messages also have some special characteristics, which may help reveal the writing style of the author. Since the length of text messages is limited and SMS messages are relatively informal compared with formal text documents, authors are more likely to leave their own “write prints” in their messages. For example, the users have to find a way of being concise their text message to communicate comprehensible messages within a limited length. The users may have to use shot spelling rather than using the standard spelling (e.g. “you” could be texted

as “u” or “y”), and often do not follow syntactic and/or standard grammatical rules. Therefore some of these special characters such as, unusual spelling words may be useful in forming a suitable feature collection for discrimination between users.

4.3 A Review of Keystroke Dynamic Technique

The research on keystroke dynamics has been conducted since the 1980’s, focussing based on QWERTY keyboard. However, factors such as the different layout, the use of smaller keys, the key shapes and the hand movement make keystroke dynamics for mobile handset significantly different from those performed over traditional keyboards. The study of the method on mobile phones began around 2002 (Clarke *et al*, 2002). Several studies have been undertaken in this area and all studies have attempted to solve the problem of providing a robust and inexpensive authentication mechanism. However, researchers tend to vary in their approach with respect to the keystroke information they utilise and in the pattern classification techniques they employ. A summary of key literature and results within the domain of keystroke dynamics on mobile devices is illustrated in Table 4.2. A number of the well-established studies on keystroke dynamics based upon keyboards have also been appended to the table for the purposes of comparison.

Study	Static/ Dynamic	Types of Keyboard	Features		Classify Approach	Users	FAR (%)	FRR (%)
			Inter-Key	Hold-Time				
Joyce & Gupta (1990)	Static	Keyboard	✗		Statistical	33	0.3	16.4
Leggett *et al* (1991)	Dynamic	Keyboard	✗		Statistical	36	12.8	11.1
Brown & Rogers (1993)	Static	Keyboard	✗	✗	Neural N.	25	0.0	12.0
Napier *et al* (1995)	Dynamic	Keyboard	✗	✗	Statistical	24	3.8 (EER)	

Source: Clarke, 2004

Table 4.2: A summary of literature and results on keystroke dynamics

Study	Static/ Dynamic	Types of Keyboard	Features		Classify Approach	Users	FAR (%)	FRR (%)
			Inter-Key	Hold-Time				
Obaidat & Macchairolo (1997)	Static	Keyboard	✗	✗	Statistical	15	0.7	
					Neural N.		0.0	
Monrose & Rubin (1999)	Static	Keyboard	✗		Statistical	63	7.9 (EER)	
Cho *et al* (2000)	Static	Keyboard	✗	✗	Neural N.	25	0.0	
Ord & Furnell (2000)	Static	Keyboard	✗		Neural N.	14	9.9	
Clarke & Furnell (2006)[2]	Static 4	Keypad	✗		Neural N.	32	8.5 (EER)	
	Static 11		✗		Neural N.	32	4.9 (EER)	
	Dynamic		✗		Neural N.	32	25.6 (EER)	
	Static			✗	Neural N.	30	18.0 (EER)	
Karatzouni & Clarke (2007)[2]	Static	Keypad	✗		Neural N.	50	15.8	9.1
				✗	Neural N.	50	34.2	36.8
Buchoux & Clarke (2008)[2]	Static	Keypad	✗		Statistical	20	57.5	15.0
			✗		Statistical	20	20.0	2.5
Campisi *et al* (2009)[2]	Static	Keypad	✗	✗	Statistical	30	16.3 (EER)	
Maiorana *et al* (2011)[2]	Static	Keypad	✗		Statistical	40	18.1 (EER)	
				✗	Statistical	40	22.9 (EER)	
			Combined		Statistical	40	4.4 (EER)	
Chang *et al* (2012)[2]	Static	Touch Screen	Combined		Statistical	100	12.2 EER)	

Source: Clarke, 2004

Table 4.2: A summary of literature and results on keystroke dynamics (Cont.)

The preliminary study undertaken by Clarke and Furnell (2006) investigated the feasibility of using keystroke dynamics as an authentication approach on mobile devices. Their study utilised two traditional keystroke features, namely inter-key time (the time between two successive key presses) and hold time (the time between pressing and releasing a single key). Two different types of input: numerical-based and alphabetic-based experiments were investigated. For both experiments, data collection was performed using a modified mobile phone handset (Nokia 5110), interfaced to a desktop computer through the keyboard connection. The pattern classification tests were performed with one user acting as the valid authorised user,

[2] Inserted by the author

whilst all the other users acted as imposters (a standard and well-accepted methodology applied to keystroke dynamics studies). A number of analyses were undertaken using a Feed Forward Multi-Layered Perceptron network (FF MLPs), Radial basis function network (RBF) and Generalised regression neural network (GRNNs) with different configurations.

The numerical-keystroke experiment utilised the inter-key time and was applied to static (number-dependent) and dynamic (number-independent) entry. A total of 32 participants were asked to enter three input scenarios including fixed 4-digit number (representing a PIN-type number), fixed 11-digit number and series of 11-digit number (representing the entry of a telephone number). For each dataset, 30 input samples were taken from each user in a single session. Two-thirds of these inputs were utilised in the generation of the reference profile, with the remaining used as a validation samples. The experimental results showed that the performance varied from an EER of 4.9% with the static telephone entry to an EER of 25.6% with the dynamic telephone entry. The performance of the dynamic-based technique for a classification algorithm is far poorer than the static-based technique as it requires a large number of training data. An individual user's performance varied with the different neural network techniques and configurations, with some users performing better with one than another. By utilising individually optimal neural network configurations, the performance improved with an EE of 8.5% and 4.9% for the four and eleven-digit numerical inputs respectively. Although these average performances showed promise, it is also noticeable that the performance drops as high as 35% for one particular user in this experiment.

The alphabetic-keystroke experiment utilised the hold time and was applied to static entry. In this particular study, the hold time is defined as the time taken from initial key press down to the final key press up (e.g. the letter 'c' requires the number 2 button to be pressed three times). A total of 30 subjects participated in the study, with each participant entering 30 text messages split over three sessions. Due to difficulty in ensuring that a user would enter enough of the character combinations in any single text to perform authentication so that this experiment utilised a static-based technique. As a result, five networks were generated with between two and six of most recurring characters (i.e. e, t, a, o, n, i), with each input representing one character. The experimental results showed that on average the gradual training schema proved most successful with an input vector of five characters achieving an EER of 18%. The best individual user performance achieved an EER of 7.2%. However, the worst individual user performance achieved an EER 42.6%.

Clarke and Furnell (2006) have shown the ability of both inter-keystroke latency and hold time for classification algorithms to collectively discriminate between the majority of users with a relatively good degree of accuracy. The use of larger input vectors resulted in a corresponding decrease in error rates as the amount of unique discriminative information and feature space increased (a well-established problem known as the curse of dimensionality (Bishop, 1995)).

A further study by Karatzouni & Clarke (2007) looked at other types of the interface and tactile environment. The research investigated the ability of keystroke dynamics based on mobile phone thumb-based keyboards. Both traditional keystroke features were analysed using Feed Forward Multilayer Perceptron Neural Network (FF-MLP) with different network configurations. Data collection based on user entered text

messages was performed using an XDA IIs handset (an early smartphone) rather than a simulated handset as in previous studies. Inter-key latency and hold-time were extracted and analysed. Six varying sized keywords were investigated to evaluate the performance (given the performance that static-based features showed over their dynamic counterparts in previous studies). The result showed that the inter-key latency gave promising results in-line with the previous study. The best performance was achieved by using the longest keyword with an EER of 12.2%. However, it appeared that a particular user achieved a high error rate with 32.4%. Although the hold-time has performed well on regular mobile phone keypads in previous research (Clarke and Furnell, 2006), this characteristic gave no promise of a potential use in this thumb-based keyboard. This was because the method calculating the hold-time was different. Moreover, the research found that the tactile interface of keyboard and the hand movement might also affect the timing of keystroke.

Whilst studies have been undertaken looking into application of keystroke dynamics on a mobile device, unfortunately these only used the mobile device to capture samples and then subsequently using desktop computing to analyse the data. Research by Buchoux and Clarke (2008) investigated the feasibility of deploying keystroke dynamics on a Smartphone. The inter-key of simple PIN and strong alphanumerical password latency were captured using Orange SPV C600 smartphone and analysed using two statistic-based and two neural network techniques. The study found that the compuatational requirements of neural networks exceeded the processing capabilities of the device. However, this is not the case as mobile devices power and storage capacity are increasing on an almost yearly basis. Furthermore, the study demonstrated that the use of a short 4-digit PIN would be ineffective for use

within keystroke dynamics. However, ensuring users use a longer PIN or combination of alphanumnerical password would enable keystroke dynamics to be applied.

Research by Maioranaet *et al* (2011) investigated the ability of keystroke dynamics for providing a verification system on mobile phones. A database consisting of six passwords, each ten alphabatical long, with a number of *K* key press/release events comprised between 20 and 26. A total of 40 users were ask to enter each password 20 times using a Nokia 6680 mobile phone during four distinct sessions. Based on 4800 records, the first 10 keystrokes captured (of each user) were used as an enrollment set while the remaining data was used to estimate the verification performance. Several experimental tests have been conducted using statistic Manhattan and Euclidean distance classifers. The discriminative capabilities of the keystroke dynamics features have been analysed and the template selection process was also applied. The results showed that keystroke dynamics is able to perform user authentication with even the number of enrolment acquisition is low. However, the performance of classification performed better when the number of enrollment samples increased. The most successful classifications implemented a combination of both inter-key and hold-time measures, illustrated that the use of both measures has a cumulative and constructive effect upon performance.

Chang *et al* (2012) research investigated the performance of keystroke dynamics to authenticate the user for touch screen mobile phone. The study proposed the use of keystroke dynamics based on the way of user typing graphical-based password to classify the user. A total of four keystroke time features were utilised: inter-key (1), hold time (2), the time interval between the stroke of key *i* and the stroke of the successive key (3), and the time interval between the release of key *i* and the release

of the successive key (4). The keystroke features were captured using three different touch screen mobile phone models. A total of 100 users participated in the study, with each user providing 10 samples for each 5 graphical passwords. For each user, five samples were collected at the same time through the same mobile phone and used in the enrolment process. The other five samples were collected over period of five weeks through the other two mobile phones. The experiments employed a computation-efficient statistical classifier to verify the user's identity. The results showed that keystroke features can be used to identify the user with an acceptable EER of 12.2%. Furthermore, an EER is reduced to 6.9% when the pressure feature is applied in addition to keystroke features. This study also demonstrated that the performance of the proposed technique will be not affected by inconsistent screen size.

Based upon aforementioned studies, keystroke dynamics does appear to perform well and the ease of collection makes the approach very transparent. In general, neural network based algorithms can be seen to outperform the more traditional statistical methods, and have become more popular in later studies. Traditional keystroke dynamics characteristics (inter-key and hold-time) have the potential to be used for user classification. Implementing a combination of both measures has a cumulative and constructive effect upon performance. A system that utilised more text prior to matching would perform better. It is also evident that static based approaches perform substantially better than their dynamic counterparts. However, Clarke and Furnell (2007) showed that implementing a static approach on commonly reoccurring text message words can provide the discriminative information required to authenticate users with a greater degree of accuracy. This technique is also word independent and would, under the majority of circumstances, achieve the requirement of continuous

and non-intrusive authentication. Furthermore, this technique can be combined with other authentication method to improve the performance of classification.

However, a number of studies have showed the failure of keystroke dynamics to perform successfully for a minority of users. These users tend to have few distinctive typing rhythms. As such, any authentication system that implements a keystroke dynamics technique would also have to consider the small number of users that will experience too high an error rate in order to ensure both security and user convenience factors required by the overall system are met.

4.4 A Review of Behavioural Profiling Technique

The research into mobile behaviour profiling started around 1995, focussing on the area of Intrusion Detection Systems (IDSs) to detect telephony service fraud. These IDSs create normal user profiles by monitoring user activities for a period of time and compares them against the current activity. If a significant deviation is observed, a possible intrusion is detected. More recently, research has also focused on the use of such approaches within an authentication system. A summary of key literature and results within the domain of fraud detection and authentication is illustrated in Table 4.3.

Study	Behaviour	Approaches	Detection rate (%)	FAR (%)
Area of Fraud Detection				
Samfat and Molva (1997)	Itinerary	Statistical	82.5	4.0
	Calls	Statistical	80.0	3.0
ASPeCT (1998)	Calls	Neural N.	50.0	0.0
Buschkes *et al* (1998)	Mobility	Statistical	87.5	NA
Boukerche and Nitare (2002)	Calls	Neural N.	97.5	4.2
Sun *et al* (2004)	Mobility	Statistical	87.5	15.0
Hall *et al* (2005)	Itinerary	Neural N.	50.0	50.0
Sun *et al* (2006)	Mobility	Neural N.	89.0	13.0

Source: Li *et al*, 2011

Table 4.3: A summary of literature and results on behavioural profiling

Study	Behaviour	Approaches	Detection rate (%)	FAR (%)
Area of Authentication				
Aupy and Clarke (2005)[3]	Device usage	Neural N.	7.1 (EER)	
Li *et al* (2011)[3]	App. Usage	Neural N.	13.5 (EER)	
	Text message	Neural N.	5.4 (EER)	
	Calls	Neural N.	2.2 (EER)	

Source: Li et al, 2011

Table 4.3: A summary of literature and results on behavioural profiling (Cont.)

Within the area of fraud detection, one of the most significant initial studies was undertaken by the European Advanced Security for Personal Communications (ASPeCT) project (Gosset, 1998), which sought to develop a fraud detection system for mobile communications. The type of features used to create user profiles were collected from a toll ticket which is issued by the network after each call and consisted of the following features: international Mobile Subscriber Identity (IMSI), start date of call, start time of call, duration of call, dialled telephone number and National or International call (Moreau *et al*, 1997). These features were constantly updated after each call and used to maintain a current behaviour profile. A number of techniques were used to study a user's calling behaviour such as supervised neural network and unsupervised neural network. Several experiments have been studied based on the dataset which was collected from the Vodafone network and contains a total of 317 fraudulent and 20212 legitimate users. The experimental results showed a detection rate of 50% and FAR of 0.02%. However, this result was achieved using a network-based approach which benefitted from a richer dataset than a host-based approach that could necessarily provide. Furthermore, utilising a network-based approach also had additional processing and memory capabilities that was limited to the host-based approach.

[3] Inserted by author

Research by Boukerche and Nitare (2002) proposed an online-security fraud detection model for the mobile phone using a Radial Basis Function (RBF) neural network. The model is able to detect possible call service fraud by using user-calling features. Once a possible fraud is identified, the model automatically sends alert notifications to both the individual user and network administrators immediately, rather than at the end of a billing month. The experiment dataset contains 4,255,973 telephone calls that were collected by an unnamed telecommunication service provider. A variety of experimental configurations have been tested with combinations of call features and RBF neural network setups. The result showed a low EER of 4.2% (obtained by using 110 neurons in the hidden layer of the RBF network).

Sun *et al* (2004) proposed a mobility-based anomaly detection model for cellular mobile networks by utilising the combinations of a high order Markov model, Ziv-Lempel data compression algorithm and Exponentially Weighted Moving Model (EWMA). The high order Markov model was used to calculate a mobile user's mobility probability from one cell to another and the Ziv-Lempel data compression algorithm was employed to create a mobility profile of the user. Since a users' mobility activity may change overtime, the EWMA technique was applied in order to modify the user normal profile constantly. The authors developed an intrusion detection algorithm by constantly comparing a mobile user's current mobility action with their mobility profile and used a threshold scheme to determine whether the mobile device is potentially compromised or not. The evaluation process was tested with different mobility speed levels by using a simulation method based on a cellular network containing 40 cells. The simulation results showed that the false alarm rate decreased with an increase in mobility speed. Conversely, the detection rate increased with an increase in the

mobility speed. The false alarm rates of the system were around 25% and 5% and the detection rates were around 80% and 95% when a user travels at a speed of 20 mile/hour and 60 mile/hour respectively. However, this system may not be accurate for some particular types of users who do not exhibit regular movement such as taxi driver.

The use of behavior profiling within IDSs is applicable because of the transparent capturing samples. By studying a user's calling or location activities behaviour-based IDS systems can achieve a high detection rate and offer the ability to detect an intruder. With the same analogy but different aims, research has also focused on the use of behavioral profiling for authentication. A preliminary study by Aupy and Clarke (2005) proposed a novel authentication technique based on the way in which the user exhibited unique behaviour whilst using their computer to provide transparent and continuous authentication. The study was conducted on a database of 21 participants over a capture period of 60 days. An application was developed that was capable of capturing system events so that a profile of the user can be built. A front-end application then was developed to extract the key features from the database. A Feed Forward Multi-Layered Perceptron (FF-MLP) was utilised for classification. Initial studies found that the four features on a single input vector were not sufficient for successful classification. As such, the researchers decided to utilise 300 actions as a single input, which corresponded to approximately 10 minutes of user interactivity. A smoothing function where one classification based upon three network outputs is also applied. Two passes and a fail equal a pass and two fails and a pass equal to a fail. The networks are generated and tested against one imposter user each time. The experiment showed encouraging results with all users achieving EERs of below 10% and

an overall average EER of 7%. However, these EERs are calculated by using testing data that also comprised of aspects of the training data; which is recognised as bad practice and will lend to results being presented more positively than could be expected in a real system.

Further research has also been undertaken on the development of behaviour profiling for mobile devices. Research by Li *et al* (2011) investigated the ability to classify users based upon the way they utilise services and applications on a mobile phone. The research employed a public dataset provide by the MIT Reality Mining project (Eagle *et al*, 2009). The dataset contains various mobile data attributes that were collected from participant's using Nokia 6600 mobile phones. The following experiments were analysed: the general use of application and specific use of voice call and text message applications. Two types of profile techniques: static and dynamic were employed. For the static profiling, each individual dataset was divided into two sets: the first half was used for creating the profile and the second half was used for testing. For the dynamic profiling, the profile contains 7/10/14 days of the user's most recent activities. A number of analyses were undertaken using Feed Forward Multi-Layered Perceptron network (FF MLPs) and Radial basis function network (RBF) with different configurations. Subsequently, a smoothing function that considers a number of entries is also employed.

For the general application experiment, the dataset contains a total of 101 individual applications with 30428 entry logs for 76 participants. Among those 101 applications, the following features were extracted from the dataset: application name, date of initiation and location of usage. The results showed the best EER at 13.5 % and it was obtained by using dynamic profiling data with 14 days of user activity with 6 log entries.

For the telephone call application experiment, the dataset contains a total of 13719 call logs with 2,317 unique telephone numbers from 71 participants. For each log, the following features were extracted: telephone number, date and location of calling. The best experimental result is an EER of 5.4 % and it was obtained by using dynamic profiling data with 14 days of user activity with 6 log entries.

For the text messaging experiment, the text messaging dataset contains 1382 logs and 258 unique telephone numbers. For each log, the following features were extracted: receiver's telephone number, date and location of texting. The best experimental result is an EER of 2.2 % and it was obtained by using dynamic profiling data with 14 days of user activity with 3 log entries.

The application name and location have proved to be valuable features that provide sufficient discriminatory information to be useful in authentication, with location being more significant characteristics. However, this might not necessarily identify the misuse for a particular user such as college, as the location information could fall within the same profile as the authorised user. The intra-application approach should help to specifically identify this type of misuse. In general, dynamic profiling achieved a slightly better performance than the static profiling as it contains a user's most recent activities. The smoothing function that treated more logs entries as one incident improves the overall performance of the technique. However, it takes a longer time for the system to make a decision so that an intruder could have more opportunities to abuse a system.

The ability to utilise behaviour profiling for network-based detection or for local PC and mobile devices has been presented. Although no commercial solutions exist outside of the domain of fraud, research studies have provided a good basis for suggesting this approach do have the ability to correctly classify users. However

template aging is a significant issue when looking to deploy the approach on its own. As such, any authentication system that implements behavioural profiling technique would also have to consider the elapsed time during the decision processing in order to ensure an appropriate balance between security and performance. Fortunately, this issue can be mitigated within multi-modal authentication system.

All of the following behavioural biometric techniques: linguistic profiling, keystroke dynamics and behaviour profiling can be implemented on mobile devices to provide continuous and transparent authentication. Since biometric data of all techniques can be captured during users typing interaction with keyboard therefore it is possible to combine these three approaches in order to provide authentication based upon any form of typed communications such as texting, email, social networking and twitter. The combination or fusion of using more than one biometric technique will be discussed in the next section.

4.5 Multi-Modal Biometrics

Authentication systems built upon multi-modal biometrics offer an opportunity to increase population coverage since multiple biometrics can solve issue with users who do not have or cannot present particular biometric traits. The use of more than one biometric technique enables the system to reduce the number of cases where the system is not able to make a decision. For example, if the user is unable to be verified by linguistic profiling due to insufficient features within the message, the user can still be verified using keystroke dynamics or behaviour profiling. The performance of the overall approach can also be improved for users who can provide all traits (Hong and Jain, 1999; Jain *et al*, 2004). Furthermore, the need to provide more than one sample will also help in preventing spoof attacks because an attacker would need to

circumvent more than one technique. Therefore, multi-modal biometrics can improve accuracy and reliability of single-modal systems.

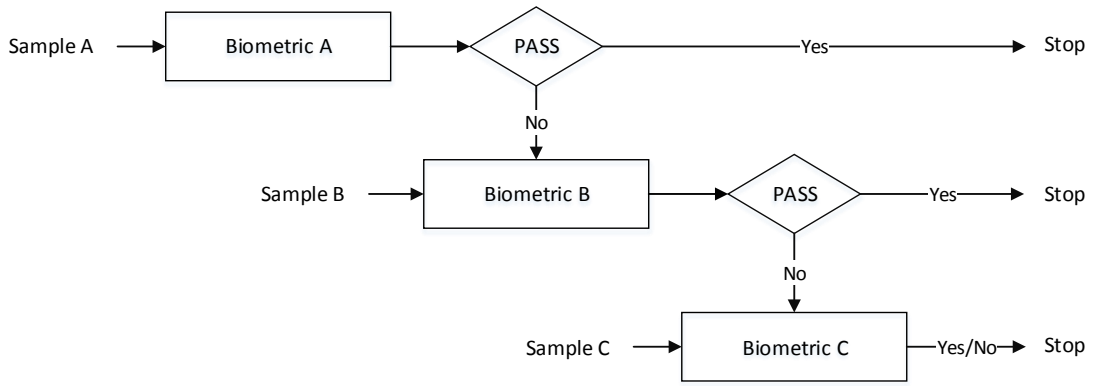
Multi-modal biometric systems require a combination of two or more biometric techniques and data. There are various scenarios that are possible in a multimodal biometric system including (Ross *et al*, 2006):

- **Multiple sensors**: the use of more than one sensor to capture a single biometric trait (e.g. optical and capacitive fingerprint sensors).
- **Multiple instances**: the use of more than one subtype of the same biometric (e.g. the left index finger and the right index finger).
- **Multiple samples**: the use of more than one sample of the same biometric (e.g. multiple face images of a person acquired under different pose/illumination conditions)
- **Multiple algorithms**: the use of more than one matcher algorithm in classification process (e.g. multiple fingerprint matchers based on minutiae or filtering)
- **Multiple biometric modals**: the use of more than one biometric modality (e.g. face, fingerprint and iris).

During each and every interaction with users, the biometric data captured will increase and incorporate a range of sample data from the limited number of sensors available on the mobile device. Therefore, it is probable that the multiple samples, multiple instances, multiple algorithms and multiple modals will be available in any given mobile device or scenarios.

The order or sequence in which biometric samples are acquired can be performed in a synchronous and asynchronous approach. Synchronous refers to capturing the biometric samples simultaneously or in parallel. Asynchronous approach refers to capturing the biometric samples sequentially. Both of these approaches could effectively be happening at any point in time as samples are captured continuously at the background. The processing of samples can also be performed in both synchronous and asynchronous methodologies.

In serial mode, the processing of samples is taking place sequentially as shown in Figure 4.1. If the sample A of the user could not be verified, the system can use sample B or C. When the system has sufficient confidence on the identity of the user after processing the first biometric sample, the user may not be required to provide the other samples. Thus a cascaded multi-modal biometric system can be more convenient to the user and generally requires less recognition time when compared to a parallel approach. However, for a transparent approach, which removes user convenience already, it is the latter parallel approach, that offers an opportunity to improve upon the verification performance of the weaker behavioural biometrics.



Source: Clarke, 2011

Figure 4.1: Serial mode of processing of biometric samples

4.5.1 Level of Fusion

In multimodal biometrics, the combination or fusion method can occur effectively at any point within the biometric system. A brief description of each of these fusion levels is presented in this section.

4.5.1.1 Feature level

If the features extracted from one biometric modality are independent of those extracted from the other, the two feature sets can be combined to construct a single features set to represent the individual, which is then passed to the classification and decision processes as a single biometric system. For example, a new multi-modal feature vector that is constructed by using the combination of the geometric features of the hand and the Eigen-coefficient of the face.

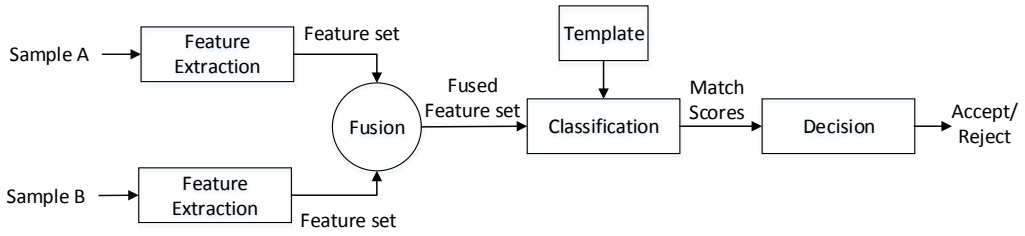


Figure 4.2: Feature level fusion

Concatenating the two feature vectors may lead to a feature vector with a very high dimensionality and can result in increasing complexity of the classification process. However, the new vector with high dimensionality does not necessarily result in an improved matching performance compared to that obtained by individual vector. Therefore, feature reduction techniques might be required to extract a small number of useful features from the larger set of features. There are several feature selection algorithms that have been used in previous studies (Jain *et al*, 2007) such as sequential forward selection, sequential backward selection, sequential forward floating search

and sequential backward floating search.

4.5.1.2 Match score level

The resulting outputs generated by each biometric classifier are combined to generate a new result that can be subsequently used to present to the decision process. For example, the match scores generated by the face and fingerprint modalities of a user are combined in order to create a new match score, which is then used to make the final decision.

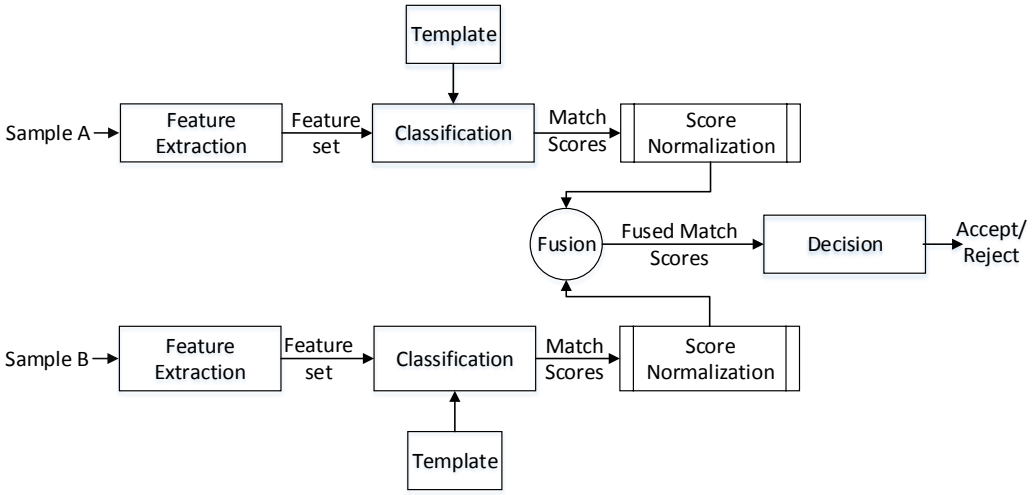


Figure 4.3: Match score level fusion

The approach enables multi-modal authentication with each modality being classified by a dedicated matching subsystem designed specifically for that approach. However, invariably the use of different classifiers results in different outputs being produced. If the range of the output result values vary then normalization is required. For example, the output of Feed-Forward Multi-Layered Perceptron (FF MLP) and Radial Basic Function (RBF) networks have an output that varies from 0 to 1 and 0 to infinity. The score normalization is a mechanism to ensure all outputs are bounded into the same interval so that they can be combined in a consistent manner. Many score

normalization strategies have been proposed in previous studies (Jain *et al*, 2005; Snelick *et al*, 2005) and it was claimed that fusion rules performance change by varying techniques. The well-known normalization methods used in previous studies are shown below:

- **Min-Max**

  This is the most simple normalization technique. The method maps the raw scores to the range between 0 and 1. The normalized scores are given by

$$n = \frac{s - Min(S)}{Max(S) - Min(S)}$$

  Where: s is a raw matching score from the set S

  S is a set of all score for that matcher

  n is the normalized score

  Max(S) is the maximum value of S

  Min (S) is the minimum value of S

- **Z-Score**

  The most commonly used score normalization technique. This method transforms the scores to a distribution with mean of 0 and standard deviation of 1. The normalized score are given by

$$n = \frac{s - Mean(S)}{Std(S)}$$

  Where: Mean(S) is the mean value of matching score S

  Std(S) is the standard deviation of matching score S

- **Tanh**

  This method is one of the most robust and highly efficient techniques (Huber, 1981). It maps the raw scores to the range between 0 and 1. The normalized score are given by

$$n = \frac{1}{2}\left[tanh\left(0.01\frac{(s - Mean(S))}{Std(S)}\right) + 1\right]$$

  Where: Mean(S) is the mean value of matching score S

  Std(S) is the standard deviation of matching score S

- **Decimal Scaling**

  This method can be applied when the scores of different matcher are on a logarithmic scale. For example, if one matcher has scores in range [0,1] and the other matcher has scores in the range [0,1000] then the following normalization could be applied.

$$n = \frac{s}{10^n}$$

  Where: n is the number of $\log_{10} Max(S)$

After the normalization method was applied, the resulting scores are then combined using one of the following methods:

- **Simple Sum**

  The scores from each modality were simply added to generate a fusion score.

$$f_i = \sum_{m=1}^{M} n_i^m \ , \forall i$$

- **Minimum Score**

  This method selects the minimum of an individual's scores to generate a fusion score.

$$f_i = Min(n_i^1, n_i^2, \dots, n_i^M), \forall i$$

- **Maximum Score**

  This method selects the maximum of an individual's scores to generate a fusion score.

$$f_i = Max(n_i^1, n_i^2, \dots, n_i^M), \forall i$$

- **Matcher Weighting**

  Weights are assigned to the individual matchers based on their Equal Error Rates. The weights are inversely proportional to the corresponding errors; the weights for more accurate matchers are higher than those of less accurate matchers. The matcher weight fused score for user *i* is calculated as

$$f_i = \sum_{m=1}^{M} \omega^m n_i^m \ , \forall i$$

  Where: ω^m is the weight assigned to the i th matcher and the weights are calculated by

$$\omega^m = \frac{\left(1 \Big/ \sum_{m=1}^{M} \frac{1}{r^m} \right)}{r^m}$$

  Where: r^m is the EER of matcher m and m =1, 2, …, M.

  Note that $0 \leq \omega^m \leq 1$, $\forall m$ and $\sum_{i=1}^{M} \omega^m = 1$.

4.5.1.3 Decision level

The fusion occurs at the end of the biometric system when each individual biometric system has provided an independent decision. Many different approaches have been proposed to combine the distinct decisions into a final authentication decision, for example, “AND” and “OR” rules (Daugman, 2000), majority voting (Lam and Suen, 1997), weighted majority voting, Bayesian decision fusion and behaviour knowledge space (Lam and Suen, 1995).

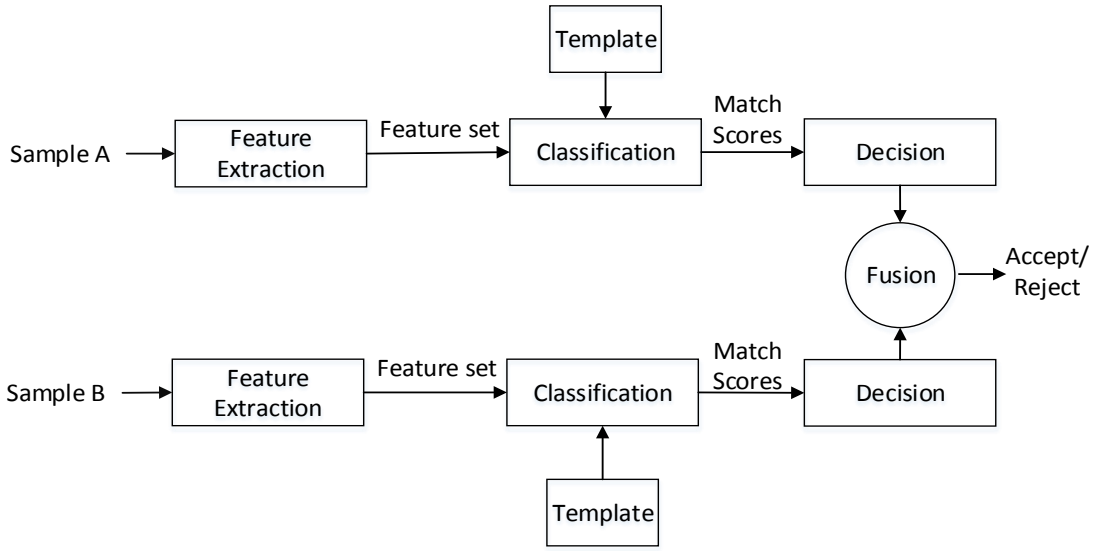


Figure 4.4: Decision level fusion

In this section, three different fusion strategies: feature level, match score level and decision level have been discussed. Fusion at feature level is difficult since the feature sets used by different biometric modalities may not be compatible. In such a case, integrations at the match score or decision levels are the only options. However, fusion at the decision level is too rigid since only Boolean information is available at this level. Therefore, integration at the matching score level is preferred due to presence of sufficient information content and the ease in accessing and combining match scores (Ross, 2007).

4.5.2 A Review of Multimodal Biometric Technique

A number of researchers have focused upon multimodal biometric systems over the last decade due to the enhancements in performance that can be experienced. Literature is available that directly compares the performance of multi-modal systems over their single-modal approaches. To highlight these aspects, this section will describe the findings from several studies conducted during last decade. A summary of key literature and results within the multi-modal biometrics domain is illustrated in Table 4.4.

Authors	Biometric Modalities	Level of fusion	Fusion Approaches	Performance	
				FRR (%)	FAR (%)
Kumar *et al* (2003)	Palmprint			4.5	2.0
	Hand Geometry			8.3	5.3
	Multi-Modal	Feature	Concatenation	5.1	2.3
		Match score	Max rule	1.4	0.0
Snelick *et al* (2003)	Face			24.7	0.1
	Fingerprint			17.0	0.1
	Multi-modal	Match score	Min-Max+Simple Sum	5.1	0.1
			Min-Max+Max Score	22.1	0.1
			Min-Max+Min Score	17.0	0.1
Ross & Govindarajan (2005)	Hand			15.0	0.1
	Face			35.0	0.1
	Multi-modal	Match score	Min-Max+Simple Sum	2.0	0.1
		Feature+ Match score	Simple Sum	5.0	0.1
Jain *et al* (2005)	Face			32.3	0.1
	Fingerprint			16.4	0.1
	Hand			53.2	0.1
	Multi-modal	Match score	Min-Max+Simple Sum	5.1	0.1
			Tanh+Simple Sum	5.6	0.1
			Z-score+Simple Sum	5.8	0.1
			MAD+Simple Sum	9.3	0.1
			Simple Sum	12.5	0.1
Koreman *et al* (2006)	Voice			3.2 (EER)	
	Face			27.6 (EER)	
	Signature			8.0 (EER)	
	Multi-modal	Match score	Min-Max +GMM	0.8 (EER)	
Kounoudes *et al* (2008)	Voice			4.1	4.2
	Fingerprint			11.4	9.9
	Hand Geometry			9.1	9.4
	Multi-modal	Decision	Majority voting	0.9	1.2

Table 4.4: A summary of literature and results on multi-modal biometrics

Authors	Biometric Modalities	Level of fusion	Fusion Approaches	Performance	
				FRR (%)	FAR (%)
Dozono & Nakakumi (2008)	Keystroke timing			0.1	0.0
	Writing speed			0.4	0.0
	Writing pressure			0.3	0.0
	Multi-modal	Feature	Concatenation	0.0	0.0
Soltane *et al* (2010)	Face			0.4 (EER)	
	Speech			0.0 (EER)	
	Multi-modal	Match score	Adaptive Bayesian	0.1 (EER)	

Table 4.4: A summary of literature and results on multi-modal biometrics (Cont.)

Research by Kumar *et al* (2003) attempted to improve the performance of palm print-based verification system by integrating hand geometry features. The experiment conducted on 1000 hand images, 10 samples from each user for 100 users. The first five images from each user were used for training and the rest were used for testing. Hand images of every user were used to automatically extract the palm print and hand geometry features. A total of 16 hand geometry features and 144 palm print features were obtained from every hand image. These features were then examined for their individual and combined performances. The feature-level approach using concatenation and match-level fusion approach using the max rule function were applied. The experimental results showed the decision level fusion scheme achieved better performance than the fusion at the feature level.

Research by Snelick *et al* (2003) evaluated the performance of a multimodal biometric system that used face and fingerprint classifiers. Normalization techniques such as min-max, z-score, MAD and Tanh were used to scale the score into the same interval. The normalise scores were then combined using different fusion methods including simple sum of scores, maximum score, minimum score, sum of posteriori probabilities (sum rule) and product of posteriori probabilities (product rule). The experiments conducted on a database of more than 1000 users showed that combining face and

fingerprint biometric classifiers reveal significant performance improvements over single biometric systems with the maximum improved FRR 20% at FAR 0.1%. The simple sum generally performs well over the range of normalization techniques. Indeed, the min-max normalization followed by the sum of score fusion method generally provided better recognition performance than other schemes with the best FRR 5.1% at FAR 0.1%. They also observed that the method chosen for fusion has a significant impact on the resulting performance as shown in Table 4.4.

Research by Ross and Govindarajan (2005) discussed the fusion of face and hand modalities where the 9-byte hand feature set and the LDA-coefficients of the grayscale face image were combined. This experiment was conducted based on the West Virginia University (WVU) dataset consisting of hand and face information pertaining to 100 users with each user providing 5 samples of each biometric. The study compared the classification performance of feature and matching levels. The experimental results showed that the matching performance of feature-level approach was seen to result in a marginally inferior performance compared to the match-level fusion approach. The researchers stated that fusion at feature level is difficult to achieve in practice because multiple modalities may have incompatible features (i.e. different features may have different dimensions and measurement).

Research by Jain *et al* (2005) studied the performance of a multi-modal biometric system that used face, hand-geometry and fingerprint modalities for user authentication. The experiment utilised the Michigan State University multi-modal database with a set of 100 users with finger, face and hand geometry samples. This study attempted to evaluate the performance of multi-modal approach against their uni-modal approach. The performance of the multi-modal biometric system has been

studied under different normalization and fusion techniques. The simple sum of scores, the max-score and the min-score fusion methods were applied on the normalized scores. The normalized scores were obtained by using one of the following techniques: simple distance-to similarity transformation with no change in scale, min-max normalization, z-score normalization, median-MAD normalization, doubles sigmoid normalization, tanh normalization and Parzen normalization. The research found that a multi-modal system employing the sum of score method provides better performance than the best single-modal system for all normalization techniques except median-MAD normalization. Among the various normalization techniques, the min-max and tanh normalization methods outperform other techniques at low FARs. The experimental results clearly showed that the normalisation approaches have a significant role to play in the performance of multi-modal approach as shown in Table 4.4. They concluded that min–max, *z*-score, and tanh normalization techniques followed by a simple sum of scores fusion method result in a superior GAR than all the other normalization and fusion techniques.

Research by Koreman *et al* (2006) described a multi-modal user authentication system on a PDA using non-intrusive biometrics. The experiment employed the multi-modal database with a set of 60 participants with voice, face and signature samples. All of these biometric data were recorded directly from PDA via microphone, camera and touch screen interface respectively. Several fusion techniques were tested for biometric evidence combination. The test results showed that fusion by the face concatenation of voice and face features led to substantially lower performance than voice verification alone. As signatures cannot usefully be time aligned with video recording therefore the match-level fusion approach was applied using Gaussians

Mixture Model (GMM) classifier together with a min-max normalisation technique. The tests showed that the combination of non-intrusive biometrics of voice, face and signature can achieve a high level of authorisation accuracy and should be acceptable to implement on mobile devices.

Kounoudes *et al* (2008) investigated the use of fingerprint, voice and palm geometry features of an individual for verification purposes. The researchers created a multi-modal biometric database containing samples from voice, palm and fingerprint for 30 users. For each user, each biometric characteristic was captured for five sessions, four of which were used for training and one for testing. The performances of single modality biometrics and multi-modal approach were evaluated. The result clearly showed that combining single modalities at decision level using majority voting technique achieved the best performance.

Dozono and Nakakuni (2008) examined the use of multi-modal behavioural biometrics sampled from keystroke timing and hand writing pattern to authenticate the user. The experiment utilised a database containing 11 users and each user provided 6 samples of “kirakira” words. The analysis of keystroke timings and pen calligraphy were evaluated by using Self Organizing Maps. By combining keystroke timing, pen speed and pressure information at feature-level showed improvement both in terms of FAR and FRR compared to the single modality techniques.

Soltane *et al* (2010) investigated the ability of multi-modal biometrics to authenticate the user by using an integration of facial and vocal modalities. A total of 30 subjected were used for the experiments. The face verification study was tested by using 16 face images from each user and 4 voice samples were used in the voice verification study. The multi-modal biometrics was evaluated by combining face and voice modalities at

the match score level fusion. By using Adaptive Bayesian fusion technique, the result showed that voice information can be used to disambiguate face information.

From all of aforementioned studies, it is clearly showed that combining more than one biometric modality together result in improved performance than using them alone. Fusion at the match-level is the most widely used however the performances of multi-modal system are varying depending on score normalization techniques and fusion strategies. An advantage of fusion at this stage is that existing and proprietary biometric systems are not affected, allowing for a common middleware layer to handle the multi-modal application but with a modicum of common information. Another advantage of using matching scores is that data from prior evaluations of single-mode biometric systems can be reused. This avoids live testing or re-running individual biometric algorithms.

In recent years, several researches have investigated the robustness to spoofing attacks of multi-modal biometric systems (Rodrigues *et al*, 2010; Akhtar *et al*, 2011; Biggio *et al*, 2012). The studies suggested that the combination of biometric modalities and the fusion methods should be taken into consideration because some integration could be less effective than others.

4.6 Conclusion

Three behavioural biometric techniques: linguistic profiling, keystroke dynamics and behaviour profiling have been identified as the most appropriate techniques that can provide transparent authentication - as users can be authenticated based on the most regularly used mobile device – typing communications. A significant number of prior studies have proved that these three biometric techniques can be used to classify a

user with a good degree of performance. There is also clear evidence that the performance of multi-modal systems significantly outperforms the single approaches. Therefore, the use of multi-modal biometric techniques based on linguistic profiling, keystroke dynamics and behaviour profiling represents an intriguing proposal given that these three approaches can provide authentication based upon any form of typed communications such as typing a phone number, texting (SMS), email and social networking. As such, the next chapter presents the results of a series of experiments that were conducted to examine the feasibility of utilising linguistic profiling, keystroke dynamics, behaviour profiling and multi-modal biometrics to verify users on mobile devices.

5 Feature Analysis of Linguistic Characteristics in Text Messages

This chapter presents a feasibility study of using linguistic profiling to authenticate a user. Whilst linguistic profiling can be applied to any text-based communication, such as email, twitter, social network communications, a focus has been given to the analysis of SMS texting messaging in the first instance due to its wide and popular usage. A key requirement in biometric design is the identification of potential features and this chapter presents a thorough examination, identification and preliminary testing of text-based features. The chapter presents a detailed statistical analysis of the acquired dataset, prior to examining and extracting possible features for further examination. The chapter subsequently performs a series of preliminary tests to determine whether the acquired features have a positive discriminative role to play.

5.1 Introduction

Given the increasing use of text-based applications such as SMS, emails and social networks on mobile handsets, linguistic profiling has the potential to enable authentication of users in a cost effective and efficient manner. If authentication of a user was possible during text-based communication then verification could be performed transparently and continuously throughout the duration of typing. Indeed, an examination of how users utilize mobile devices demonstrates that text-based input is a significant user activity with a growing range of data-based services with which the user can interact (e.g. social network postings, twitter, text messaging, email and internet searching). It is, however, anticipated that a user's linguistic profile will vary depending upon the application the user is interacting with – the composition and

nature of the language used in a twitter post can differ significantly from a text message, email or social network post. As such, this research restricted its examination of linguistic profiling to a single application scenario – text messaging. This was selected for three reasons:

1. The length and nature of text messaging fits largely in between other forms of text-based communication. It is larger than twitter (and many social network) posts but smaller than typical emails.
2. It was felt text messaging would permit the examination of linguistic profiling in an environment where users have modified language use. The use of abbreviations, emotions and "text-speak" are well documented and it was deemed an interesting research question to examine what role they would serve in identifying individuals (Rafi, 2009). The composition of emails tends to be more professional and in line with the norms of correct grammar and language.
3. Text messages represent a fair share of written communication with over 6 billion messages sent per day only in the US in 2011 (Grady, 2012). On average, text messaging users send or received 35 messages per day.

Compared with formal text documents such as student essays, SMS messages pose special challenges due to their special characteristics of size and composition. SMS messages are short in length, limited to 128 characters. However, many mobile systems are capable of sending and receiving multiple messages appended. Nevertheless, the medium does result in a shortening of normal communication as evidenced by widely utilized "text speak" (Doring, 2002). SMS messages are generally informal in style and the users had to find a way of being concise their text message

with in a limited length. The users may have to use varied spelling rather than using the standard spelling and often do not follow syntactic and/or standard grammatical rules. Therefore, the analytical techniques that are successful in previous works may not be applicable in the context of text messaging. However, certain characteristics such as unusual spelling usage (e.g. “you” text as “u”), patterns of vocabulary usage and structural layout could be useful features for discrimination between users. Ledger and Merriam (1994), however, highlight that authorship analysis would not be significant for texts containing less than 500 words, or approximately 2500 letters.

5.2 Dataset

The purpose of this investigation is to determine the feasibility of identifying a user based on SMS messages. Given the importance the dataset has in ensuring appropriate features are identified, the research sought to examine two datasets. The first is a publicly available SMS corpus collected for research at the Department of Computer Science at the National University of Singapore and the second was conducted using a SMS corpus collected by the author at Plymouth University.

5.2.1 The National University of Singapore SMS Corpus (NUS Corpus)

The first experiment employed a publicly available SMS corpus provided by the National University of Singapore (How and Lee, 2004). These messages were collected from two sources. The first source was a large group of 166 undergraduate students who together contributed 9515 messages to the corpus. The majority of these students were Singaporeans aged between 18 and 22. These students were aware that their messages would be made public. In order to collect the messages for the corpus, the students were asked to upload up to 75 messages to a website for which they

would receive nominal compensation. They were asked to only submit conversational English messages and were also asked not to upload repetitive messages. The second source was collected from Yahoo's SMS chat website which broadcasts live SMS chats of certain SMS chat rooms. Form this source; an additional 602 messages were collected from an estimated 30 people. Repetitive messages were removed by the collectors. To this end, a total 10117 SMS messages were collected with user ID number tagged. The messages tagged with user ID number could be grouped into 133 users. Each user had a different number of messages. In order to maximise the number of users and SMS messages to ensure that there were sufficient samples to use, a total of 26 users who had more than 60 messages were selected. For each user, a subset of 60 SMS messages was randomly chosen. Utilising a fixed-size dataset assisted in simplifying the code necessary to undertake the analysis and provided a consistent methodology. Table 5.1 demonstrates the final dataset for utilising in the analysis.

Number of participants	26
Number of SMS text messages	1560
Number of messages per user	60
The maximum length of text messages	200 characters
Average length of messages per user	12 words
Average length of messages	56 characters

Table 5.1: Description of the final dataset of NUS SMS corpus

Since the majority of participants are Singaporean, the "Singlish" language has been used in SMS text messages. Singlish is the English-based creole spoken and written colloquially in Singapore and has its own unique slang and syntax, which are more pronounced in informal speech (AussiePete, 2008). The vocabulary of Singlish consists of words originating from English, Malay, and Hokkien (the native Chinese dialect). A list of widely used Singlish vocabularies in this NUS SMS corpus is "Lor", "Leh", "Mah"

and “Liao”. Examples of SMS messages are shown in Table 5.2 with a full extract of all messages located in Appendix A)

SMS text messages
Of course got use lah. U must jia you k. 2 more days and u can enjoy liao oh. :>
I'm still thkin... I'll consider after all e talks lor, c which one more interestin...
Huh!? Aiyah i tot can earn some money... nw no money liao. Nvm then thanks anyway.
Oki... Tink i'm confused... I only know that it's at paragon... Dunno where you referring to...
Hi. They not going tonite cos they have a morning class tml. Think i wont be going either. Have fun.
which part of town would you be in ?
It is not ur fault, no need 2 zi ze lor. Ming zhi dao ni mei you cuo hai ying yao wo yuan liang. I will see.
Hey pinky... E flowers frm ur mum huh, very nice leh...
Thanx a lot 4 ur help!
Hi. We are going to play pool. May not be home so soon.

Table 5.2: Examples of SMS text messages from NUS corpus

5.2.2 The Plymouth University SMS Corpus (PU Corpus)

Based upon the requirement to keep the number of samples to a minimum, a total of 30 participants were required to send at least 16 messages to each other using a non-predictive text input method. Participants were encouraged to discuss any subject of their choice during the experiment in order maximise the potential level of detail and complexity within each transmitted message and to increase the likeliness of users adopting their own unique texting style and linguistic composition. A real time SMS simulation program was developed to collect the SMS input data, a screenshot of which is illustrated in Figure 5.1 A copy of the source code can be located in Appendix B.

(a) An XDA IIs Smartphone (b) Screenshot of user interface of experimental software

Figure 5.1: Screenshots of application deployed on mobile devices

A program was deployed onto two Smartphones (XDA II and SPV M2000). The program enabled text based communication between pairs of participants using a Bluetooth communication channel. Bluetooth was used to simulate a GSM based SMS service in order to reduce cost. The maximum length of the SMS text message is limited to 128 characters. The messages were saved in a text file prior to transmission. Table 5.3 demonstrates the final dataset for utilizing in this experiment.

Number of participants	30
Number of SMS text messages	487
Average number of messages per user	16
The maximum length of text messages	128 characters
Average length of messages per user	17 words
Average length of messages	83 characters

Table 5.3: Description of the final dataset of PU SMS corpus

The majority of participants were students at Plymouth University and the use of English language was stipulated. Examples of their messages are shown in Table 5.4 (and a full listing of all messages can be located in Appendix A).

SMS text messages
hello how are you tonight
LOL oook, aall is clear now. well, relatively. Is that what we''re having at cake time then? And I''ve decided this office
hey im here now. i cant see you so im going to give you a call bye
ahaaaa great indeed.i am happy to try :D u stay in university?
good. I stay off campus but just 10 mins walk 2 d uni
actually i''m abit tipsy now. dont mind what i said earlier. hahaa. never knew you have a maid. :p
that good so u work hard and also study at the same time!!take care ur self :) I love money but cant find a job!!!
yeah actually i''m not the biggest fan of running either haha x x x
no I actually had breakfast and u?
its new...just discovered last year...hehehe

Table 5.4: Examples of SMS text messages from PU corpus

5.3 Methodology

The preliminary feature analysis utilised a descriptive statistic approach to analyze the raw information presented by the datasets. The statistical analysis has the ability to determine potential positive features for forming unique patterns to discriminate individual users (Jain *et al*, 2000). Furthermore, it is a critical process to select effective features for subsequent use in pattern classification because the performance of the system is closely related to the feature selection. Furthermore, the aim of this task was to identify an optimum set of features since a large number of features will increase the complexity and the size of classifier resulting in problem known as the *curse of dimensionality* (Fu *et al*, 1970). The second aspect to this analysis was to conduct a preliminary experiment (fixing the classifier and varying the features) to determine the positive linguistic profiling features.

In this study, the individual user word profiling and stylometric features were proposed for use in authenticating users. In order to create individual user word profiling, special keywords such as abbreviation and emotion based words that a user uses in their message were selected as these special words may provide some useful insight into

the identity of the author. Although more than 1000 features including lexical, syntactic, structural, content-specific, and idiosyncratic characteristics have been evaluated and compared in various studies. This research focused on features that cover a wide range of linguistic levels and are easy to implement. The following features: lexical, syntactic, and structural features were integrated into a stylometric feature set for discriminating between users. As a result, four different types of features were employed and a brief description of each type of these features is given below.

- User's word profiling features: are collections of abbreviation and emotional words that are frequently used for each user.
- Lexical features: can be divided into character-based or word-based features. In this research, character-based lexical features used in De Vel *et al* (2001) and word-length frequency features used in Mendenhall (1887) and De Vel *et al* (2001) are investigated. In total, 33 lexical features were adopted into the key feature set.
- Syntactic features: are used to measure an author's writing style at the sentence level. Since SMS language tends to use syntactic short forms (Rafi, 2009), these features may be useful for discrimination. In total, 23 features were integrated into key feature set.
- Structural features: are used to measure the overall appearance and layout of the messages. In total, 8 features used in De Vel (2000) were considered in this study.

As a result, a total of 64 stylometric features were employed in the experiment. These are listed in Table 5.5.

Features		Example
Word-profiling features		
	Total number of abbreviation and emotional words in user's word profiling used	
Lexical Features		
Character-based features		
	Total number of characters (C)	a-z, A-Z, 0-9
	Total number of alphabetic characters	a-z, A-Z
	Total number of alphabetic characters/C	
	Total number of digit characters	0,1, 2, 3, 4, 5, 6, 7, 8, 9
	Total number of digit characters/C	
	Total number of symbol characters	
	Total number of symbol characters/C	
	Total number of capital characters	A-Z
	Total number of capital characters/C	
Word-based features		
	Average Word Length (number of characters)	
	Word length frequency distribution (11 features)	1,2,3,4,5,6,7,8,9,10,long word
	Word length frequency distribution/M (12 features)	
Syntactic Features		
	Frequency of punctuations (20 features)	".", "!", "?", "!!", "??", "…", ",", ":", "☺", "-","/", "*", "**", "+", "++", "(", ")", "$", "£"
	Total number of space after punctuation	".[Space]"
	Total number of punctuation after space	"[Space]."
	Total number of no space after punctuation	
Structural Features		
	Average Sentence Length	
	Message Length	
	Total number of sentences	
	Message Length/ Total number of sentences	
	Total number of words in message (M)	
	Total number of words/ Total number of sentences	
	Total number of space characters	
	Total number of space characters/C	

M = total number of words, C = total number of characters

Table 5.5: Summary of stylometric features

Before being analysed, all SMS text messages were passed through a program to extract all of the features. Manual feature extraction is too labour intensive and time consuming; an automated feature extraction program was developed using MATLAB developed by MathWorks. This is a common and well-accepted tool used throughout scientific and engineering communities in the analysis of mathematical problems. Indeed, MATLAB was utilised extensively for the modelling and evaluation of this study.

5.4 Descriptive Statistics

Descriptive statistics is an approach that enables the biometric-designer to visualise the nature and complexity of the data so that an appropriate set of features and classifier can be selected. Unfortunately, this is not a definitive science, with a set problem leading to a definitive set of features or type of classifier, but is a process subject to trial and iterations (Jain *et al*, 2000). Descriptive statistics aids the designer into selecting suitable features and classifiers to test.

5.4.1 NUS Feature Analysis

5.4.1.1 Words profiling features

Abbreviation and emotional words are commonly used in text messages. Abbreviation, which commonly uses single letters, numbers or combination to represent the whole word (for example, “See You” can be texted as “CU”, “Mate” can be texted as “M8”). Emotion or verbal effect represents body language such as “:)” for “smiley face”, “hehe” or “haha” for laughter. Therefore, the frequency distributions of each abbreviation, emotional words were used to create user profiles. By manually observing and analysing historical messages on NUS SMS database, a total of 317 abbreviation and emotional words were identified. Through performing a total count of each word used in messages, it is possible to estimate the similarity of the quantitative information between users.

Whilst theoretically it is possible to identify users based upon abbreviation and emotional words if the user used abbreviation and emotional words in every single message and no two users used the same word, this is unlikely to occur frequently. For illustration purposes, the figure below shows the total count of abbreviation and emotional words used in text messages for each user.

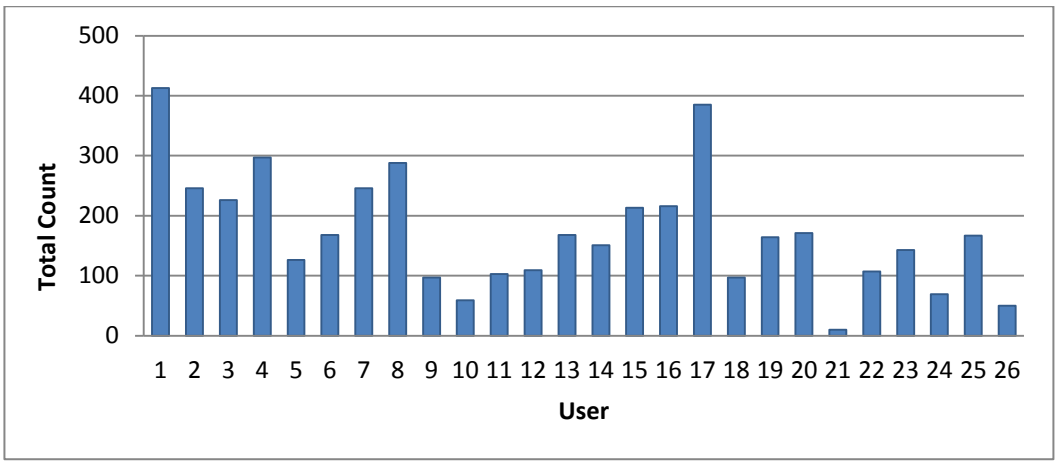
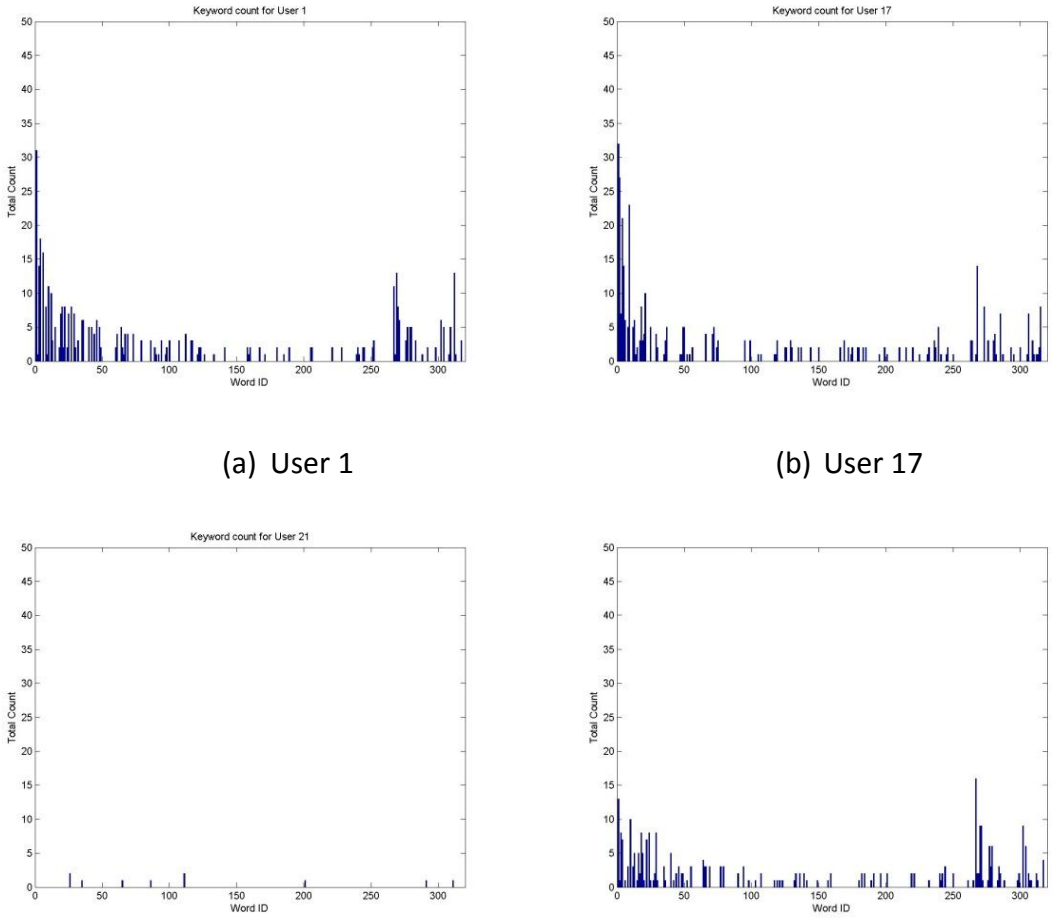


Figure 5.2: Total abbreviation and emotional word count for each user

The analysis of possible features at this stage needs to focus upon two particular biometric characteristics; their ability to be universal and unique. With respect to universal (i.e. the feature appears throughout the user population) it can be seen in Figure 5.2 on average, each user has more than 160 abbreviation counts across their 60 messages which suggest that abbreviation words have been used multiple times in a message. The majority of users (88%) used abbreviation words in a single message. If they used different abbreviation and emotional words they could be discriminated between each other. In order to examine the uniqueness of abbreviation and emotional word usage between users, the total count of each abbreviation and emotional word occurrence for an individual user were analysed (and illustrated in Figure 5.3).



(a) User 1 (b) User 17

(c) User 21 (d) User 8

Figure 5.3: Abbreviation and emotional word usages for individual user

Figure 5.3 demonstrates that users use different combinations of abbreviation and emotional words in their messages. In general each user used 65 words in their message profile; however, they shared 62 words with another user. For the worst case, *User 1* and *User 8* have shared 65 words in their SMS messages but they used these words (in term of quantity) very different (as illustrated in Figure 5.3 (a) and (d)). There are also some users (e.g. *User 21*) that rarely use abbreviation and emotional words in their messages as also shown in Figure 5.3 (c). However, the lack of using abbreviation and emotional words could in itself be a potential feature to discriminate the user (i.e. the unique feature being the lack of use).

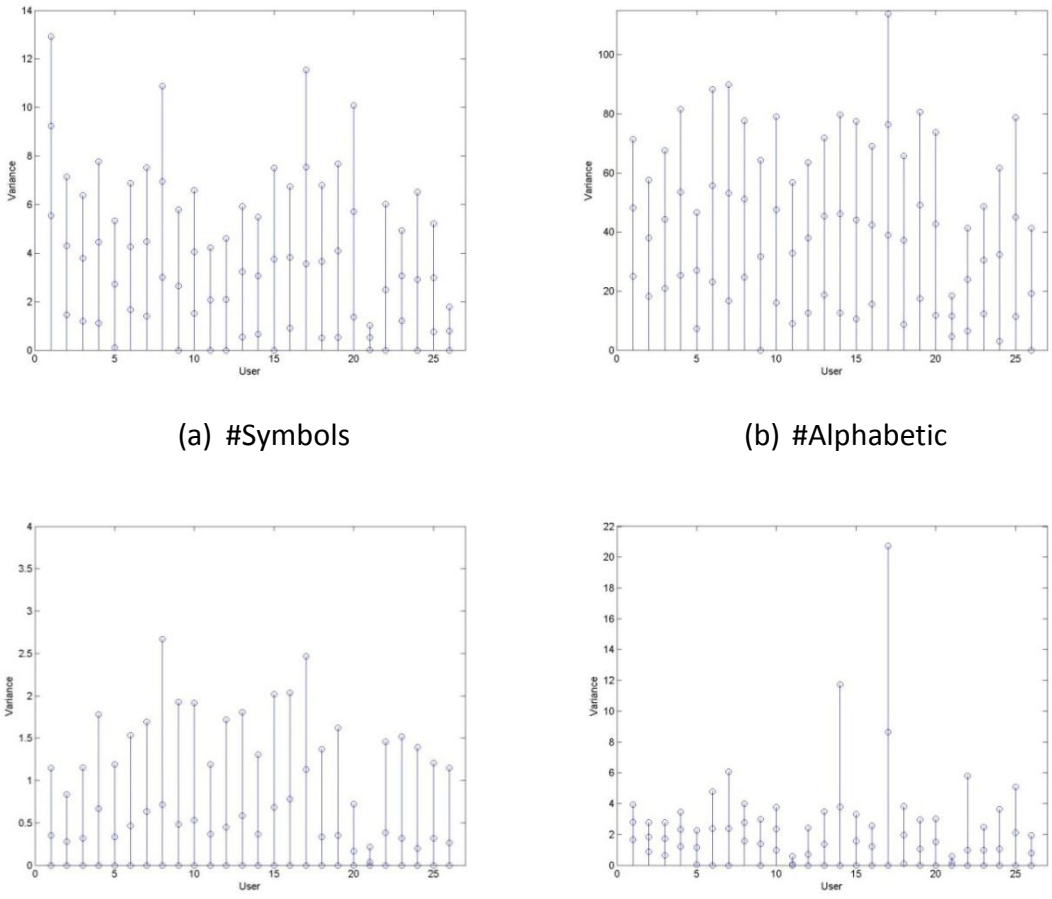
Further investigation observed that abbreviation and emotional words rarely appeared in every message, with users using different words in each message. This is expected as it is improbable that a user would compose a message with the same set of abbreviations and emotions. Furthermore, the majority of abbreviation and emotional words (72 %) are used only a single time in a message. This suggests that whilst there appears to be some value in abbreviations and emotions, their exclusive use in isolation is unlikely to result in an overly favourably outcome.

5.4.1.2 Lexical Features

In order to examine the possibility of discrimination based on lexical features, descriptive statistical analysis on both character-based and word-based features was conducted. For the character-based features, the total number of characters, alphabet only, digits, capitals and symbols character features were examined. For the word-based features, average word length in a message and length of word usage distributions were analysed. By performing mean and standard deviation examinations it is possible to establish the degree to which input data is similar or dissimilar between the users.

Figure 5.4 shows a mean and standard deviation plot of lexical character based features. The figure presents each users mean value and also the variance of lexical character based features by calculating the standard deviation, providing an estimate as to the similarity between the users input vectors. The classification process will be much more complex if the latency vectors observed from a single user incorporate a fairly large spread of variance, which suggests the samples do not exist on a clearly definable discriminative region. Indeed, from the users input data, two types of variance can be extracted: inter-class and intra-class variance. It is hypothesised that it

would be easy to classify a user if the intra-class variance ideally was zero so that every sample a user input would be identical and the inter-class would be as large as possible in order to widen the boundaries between users.



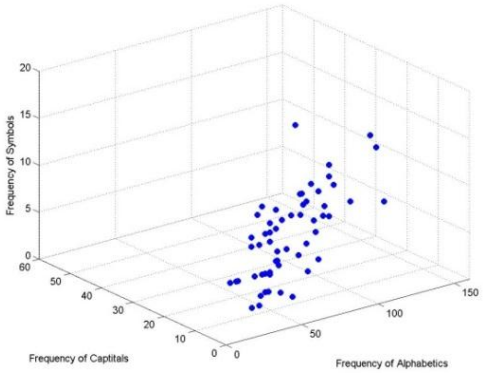
(a) #Symbols (b) #Alphabetic

(c) #Digits (d) #Capitals

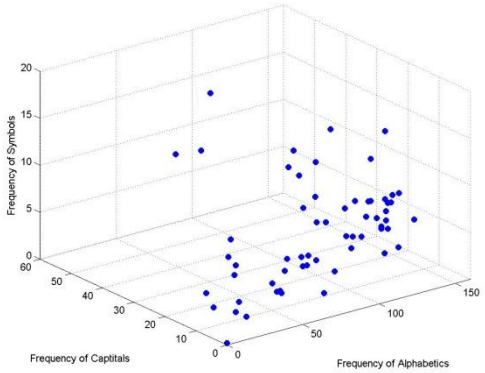
Figure 5.4: Mean & Standard deviation plot for lexical character based feature

Although an initial analysis of the intra-class variance for all features indicates that they are not ideal as no users had standard deviation close to zero for all features; however some users obviously have smaller intra-class variances than others. Furthermore, for each feature the majority of users have latency spreads that coincide with a number of others demonstrating that they have very low inter-class variance. This will make classification more difficult as users input vectors are more likely to be similar or within similar boundaries as other users.

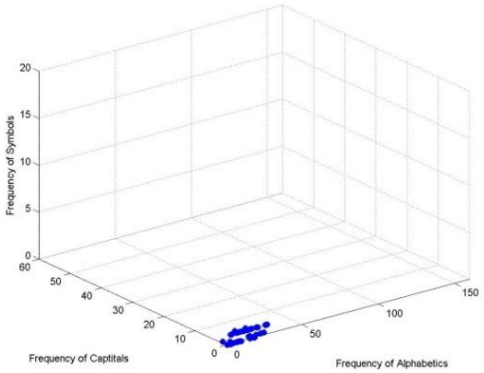
However, analysing features in a single dimension does not convey the uniqueness that can be obtained through combining features and thus moving the discrimination into multi-dimensional space. From an illustrative perspective, it is only possible to present three dimensions. Figure 5.5 illustrates a plot of an individual users' lexical character based feature data utilising the alphabetic count, digit count and symbol count values in each input vector.



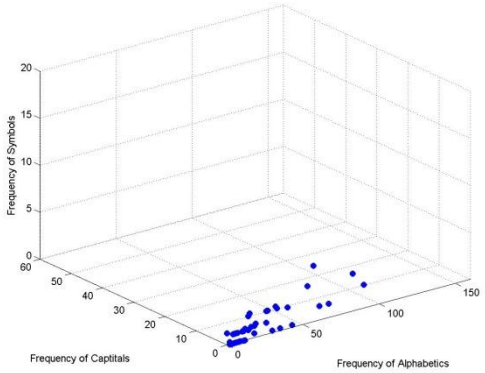
(a) User 1



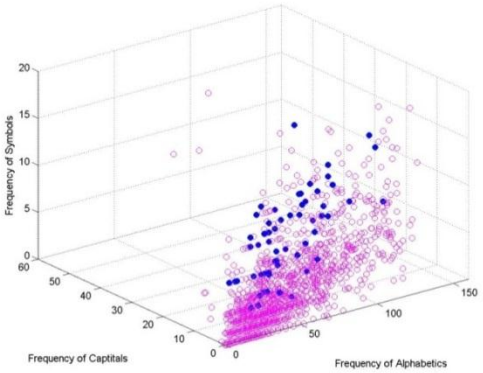
(b) User 17



(c) User 21



(d) User 26



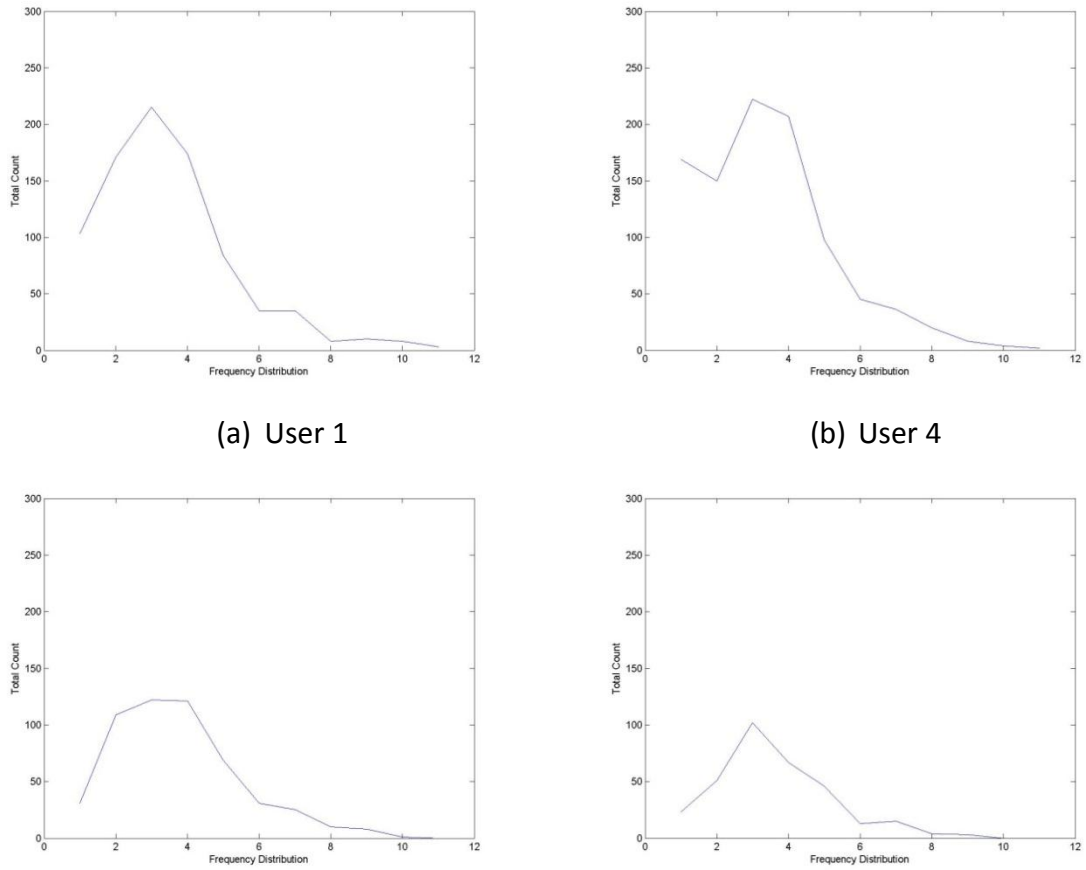
(e) All users

Figure 5.5: 3D plot of lexical character based feature

Figure 5.5 (e) shows a fairy complex problem with respect to deriving efficient decision boundaries. The correct classification of users in a multi-dimensional space is no simple task. However, there is some level of discriminative ability for some users. For example, it is possible for *User 1* and *User 21* (see Figure 5.5 (a) and (c)) to be discriminated from

each other as the graph clearly shows that the area plot of data do not coincide between each other. On the other hand, it can be difficult to classify *User 1* from *User 17* (see Figure 5.5 (a) and (b)) or *User 21* from *User 26* (see Figure 5.5 (c) and (d)) as they have the same area of data.

To examine the way in which users use short or long words in messages, the distributions of word length usage for each user were investigated. Figure 5.6 illustrates examples of the word length distribution usage for an individual user.



(a) User 1 (b) User 4

(c) User 24 (d) User 26

Figure 5.6: Word length usages for individual user

An initial analysis of word length feature indicates that the majority of users used words that were an average of 3-4 characters long in their messages. As can be seen in Figure 5.6, all users tend to use short words in their messages. This is expected, as

users had to find a way of being concise in their text message with in a limited length. Further analysis was investigated to examine similarity between users; Figure 5.7 shows a mean and standard deviation plot of lexical word based features.

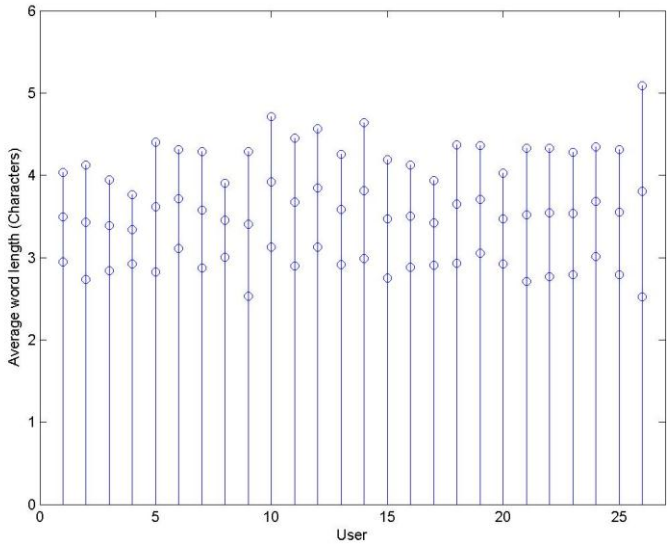


Figure 5.7: Mean & Standard deviation plot for average word length feature

Figure 5.7 demonstrates that users tend to share similar mean and standard deviation plot between each other, indicating that input vectors are more likely to be similar between users. Therefore, this will make classification more difficult. As a result, it is not expected that this measure would provide a robust feature.

5.4.1.3 Syntactic Features

In order to examine the universe of punctuation features the quantity of punctuation usage was counted and shown in Figure 5.8.

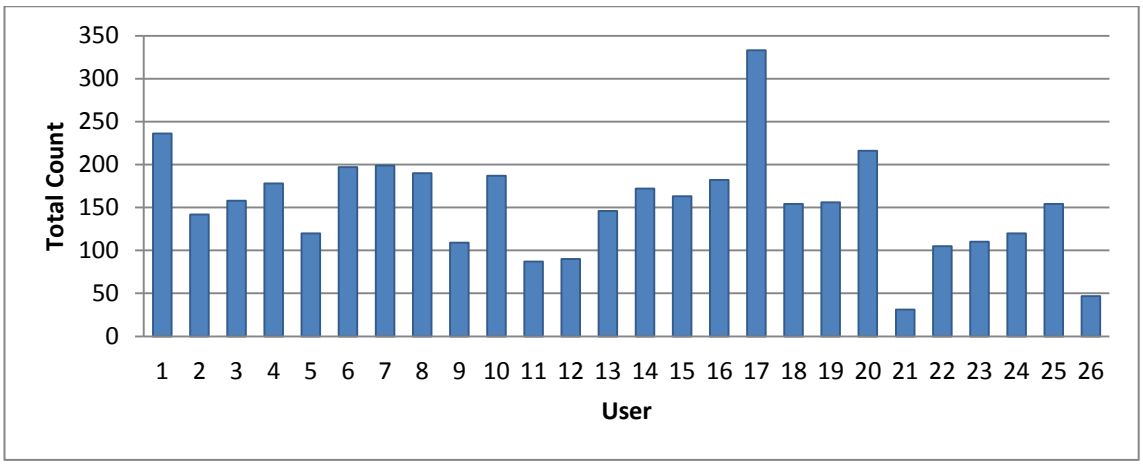
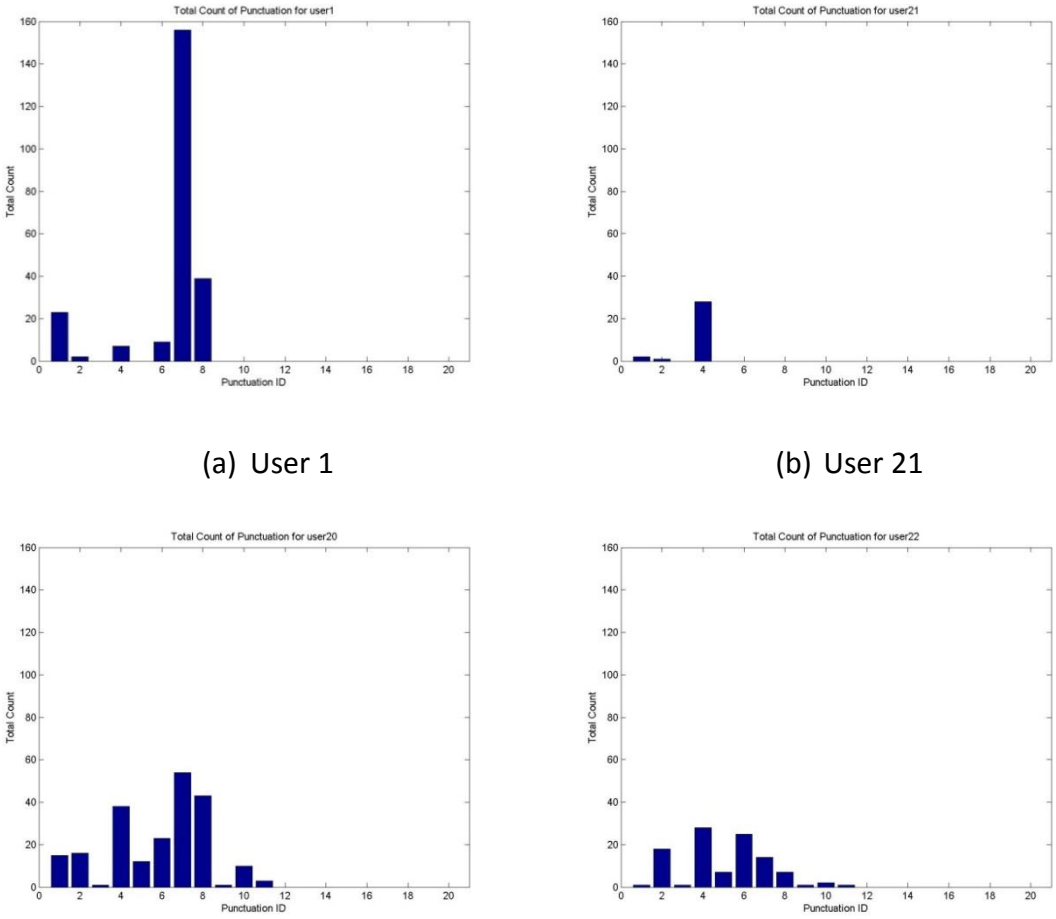


Figure 5.8: Total punctuation mark count for each user

Figure 5.8 demonstrates users used punctuation marks multiple times in a message. The majority of users (73%) used punctuation in a single message at least two times. This indicates that punctuation marks are widely used so that it is possible to be used as a classification feature. In general, each user used 8 types of punctuation mark and they were shared with others user. However, the user could be classified from each other if the quantity they used is different. Further in-depth analysis was conducted to find the similarity of punctuation mark usage between users.



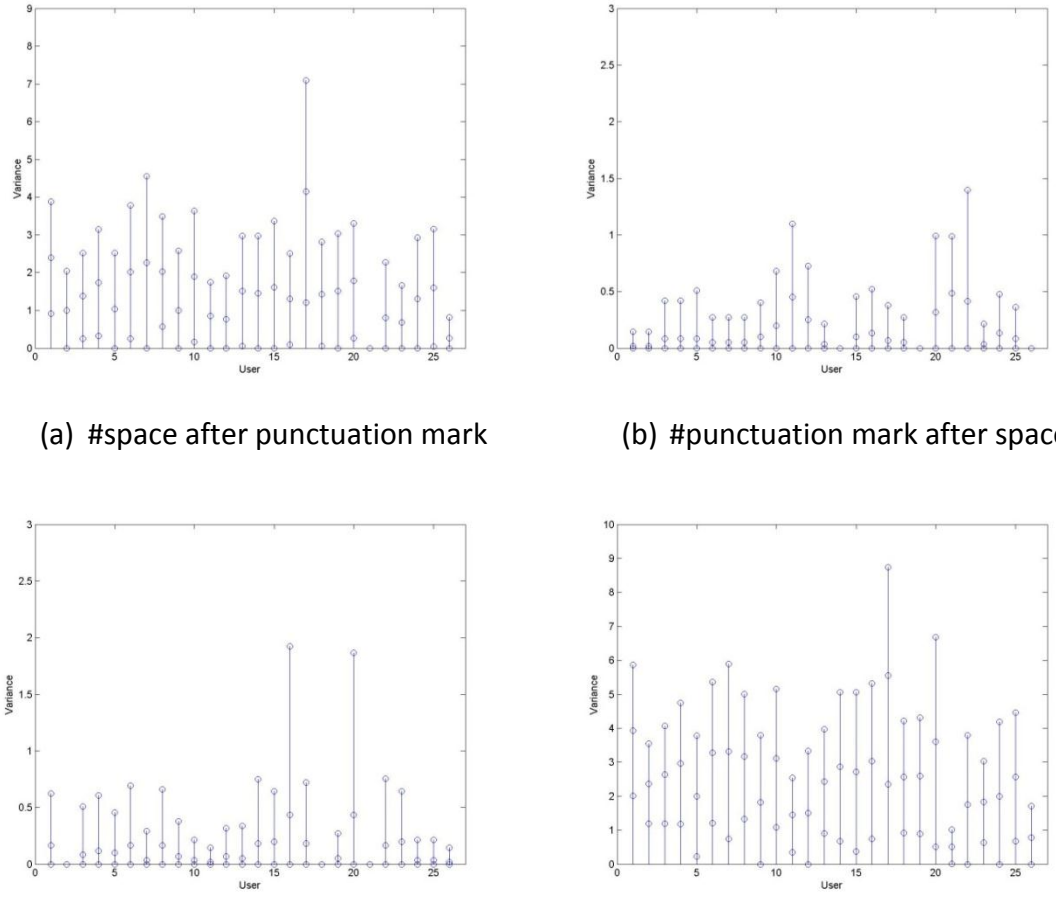
(a) User 1 (b) User 21

(c) User 20 (d) User 22

Figure 5.9: Punctuation mark usage for each user

An initial analysis based on individual punctuation usage indicates that users used similar common punctuation such as full stop, ellipse, question mark and comma in messages. Although users used a similar set of common punctuations their usage in terms of quantity is different. For the worst case, *User 20* and *User 22* used the same 11 punctuation marks in their message profile however; their usage in terms of quantity is different, for example, *User 20* used punctuation ID 6, 7, 8 in total 23, 54 and 13 times while *User 21* using punctuation ID 6, 7, 8 in total 25, 14 and 7 times as shown in Figure 5.9 (c) and (d). Therefore it is possible to separate their usage of these punctuations from their historical usage. However, it might be difficult to classify *User 21* as their punctuation mark usage rarely appeared in messages.

As can be seen in Figure 5.10, although the intra-class variances of syntactic features are not ideal there are some users that can be observed to have smaller intra-class variances than others. However, there are some users that have a latency spread that does not coincide with other users, demonstrating that they have very high inter-class variance. As a result, this may make classification easy as users input vectors clearly exist within different classification boundaries.



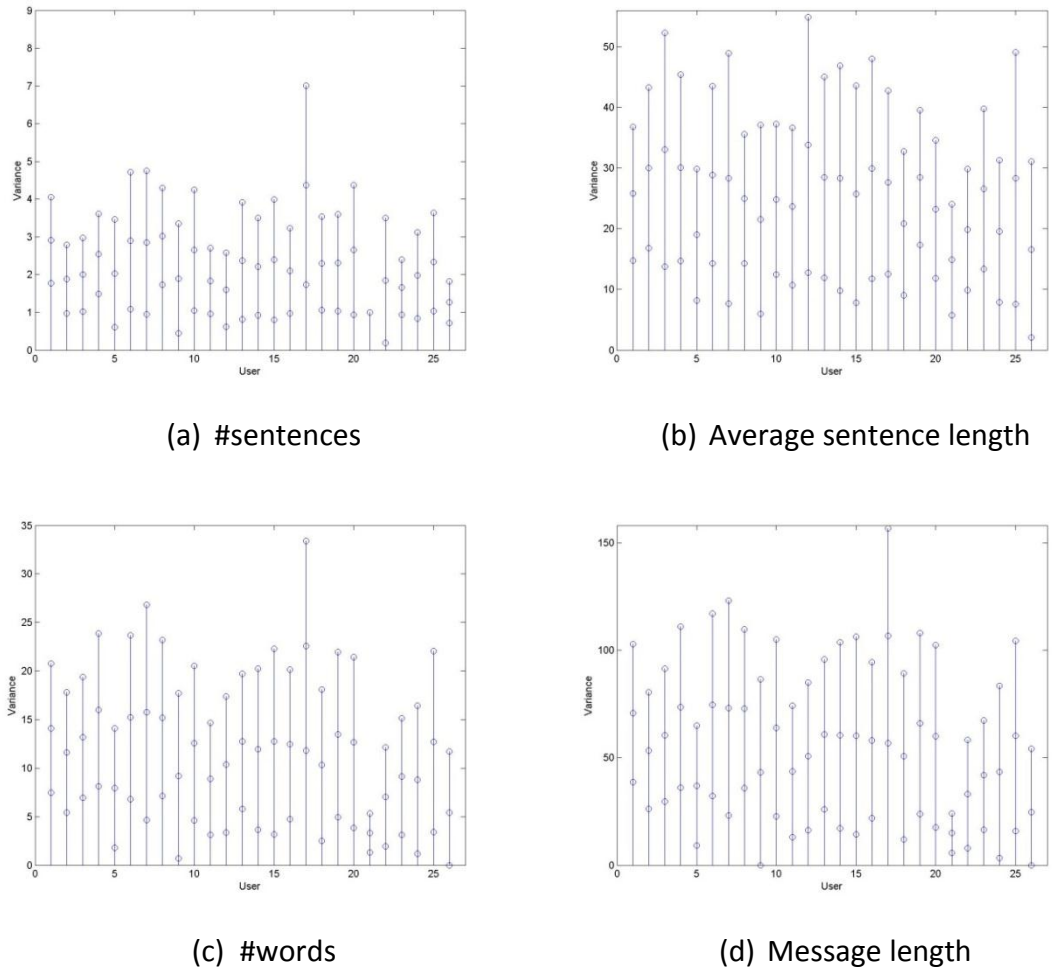
(a) #space after punctuation mark (b) #punctuation mark after space

(c) #no space after punctuation mark (d) Total #punctuation mark

Figure 5.10: Mean & Standard deviation plot for syntactic feature

5.4.1.4 Structural Features

In order to examine the similarity based on the way a user organises the layout of messages between users, inter and intra class variance for each user was calculated. Figure 5.11 shows a mean and standard deviation plot of structural features.



(a) #sentences (b) Average sentence length

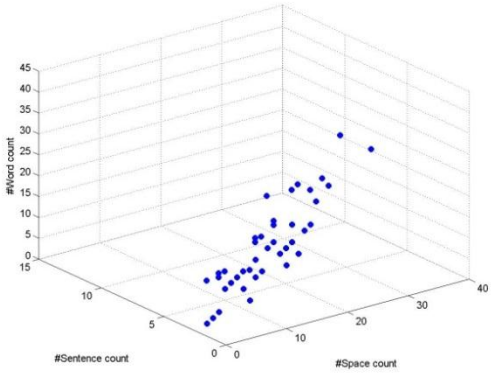
(c) #words (d) Message length

Figure 5.11: Mean & Standard deviation plot for structural feature

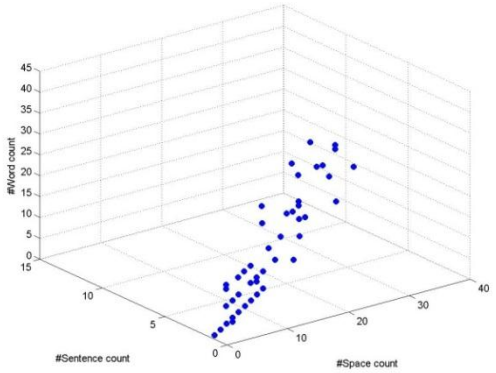
An initial analysis on structure features demonstrates that individual users have a fairly large intra class variance and also there are a number of users that have latency spreads that coincide with other users – demonstrating that they have very low inter-class variance. As a result, this may make classification more difficult as users input vectors are more within similar boundaries as other users.

Further analysis was carried out on a plot of an individual users' structural feature data utilising the space count, sentence count and word count values in each input vector in Figure 5.12. Figure 5.12 (e) shows a fairly complex classification space demonstrating that it would be difficult to classify all users based on structural features. However, there is some level of discriminative ability for some users. As can be seen from

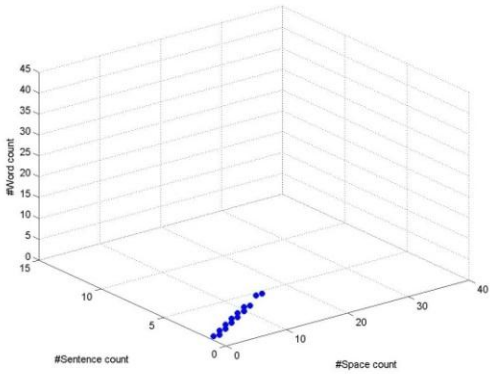
Figure 5.12 (c) *User 21* has a very small area plot of data compared to *User 17* (see Figure 5.12 (d)) which has a very wide spread of data. Therefore, it is possible to discriminate these users from each other. On the other hand, it could be difficult to classify *User 1* and *User 10* (see Figure 5.12 (a) and (b)) due to the area plot of data coincide between each other.



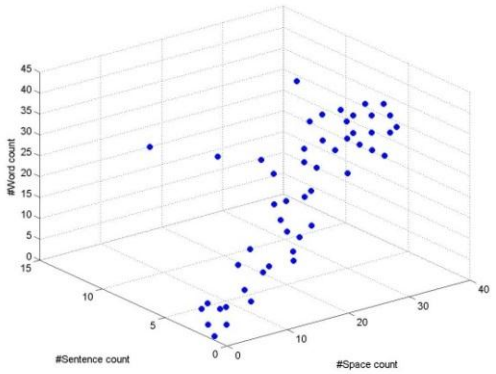
(a) User 1



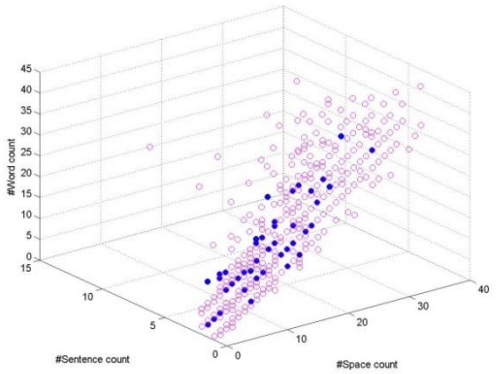
(b) User 10



(c) User 21



(d) User 17



(e) All users

Figure 5.12: 3D plot of structure features

From the above analysis, it demonstrates that a number of features have the potential to be used for discriminating users. For word profiling features, a user could be classified via abbreviation and emotional words because each user uses different set of

words in their message profile. Unfortunately, the classification process can be difficult as the quantities of each word usage are very low and fairly similar between users. Notably from analysis based on word usage by an individual user indicates that the user used only subsets of abbreviation and emotional word features. Therefore, it is possible to create an individual word profiling based on their historical messages. This can provide a degree of uniqueness between users. By counting the frequency of words in a profile that appears in a message and represent it as a dynamic word profiling feature could provide strong discriminative information. However, using only a dynamic word profiling feature to classify user might be difficult because it is possible for some users to have similar usage quantities. As a result using combination of this feature and others strong discriminative features could provide successful classification.

When analysing lexical features, all of the features other than the digit feature could contribute positive information toward the classification process for some users. However, it could be difficult to discriminate users from each other based on lexical word-based features as all users tend to use short words in their message. For syntactic features, some of the common punctuation features and features based on the way the user used punctuation in messages could provide additional discriminative information for distinguishing users. Although structural features show a fairly complex classification space however, sentence count, word count, space count, average sentence length and message length have some level of discriminative ability to be used to classify users.

Although, the analysis shows a strong indication that a number of features could be very useful for user classification, there are some features that contain no

discriminative information. Considering individual features from the above analysis, the discriminative power varies between users. For example, as can be seen from Figure 5.9 (a), punctuation 7 shows a very strong discriminative component for *User 1* as the quantity of usage is very high in comparison to other users. In contrast, this feature might not be as useful for *User 19* as the quantity of usage is similar to other users. By using all linguistic features for everyone (or a traditional static feature approach), the input vector may contain no discriminative information to discriminate a user and result in poor classification performance. Therefore, the data suggests a dynamic-based feature approach should be considered in this problem. It is expected that the classification performance will be improved by using a dynamic feature approach because only the selected useful features are used as part of the input vector but are dependent upon an individual analysis rather than a population. Furthermore, as the number of features is reduced, the classification process may require less complexity and result in less time consuming computation.

5.4.2 PU Feature Analysis

Due to the writing style of English language differs from the Singlish language on some important aspects (e.g. grammar, syntax). Therefore, this study investigated the writing-style features in SMS messages from the PU corpus in an attempt to examine the feasibility and reliability of linguistic profiling in a multiple-language context. The results and discussion of feature analysis will be presented in this section.

5.4.2.1 Keyword profiling feature

By manually observing and analysing historical messages on the PU SMS database, a total of 133 abbreviations and emotional words, were identified. In order to examine

the universal characteristic of features, the frequencies of words occurrence in messages were counted and are illustrated in Figure 5.13.

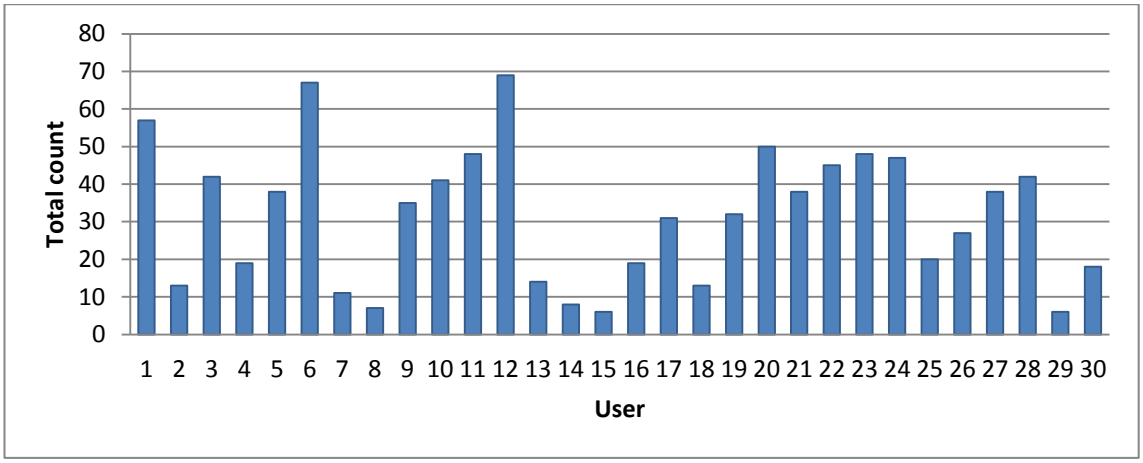
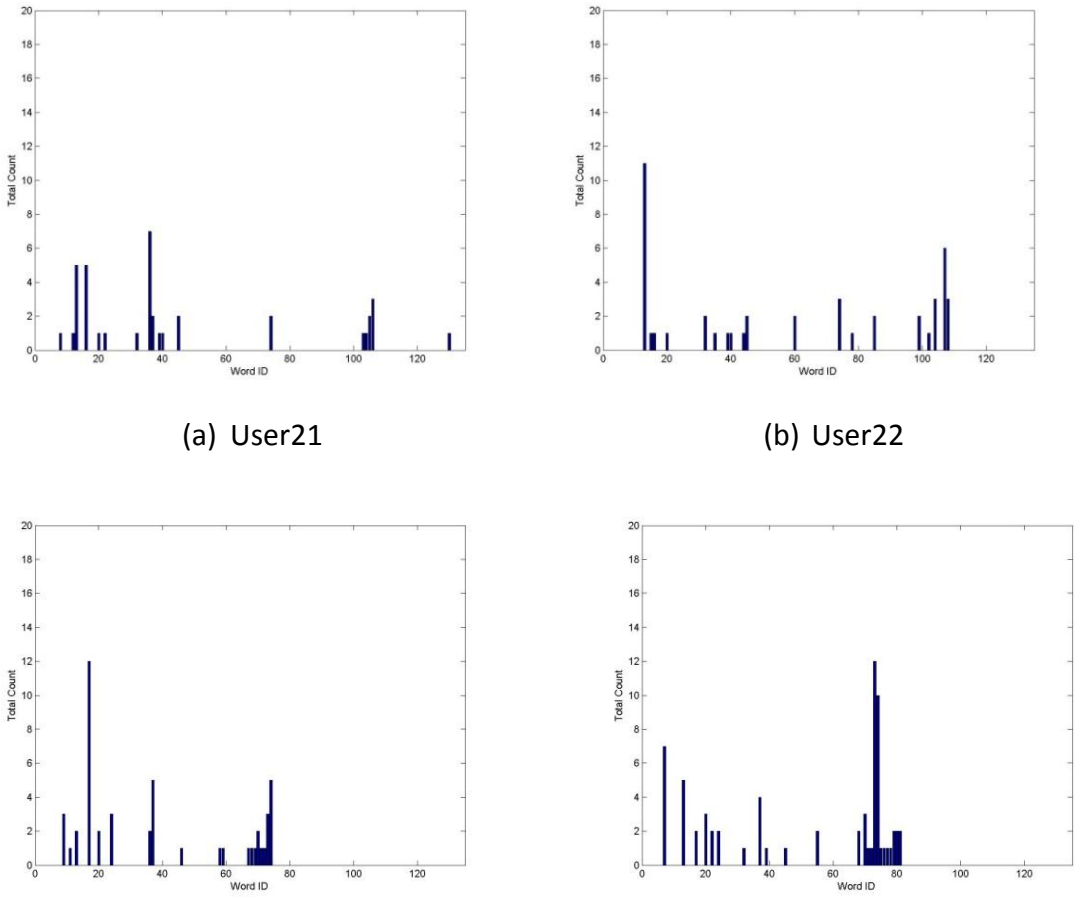


Figure 5.13: Total abbreviation and emotional word count for each user

Figure 5.13 illustrates that abbreviation and emotional words have been used in SMS messages by all users. On average, each user used abbreviation and emotional word multiple times in a message. This indicates that it is possible to use these abbreviation and emotional words as discriminative features. However, a user could be discriminated from each other if they use different abbreviation words. In order to examine the uniqueness of abbreviation and emotional word usage between users, the total count of each abbreviation and emotional word occurrence for an individual user were analysed and is shown in Figure 5.14.



(a) User21 (b) User22

(c) User11 (d) User12

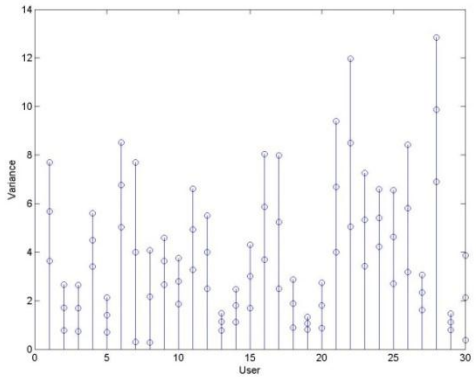
Figure 5.14: Abbreviation and emotional word usages for individual user

The finding showed similar results with previous studies demonstrating that users use different set of abbreviation and emotional words in their messages and share a number or words with other users as shown in Figure 5.14. However, the quantities of shared word usage are different between the users. Therefore, it is possible for these features to be used to classify users. Unfortunately, the majority of words (77%) appeared only a single time in a message so that the classification process could be difficult as the input vector contained only values 0 and 1.

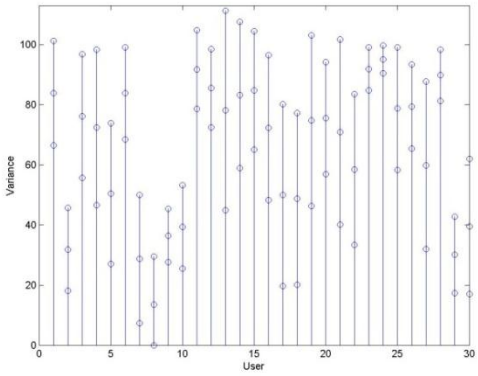
5.4.2.2 Lexical feature

In order to examine the similarity of input data between the users, inter-class and intra-class variances performed by mean and standard deviation were calculated.

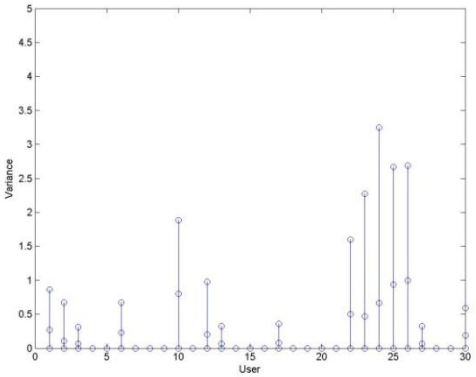
Figure 5.15 shows the relationships with each user's mean and standard deviations plotted for lexical character-based features. It can be seen that, many users share a similar mean and standard deviation plot, e.g. *User 2* and *User 3* with #symbol feature and *User 16* and *User 19* with #alphabet feature thereby giving rise to poor classification performance due to users input vectors are similar or within similar boundaries as other users. However, this is not necessarily the case as the similarity in mean and variance does not continue throughout all features and since the input vector to classification process is constructed from a number of features it is hoped this should provide the sufficient disparities required in input data for a classifier to discriminate against users correctly.



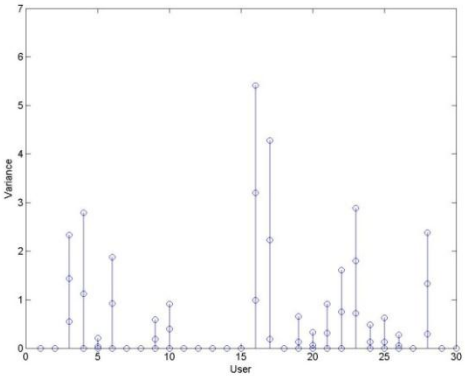
(a) #symbols



(b) #alphabets

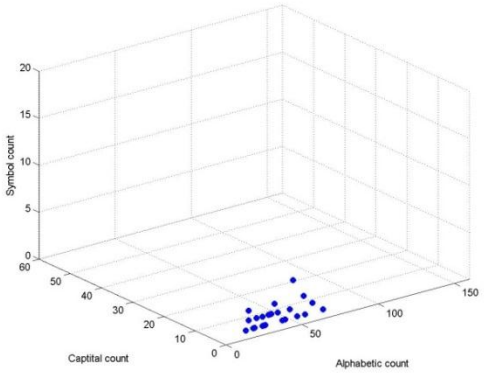


(c) #digits

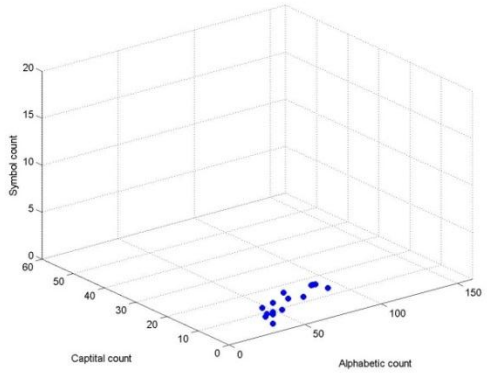


(d) #capitals

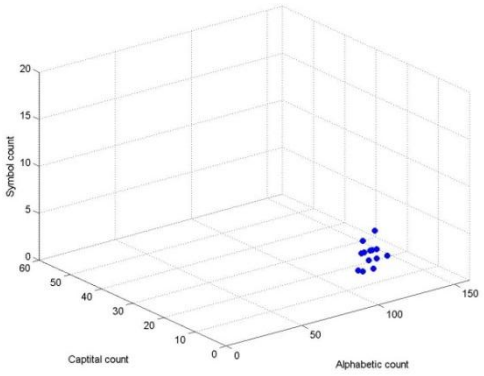
Figure 5.15: Mean & Standard deviation plot for lexical character-based features



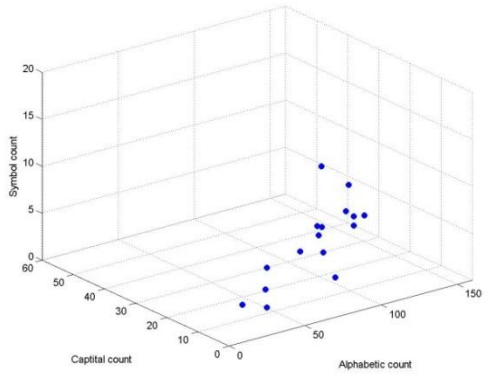
(a) User2



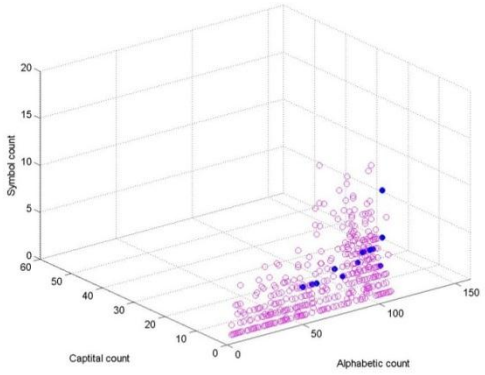
(b) User10



(c) User24



(d) User22



(e) All users

Figure 5.16: 3D plot of lexical character based feature

In order to appreciate the difficulties involved in discriminating between users successfully, Figure 5.16 illustrates 3D plots of an individual users' lexical character based feature data utilising the alphabetic count, digit count and symbol count values in each input vector. The Figure 5.16 (a) – (d) shows a 3D plot for an individual user

and Figure 5.16 (e) shows a 3D plot for all users' input data. As can be seen from the above Figure, it is possible for *User 2*, *User 10* and *User 24* (see Figure 5.16 (a)-(c)) to be discriminated from each other however, it can be difficult to classify *User 22* from other users as their input data shows a fairly large spread of values. This illustrates that lexical character-based features have some level of discriminative ability for some users.

Figure 5.17 shows a mean and standard deviation plot of lexical word based features for each user. It demonstrates that many users share latency spread of intra-class variance. For example, *Users 3, 6, 10, 23,* and *27* have overlapping latency spreads that coincide with each other. Therefore, it is not expected that this feature would provide a robust feature for these users. However, some users have latency spread that slightly differs from others e.g. *User 14*, thereby demonstrating this feature could provide additional discriminative information to this user.

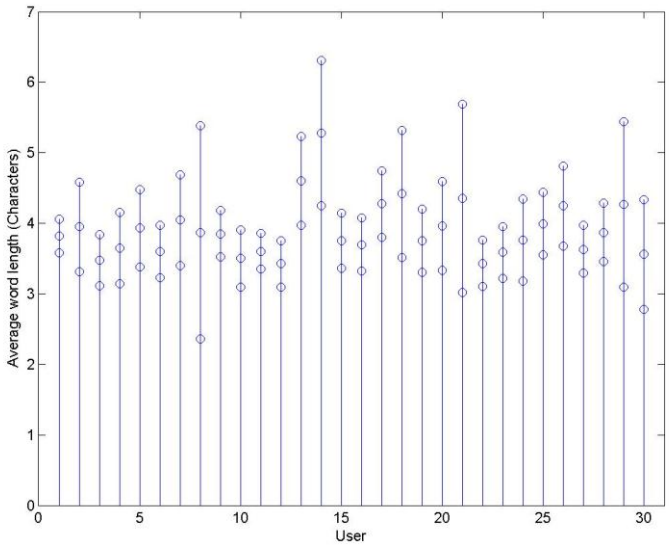
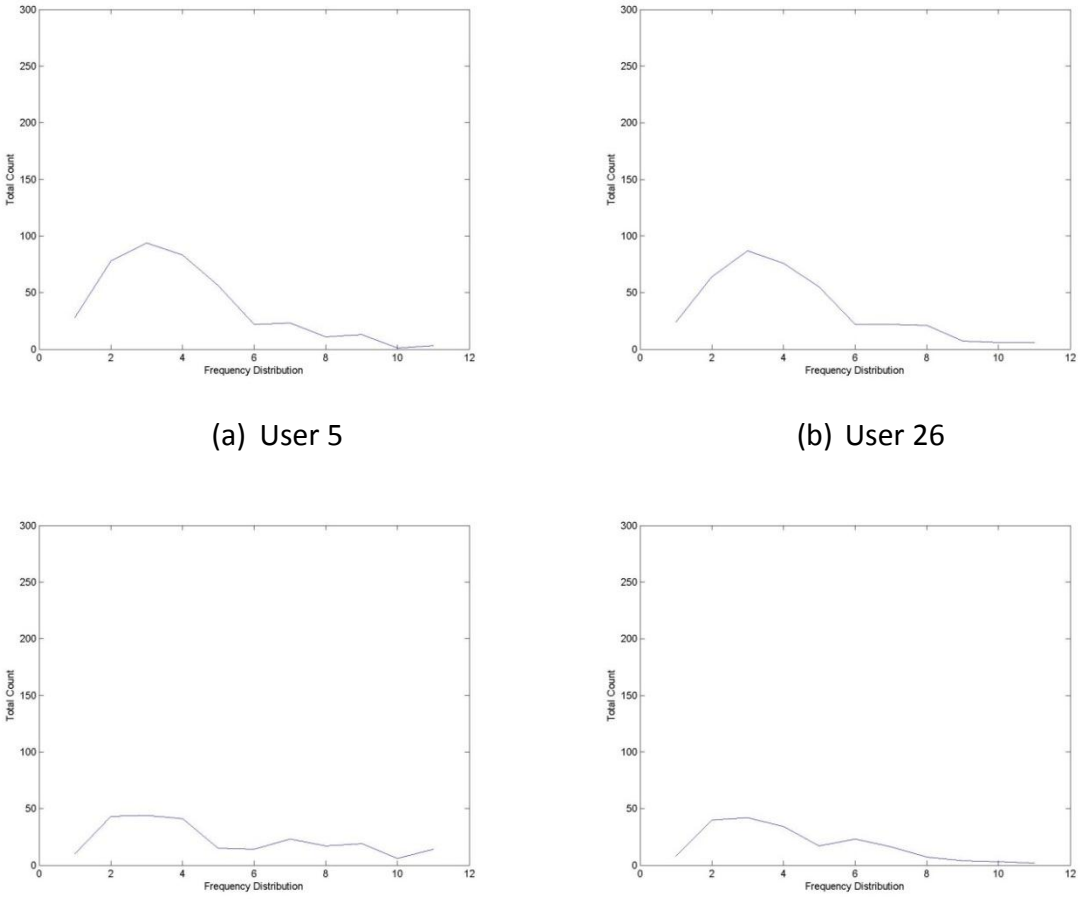


Figure 5.17: Mean & Standard deviation plot for average word length feature

Figure 5.18 illustrates examples of the word length distribution usage for an individual user. In general users tend to use short words in their messages with an average 4

characters. Therefore, the usage of short words may not be useful for discriminating between users. However, the usage of long words could provide some level of discrimination between users. For example, it is possible to discriminate *User 14* and *User 18* from each other as *User 14* tends to use more than 8 characters in their message compared to *User 18*.



(a) User 5 (b) User 26

(c) User 14 (d) User 18

Figure 5.18: Word length usages for individual user

5.4.2.3 Syntactic features

Figure 5.19 shows the quantity of punctuation usage in messages for each user. As can be seen from the figure, punctuation usages vary between users. Some users use greater than four times in their messages. On the other hand, some users rarely used

punctuation in their messages. Therefore these features might provide some level of discrimination for some users.

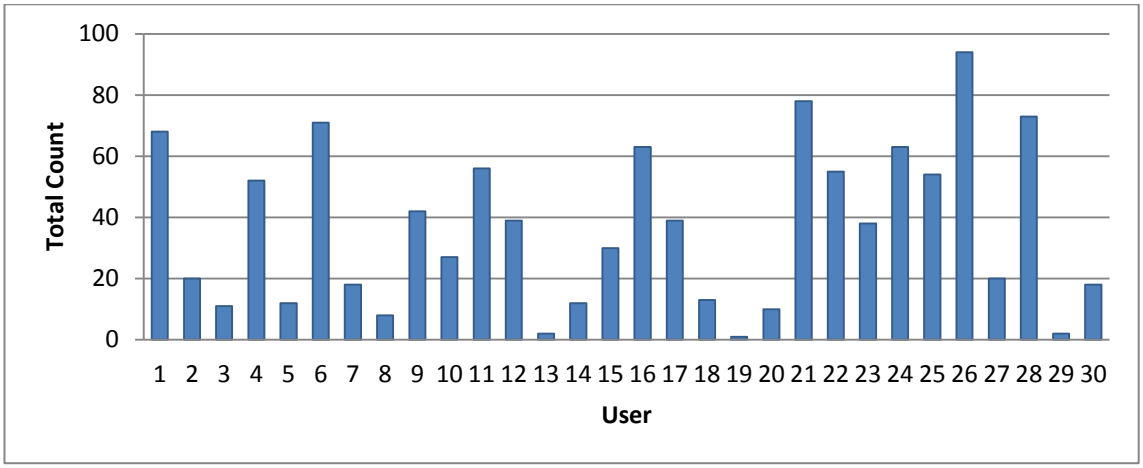
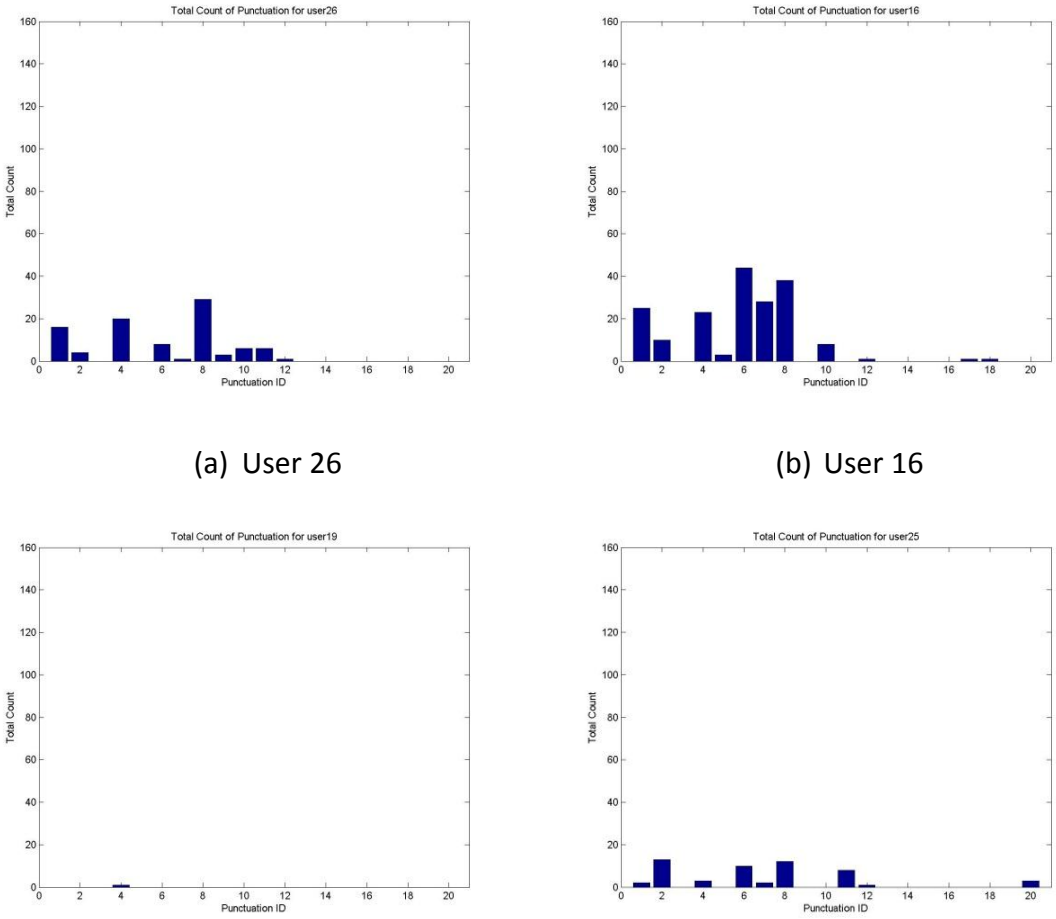


Figure 5.19: Total punctuation mark count for each user

Punctuation mark usage features can be useful for classification if a user uses that punctuation mark in a single message but no other users used it. As a result, further in-depth analysis was conducted to find the similarity of punctuation usage between users. Figure 5.20 shows examples of individual punctuation usage. An initial analysis based on individual punctuation mark usage indicates that users used similar sets of common punctuations such as question mark, full stop, and commas. In general, each user shared 5 types of punctuation marks with other users. For the worst case, *User 16* and *User 26* shared 8 types of punctuation marks however, their usage in terms of quantity is different as shown in Figure 5.20. Therefore, some of these common punctuation usages could be useful for identifying users. In addition, the lack of punctuation usage could also provide additional discriminative information.

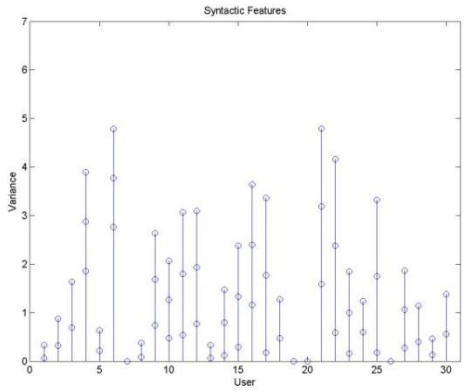


(a) User 26 (b) User 16

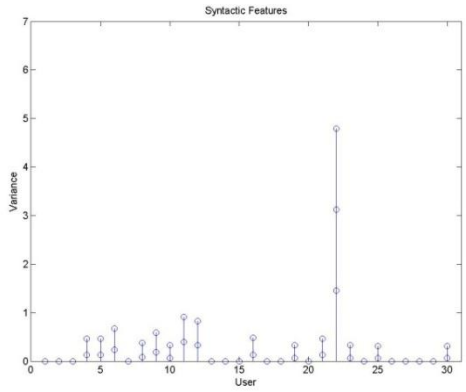
(c) User 19 (d) User 25

Figure 5.20: Punctuation mark usages for individual user

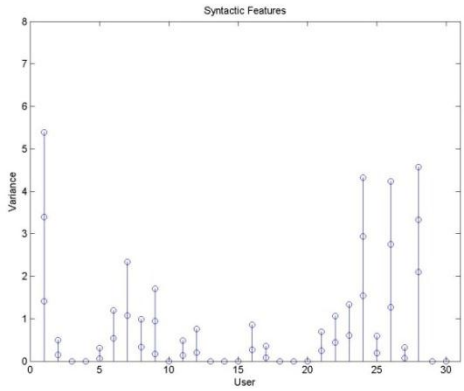
Figure 5.21 shows a mean and standard deviation plot of syntactic features for each user. An initial analysis based on inter class and intra class variances shows that the similarity between a users' individual input data, or intra-class variance are not ideal but some users have smaller intra-class variance than others. It should be noted that, users may have intra-class variance zero because they do not have data on that features. Although many users have an overlapping latency spread with each other, there are some levels of discriminative ability for this linguistic feature category to be used as useful feature. For example, the use of punctuation after a space feature can provide additional discriminative information for *User 22* because the user's latency spread showed a very clear discriminative boundary.



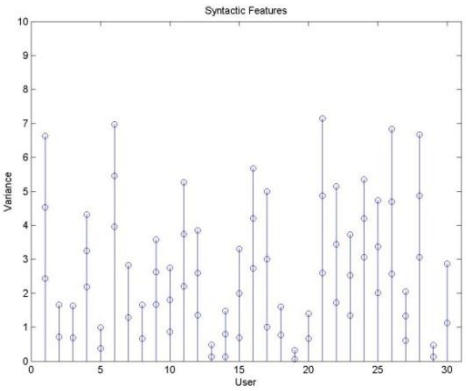
(a) #space after punctuation mark



(b) #punctuation mark after space



(c) #no space after punctuation mark

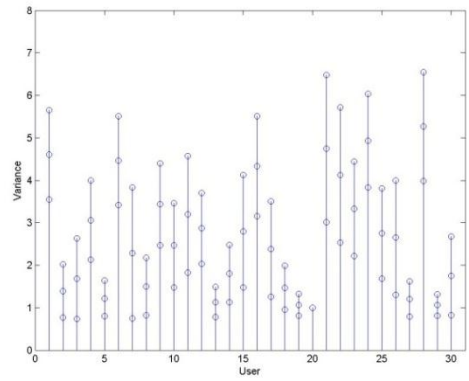


(d) Total #punctuation marks

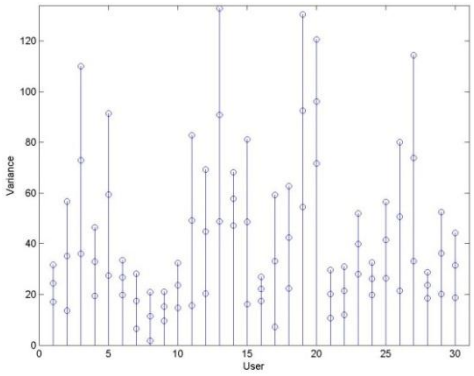
Figure 5.21: Mean & Standard deviation plot for syntactic features

5.4.2.4 Structural features

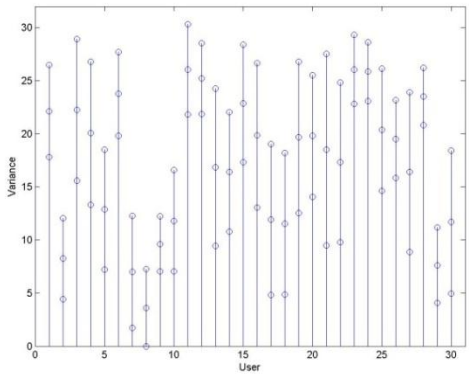
Figure 5.22 shows a mean and standard deviation plot of structural features for each user. An initial analysis of structural features demonstrates that individual user has a fairly large intra class variance and also there are a number of users that have latency spreads that coincide with other users demonstrating that they have very low inter-class variance. However, some features showed positive information could be used for discriminating between users. For example, the number of sentence feature can contribute discriminative information to classify *User 21* from other users.



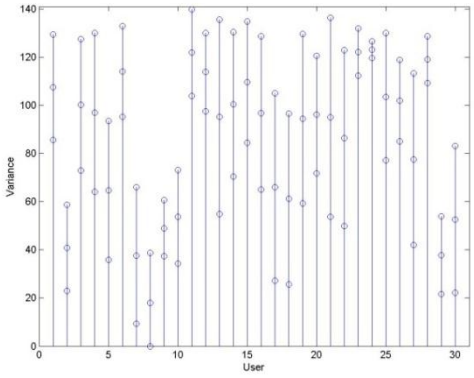
(a) #sentences



(b) Average sentence length



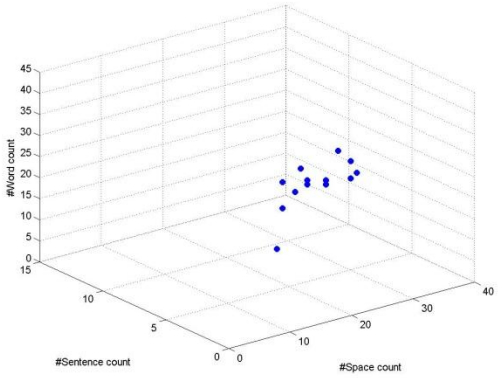
(c) #words



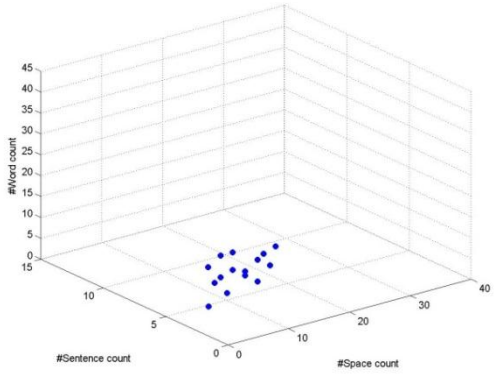
(d) Message length

Figure 5.22: Mean & Standard deviation plot for structure features

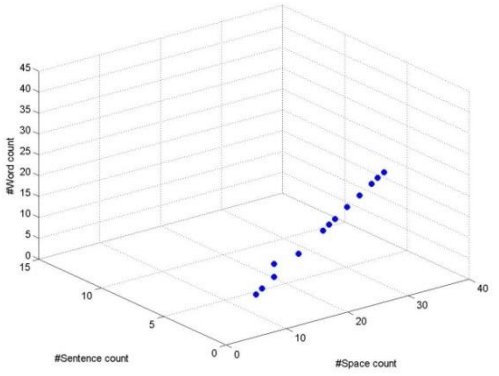
Figure 5.23 shows an example of a plot of an individual users' structural feature data utilising the space count, sentence count and word count values in each input vector for each user and for all users. Consider a 3D plot for all users in Figure 5.23 (e), it would be difficult to classify all users as the figure shows a fairly complex classification space. However, when considering a 3D plot for individual user, these features can be useful for classification. For example, it is easy to discriminate *User 6* and *User 9* from each other as their plotting areas do not coincide as is shown in Figure 5.23 (a) and (b) . On the other hand, it could be difficult to classify *User 19* and *User 20* as they share the same area plot of data between each other as is shown in Figure 5.23 (c) and (d).



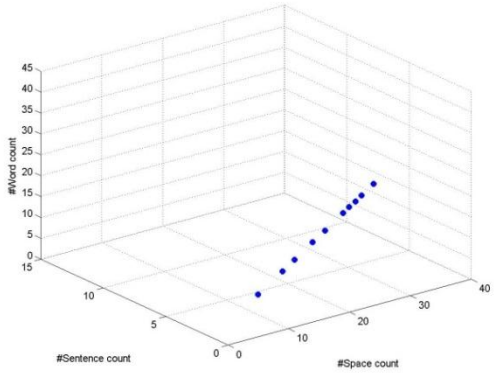
(a) User 6



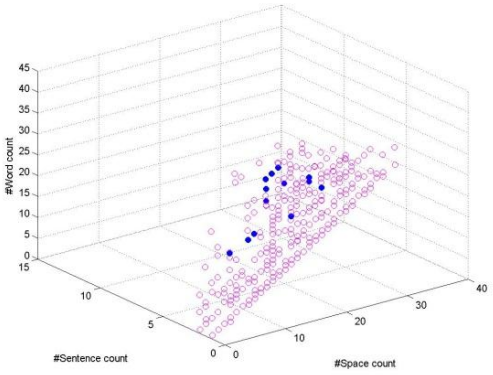
(b) User 9



(c) User 19



(d) User 20



(e) All users

Figure 5.23: 3D plot of structure features

From the above analysis, there is similarity with previous studies. The result demonstrates that a number of features have the potential to be used for discriminating between users. An initial analysis based on user word profiling features

showed that users have used different sets of abbreviation and emotion in their message and the quantity of each words usage is different between users. A lexical character-based feature could provide discriminative information to classify users. Although users tend to use short words in messages, some of the word based features such as long word distribution could be useful for discriminating some users who tend to use long words in their message. Syntactic feature analysis showed that the way of punctuation usage features could contribute additional discriminative information. There is some level of discriminative information from structure features on sentence count, word count, space count, average sentence length and message length features to be used for classification.

Similar to the previous feature analysis, there are some features contain negative information to classify user and can result in poor classification performance. Therefore, it is suggested that selecting only useful features as input vector but are dependent upon an individual analysis will improve the performance of classification with less classification complexity and time consuming.

5.5 Feature Identification and Evaluation

5.5.1 NUS Dataset

The descriptive statistical study identified several features which may provide positive information for a classification process. In order to analyse the impact of the linguistic features when evaluated in a multi-dimensional fashion, three preliminary studies were conducted. The methodology employed two types of profile techniques: static and dynamic. For the static profiling technique, two types of static profiling approaches were examined. In the first experiment, individual user profiles were

created using a static feature template. This technique used all the 381 linguistic features (317 keyword profiling features, 33 lexical features, 23 syntactic features and 8 structural features) as the input vector. In the second experiment, the profiling technique was based upon a dynamic keyword profiling and static feature template. This technique utilised a dynamic keyword profiling, which is the frequency of keywords used in a message based upon individual keyword profile, and the remaining features from lexical, syntactic and structural features to create individual user profile. As a result, a total number of 64 features were used as an individual input vector. Finally, the last experiment employed a dynamic profiling technique. In order to make the user profile dynamic, a feature selection technique was employed. In this study, the t-test statistic was applied to select a subset of useful features to create individual user profiling. This is because the technique is fast to compute and very efficient (Guyon and Elisseeff, 2003). As a result, a classification result is made based upon the most significant discriminative features rather than using all features which may contain not useful or no discriminative information for a particular user. Since there is no single classification method that can solve all given problems (Wolpert and Macready, 1997), a Feed-Forward Multi-Layered Perceptron (FF MLP) neural network classification approach was employed. This approach has also been extensively employed by classification problems such as this (Jain *et al*, 1999b; Cho *et al*, 2000; Clarke and Furnell, 2006; Tsimboukakis and Tambouratzis, 2010). To perform classification, the following FF MLP network parameter variables were defined:

- Transfer Function – Hyperbolic tangent sigmoid function
- Training Algorithm- Gradient descent with momentum and adaptive learning rate

- Number of network layers - 3
- Number of neurons per layer –20, 5, 1
- Number of training epochs - 2000

All three experiments were examined using FF MLP neural network with an identical configuration to ensure a consistent and meaningful evaluation of the feature vector. For individual users, data was divided into two datasets: 2/3 of data was used to generate a user profile and the remaining 1/3 was used to evaluate the classification's performance. The pattern classification tests were performed with one user acting as the valid authorised user, whilst all the other users were acting as imposters. This is a standard methodical approach (Clarke *et al*, 2002; Indovina *et al*, 2003; Halteren, 2004). The results for the preliminary study are presented in the following sections.

5.5.1.1 Static feature profiling approach

In this experiment, all of the 381 linguistic features were used as the input vector to discriminate between users using a FF MLP neural network. Table 5.6 demonstrates the average results for the FAR, FRR and EER from all 26 users. In general, the results show that the classification algorithm is unable to discriminate between the valid and invalid user as the performance is fairly poor with an average EER 39.3%. The best case individual user did achieve an EER 10%. However, the result of the worst case individual performance is an unacceptable EER of 61%.

#Features used as input vector	FAR (%)	FRR (%)	EER (%)
381	42.5	36.1	39.3

Table 5.6: Experimental results by employing static feature approach

5.5.1.2 Dynamic word profiling and static feature approach

In this experiment, a subset of abbreviation and emotional words based on user's usage from historical messages was selected to create word profiling for each user. Then the frequencies of these word appear in a message has been counted to represent a user keyword feature. Therefore, the number of word profiling features has been reduced from 318 to 1 feature. In order to examine the possibility of this profiling technique for discriminating between users, a total of 65 linguistic features were used as input vector. Table 5.7 demonstrates the average classification performance results using FF MLP neural network classifier with an identical configuration as the previous study. By using dynamic word profiling combined with static feature template, the overall performance of user classification achieved EER 34.5%. The best result of individual performance was achieving an EER 10%. However, the result of worst case individual performance showed unacceptable EER 59%.

#Features used as input vector	FAR (%)	FRR (%)	EER (%)
65	34.7	34.4	34.5

Table 5.7: Experimental results by employing dynamic word profiling

5.5.1.3 Dynamic feature profiling approach

In this experiment, a statistical t-test and its p value for feature selection were applied. The t-test is used to assess whether the means of two classes are statistically different from each other by calculating a ratio between the different of two class means and the variability of the two classes. In order to select effective subset of features for individual user, t-test statistics are performed on each feature from a total of 65 features and the p-value examined. According to their p-value, the features can be sorted with the less p-value being a more discriminative feature. In order to select effective subset of features for individual user, two types of dynamic profiling

techniques were examined. The first technique utilised top 10 of the most discriminative features as an input vector and the second technique utilised all features that have p-value less than 0.05 ($p<0.05$) as an input vector. By using the second technique, the number of features used as an input vector is varied between users. The t-test approach is illustrated in Equation below.

$$t_i = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{{S_x}^2}{n} + \frac{{S_y}^2}{m}}}$$

Where:

t_i is the t- statistical value for $Feature_i$

$\bar{x}$ is the mean value of valid user input samples

$\bar{y}$ is the mean value of invalid user input samples

S_x is the standard deviations of valid user samples

S_y is the standard deviations of invalid user samples

n is the size of valid user input samples

m is the size of invalid user input samples

The standard deviation is illustrated in Equation below

$$S = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$

Where:

S is the standard deviations of input samples

N is the size of input samples

x_i is the value of input sample at element i

$\bar{x}$ is the mean value of input samples

Table 5.8 demonstrates the average classification performance results using FF MLP neural network classifier with the same configuration as previously studied. By using the most top 10 discriminative feature to classify a user, the classifier produced overall classification performance with EER 29.3%. The best case of individual performance achieved was an EER of 9%. However, the worst case of individual performance achieved an EER of 45%. As shown in Table 5.8, by using dynamic feature profiling based on all features that have p-value less than 0.05 ($p<0.05$) to discriminate user, the overall classification performance shows an EER of 28% with an average 20 dimensional input feature vector. The best case of individual performance achieved an EER of 8%. In contrast, the worst case of individual performance achieved an EER of 42%.

Feature selection method	#Features used as input vector	FAR (%)	FRR (%)	EER (%)
Top 10	10	29.3	29.2	29.3
p-value <0.05	Varied between 8-44	27.3	28.6	28.0

Table 5.8: Experimental results by employing a dynamic feature approach

5.5.2 Discussion on preliminary studies

By using a static feature template approach to profile individual user, the classifier produced a fairly high EER or poor performance as shown in Table 5.6. A possible reason for this is that the high dimensional input feature vector may contain a combination of positive, negative and no discriminative information (Clarke, 2004). This is expected since there are a number of abbreviation and emotional words that

rarely appear in a message. Alternatively, these large input vector neural networks might require a larger and more complex neural network to compensate for the increased dimensionality of the classification problem. It is interesting to note that, *User 21* who is the most rarely used abbreviation and emotional words in their message profile achieved the lowest EER. This indicates that the lack of use can provide strong discrimination power.

As can be seen from Table 5.7, the dynamic word profiling and static feature template technique have performed better than the static feature template profiling approach by decreasing the EER by 4%. This indicates that merging all words from individual word profiling and treating them as a single feature can provide additional discriminative information towards the classification process. However, the 62 dimensional input vector features may still contain negative discriminative information, as the overall classification performance was fairly high. Considering individual classification performance, *User 21* achieved the lowest EER.

The dynamic feature profiling approach has performed significantly better than other profiling techniques. This is expected since the input vector contains only strong discriminative features for each individual user. Indeed, using the feature selection technique based on $p<0.5$ achieved the best classification performance. One possible reason is that by using a fixed number of features to select in order to create a user profile, the input vector can contain negative features. For example, there are two users, *User 19* and *User 25* that have achieved slightly better results when they utilise only 8 features compared to 10 features. Therefore, care is required in determining how many of the best features to select for the best performance.

5.5.3 PU Dataset

From the descriptive statistical study, the findings showed that several features may provide discriminative information for a classification process. Therefore, a preliminary study was conducted to explore the impact of linguistic features on the performance of classification. Building upon the finding of a previous preliminary study, the methodology in this study employed three user profiling techniques and a classification algorithm utilised are based on feed-forward MLP neural network. An identical network configuration with the previous study was utilised in all three experiments. For the actual experiment, an identical research methodology was employed. The results for the preliminary study are presented in the following sections.

5.5.3.1 Static feature profiling approach

In this experiment, all of 197 features were used as user's input vector. The FF MLP neural network was utilised as a classifier to discriminate between users. Table 5.9 demonstrates the average results for the FAR, FRR and EER from all 30 users. The results show that the classification algorithm produced fairly poor performance with an average EER of 27.8%. The best case individual user was achieving an EER of 10%. However, the result of worst case individual performance showed unacceptable an EER of 67%.

#Features used as input vector	FAR (%)	FRR (%)	EER (%)
197	29.1	27.5	27.8

Table 5.9: Experimental results by employing static feature approach

5.5.3.2 Dynamic word profiling and static feature template approach

In this experiment, a dynamic word profiling technique that represents user word profiling features as a single vector was combined with a static feature template technique were used to create a user profile. As a result, a total of 65 linguistic features were used as the input vector. Table 5.10 demonstrates the average classification performance results using a FF MLP neural network classifier. By using dynamic word profiling combined with a static feature template, the overall performance of user classification achieved an EER of 22.3%. The best result of individual performance was achieving an EER of 0.6%. However, the result of worst case individual performance showed an unacceptable EER of 58%.

#Features used as input vector	FAR (%)	FRR (%)	EER (%)
65	21.5	22.6	22.3

Table 5.10: Experimental results by employing dynamic word profiling

5.5.3.3 Dynamic feature profiling approach

In this last experiment, the dynamic feature profiling technique was utilised to create a user profile. A statistical t-test method was applied and its p value can be used to explore the discriminative power of features. Two ways of selecting the effecting subset of features for individual user were examined. The first technique used top 10 of the most useful discriminative features and the latter technique utilised all features that have p-value less than 0.05 ($p<0.05$) to create the user input vector.

Table 5.11 demonstrates the average classification performance results using a FF MLP neural network classifier. By using top 10 discriminative features to classify a user, the classifier produced overall classification performance with an EER 16.85%. The best case of individual performance achieved an EER of 53%. However, the worst case of individual performance achieved an EER 0.06%. As also shown in Table 5.11, by using dynamic feature profiling based on all features that have p-value less than 0.05 ($p<0.05$) to discriminate a user, the overall classification performance shows an EER of 16.41% with an average 27 dimensional input feature vector. The best case of individual performance achieved an EER of 0.06%. In contrast, the worst case of individual performance achieved an EER of 61%.

Feature selection method	#Features used as input vector	FAR (%)	FRR (%)	EER (%)
Top 10	10	18.4	14.7	16.8
p-value<0.05	Varied between 13-41	14.8	18.2	16.4

Table 5.11: Experimental results by employing dynamic feature approach

5.5.4 Discussion on preliminary studies

As can be seen from Table 5.9, using a static feature template approach to profile an individual user, the classifier produced a fairly high EER. A dynamic word profiling technique can improve the discriminative power to user word profiling features because the second experiment showed better the classification performance as shown in Table 5.10.

Based upon the above three set of results, it demonstrates that the dynamic profiling technique has performed the best in term of classification performance. Utilising features selection based on p value<0.05 achieved slightly better performance than using top 10 of the discriminative power features. Since FRR rate of the p-value method is lower than using 10 features, the p-value method can provide more tight security as the number of users was more rejected.

5.6 Conclusion

In general, utilising the dynamic feature approach can achieve a good level of accuracy and provides better classification performance than using a static approach. By employing a word profiling technique, individual user word profiling can be added or deleted over a period of time without affecting the user word profiling feature. Since a user's typing style can change over time therefore by using a feature selection technique, a user profile can contain a user's up to date discriminative information rather than using old feature that might not be useful anymore. As a result, the dynamic profiling technique is the best solution to create user profile.

From both preliminary studies, using the p value<0.05 criteria to select an effective subset of features produced better performance than using top 10 of the

discriminative power features. Although using the p value<0.05 method produced varying size of input vector between users and can be up to 40 features, it is ensure that maximise of significant features were used as input vector and result in optimal classification performance.

By using an identical network configuration, FF MLP neural network provided a good level of overall performance however, according to previous study (Clarke and Furnell, 2006) the classification performance for an individual user can be improved if the most effective network configuration for each user is employed. Therefore, a number of experiments with various network configurations should be conducted and investigated.

Based upon the above two set of results, the studies have shown the ability for classification algorithms to correctly discriminate between users with a fairly good degree of performance based on a dynamic profiling technique. By using the descriptive statistics method, the findings showed that a user has a different style of composing SMS and there are a number of features may provide discriminatory information toward a classification process. Two preliminary studies were formed to examine the effectiveness of linguistic features by employing three profiling techniques: static feature, dynamic word profile with static template features and dynamic profiling. All experiments utilised an FF MLP neural network with an identical configuration. Based upon the experimental results, dynamic profiling techniques have been shown to be the best solution to create user profiling for classification because of its performance. However, a further investigation is necessary in order to optimise the classification performance. The next chapter presents a complete experiment of the linguistic profiling technique to classify users.

6 Text-based Multi-Modal Biometrics

6.1 Introduction

This chapter presents an in-depth investigation into approaches that enable text-based authentication. The study comprises of two primary objectives: to evaluate the performance of linguistic profiling, and determine the value of utilising a multi-modal approach. Building upon the previous chapter, that highlighted the features and initial viability of linguistic profiling; this chapter will proceed to evaluate appropriate classifiers to determine the overall performance that can be achieved. However, as the initial studies demonstrated, the highly behavioural nature of the biometric characteristics is likely to require a composite approach that involves additional biometric-based information. As such, this research also investigated two further behavioural biometric techniques, looking to apply these techniques as a multi-modal biometric authentication method for mobile devices. The performance of keystroke dynamics, behaviour profiling and the combination of these three techniques is presented.

6.2 Performance of text-based biometric techniques

This section presents the baseline performance of the individual technique. Then, the baseline result will be used in comparison in order to determine how much of the multi-modal technique improvement comes from combining results from different modalities versus a single modality. The following section details the description of the experiments used in each technique and provides an analysis of the results.

6.2.1 Linguistic profiling

A number of linguistic features can provide discriminative information and dynamic profiling techniques providing the opportunity to develop a more meaningful user profile. The primarily results presented by using a common classifier FF MLP neural network showed that a user can be correctly identified with a good degree of success. However, from the literature reviews in Chapter 4, a number of pattern classification techniques were introduced that performed well in studies. In order to identify the most appropriate classifier- so that the optimal level of system performance can be achieved, a number of algorithms were applied. The study results are illustrated in the following section.

6.2.1.1 Dataset

In this experiment, the SMS corpus collected at Plymouth University was used in the analytical process in order to maximise the number of users. A total of 30 participants were obtained with total of 487 text messages. Table 6.1 demonstrates the final dataset for utilizing in this experiment.

Number of participants	30
Number of SMS text messages	487
Average number of messages per user	16
The maximum length of text messages	128 characters
Average length of messages per user	17 words
Average length of messages	83 characters

Table 6.1: Description of the final dataset of PU SMS corpus

6.2.1.2 Procedure

A number of program scripts were generated to perform various tasks such as manipulation of the input data, generation of neural networks and evaluation of the

performance. Several of these scripts were utilised in the classification algorithms and form the following common components:

- **Data extraction function**: extracts user's linguistic features from SMS message database.
- **Normalise function**: normalises data into the range of 0-1. Due to a previous study (Snelick *et al*, 2003) have proven that it is necessary to normalise input data into the same range as the output on order to reduce the complexity of the resulting classifier and (often) the performance.
- **Dataset splitting function**: Splits the data into two sets: approximately 2/3 of the data is used for training a classifier and 1/3 is used to validate the performance of a classifier.
- **Feature selection function**: based upon an individual user, useful features are selected from a feature database.
- **Classifier selection function**: Chooses a classifier and sets up the parameters of the classifier.
- **Evaluation function**: Calculates the FAR, FRR and EER to evaluate the performance of the selected classifier.

The calling sequence for each function is illustrated in Figure 6.1: data information flows from one function to another via the directional arrows.

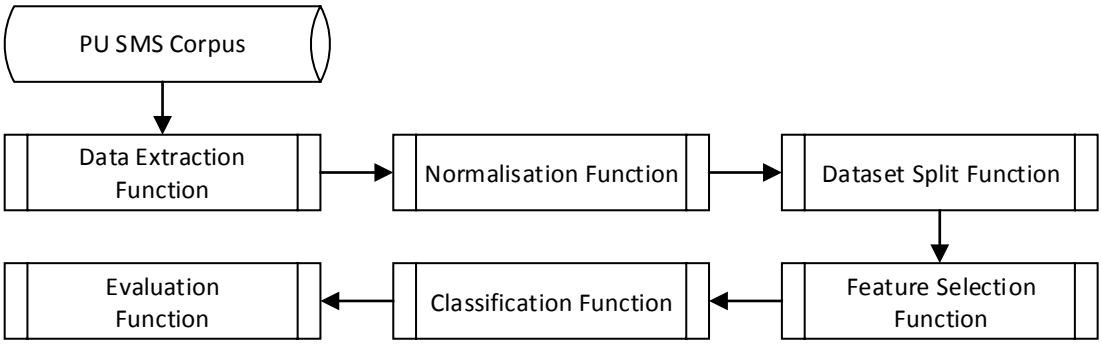


Figure 6.1: Linguistic profiling system functions flow diagram

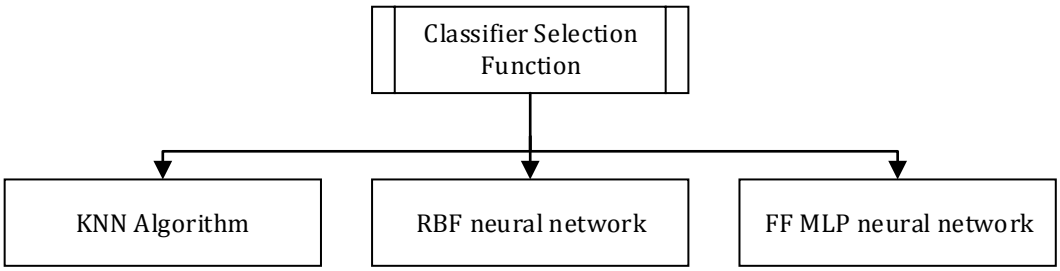


Figure 6.2: The three classifiers being employed

In order to investigate the linguistic profiling's effectiveness, four types of linguistic features were examined: (1) user's words profiling, (2) lexical, (3) syntactic and (4) structural were utilised. The frequency distribution of a total of 133 abbreviations and emotional words including 64 discriminating characteristics of every possible type of feature were used to create user's words profiles. It is evident from Chapter 5 that a dynamic feature-based approach performs substantially better than other methods and as such this method was applied. To create a user profile, the t-test ranking measure was utilised to rank input features according to its discriminative capability. From the ranking list, features with p value less than 0.05 ($p < 0.05$) were selected to create input vectors thereby minimising unnecessary feature information. Therefore, the number of linguistic features required for discrimination will vary between users. For the actual experiment, in line with standard methodological approaches, each

user's data was divided into two dataset: 2/3 of data was used to generate a user profile and the remaining 1/3 was used to evaluate the classification's performance. The pattern classification tests were performed with one user acting as the valid authorised user, whilst all the other users were acting as imposters and each user was given the opportunity of acting as the authorised user. The final results presented are an average across this population.

6.2.1.3 Performance of Linguistic Profiling

Many classification techniques from statistical methods to advanced neural networks have been used to accomplish classification tasks in previous studies (as indicated in the literature review in Chapter 4). Since there is no single classification method that can solve all given problems, the most effective and well-known pattern classification approaches were employed as the linguistic profiling technique: K-Nearest Neighbour (K-NN) classification algorithm, Radial Basis Function (RBF) and Feed Forward Multi-Layered Perceptron(FF-MLP) neural networks. From previous researches (Diederich *et al*, 2003; Tuurkoglu *et al*, 2007; Calix *et al*, 2008; Jockers and Witten, 2010; Tsimboukakis and Tambouratzis, 2010; Kusakci, 2012) these classifiers have shown the ability to deal efficiently with a high-dimensional and small size of data. Each of the classification algorithms performed numerous iterations, changing the various network parameters in order to optimise the performance of the classifier. The results for the study are presented in the following sections.

A. K-Nearest Neighbour Algorithm (K-NN)

The *K-NN* is one of the most popular and simple classification methods used in the text classification problem domain. Many researchers have found that the *K-NN* algorithm achieves very good performance in their experiments on different data sets (Yang and

Liu, 1999; Joachims, 1998; Li *et al*, 2002; Calix *et al*, 2008; Kusakci, 2012). In *K-NN* classification, the similarity between a validation sample vector and training vector is computed using distance measure techniques. The class of feature vector is determined by selecting the class that has the majority among the k-nearest neighbours. The performance of this algorithm mainly depends on two factors, a suitable similarity function and an appropriate value for the parameter *k.* This experiment utilises a Euclidean distance measure technique by default. The number of *k* used in the classification was set to different values between 1 and 5. This range gave the best results, whereas the performance decline with higher values. As each user creates an individual neural network and therefore respective FA, FR and EE rates, the results illustrated are an average of all the users' EERs. The best overall performance obtained using K-*NN* network is demonstrated in Table 6.2.

Average			**Worst Case**				**Best Case**			
FAR (%)	**FRR (%)**	**EER (%)**	**User**	**FAR (%)**	**FRR (%)**	**EER (%)**	**User**	**FAR (%)**	**FRR (%)**	**EER (%)**
2.0	57.9	30.0	19	1.8	100.0	50.0	11	0.6	0.0	0.3

Table 6.2: The classification results by using the most successful *K-NN* network

As can be seen from Table 6.2, the results showed that using the KNN classifier to classify users gave a fairly poor classification performance with an average EER 30%. This was achieved utilising the most successful network using a configuration with 2 nearest neighbours (*k*=2). An analysis of the individual error rate gave rise to a large spread of results with some users performing better than others. The best individual result achieved the lowest EER 0.3 % by *User 11*. Further analysis shows that 2 users achieved an EER of less than 10% and 7 users have an EER smaller than 20%. However, the worst individual result showed a fairy high EER of 50% by several users including

Users 13,17,19,27. The error rate for each user obtained using the most successful network configuration is shown in Figure 6.3.

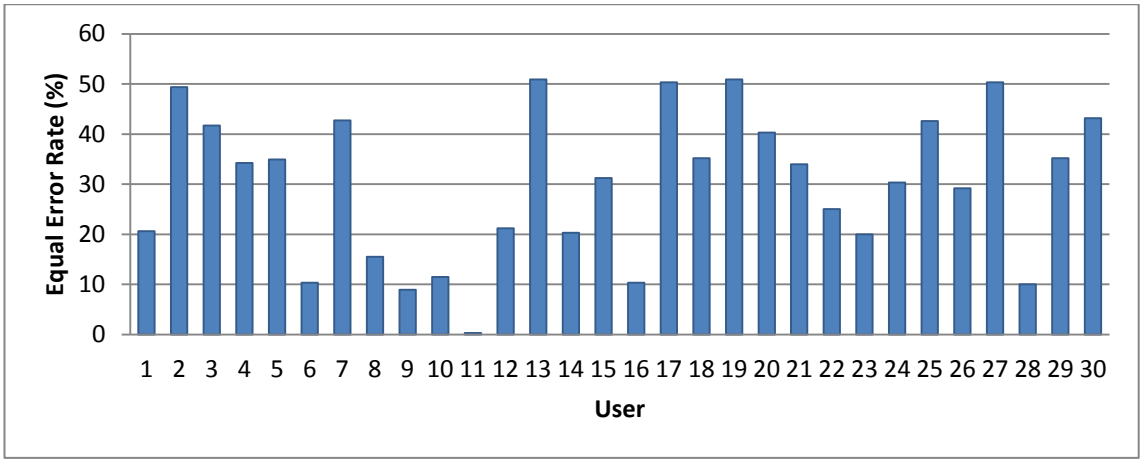


Figure 6.3: Individual results by using the best K-NN network configuration

It is interesting to note that as the *k* increases, the performance of the network tends to decrease dramatically. When *k*= 5, the overall performance obtained the worst EER of 35.9%. This could be caused by the training sample dataset being larger for the imposters' class than for the authorised user class. As such the results could favour the imposters as they have largest number of members with *k* nearest neighbours.

From an analysis of all the KNN investigation results, it was noticed that individual network performance varied considerably between network configurations and individual users obtained the optimal EER using different network configuration. From a feasibility perspective, an analysis of the optimal results would determine the best possible result that can be achieved in any one given network configuration per user. Figure 6.4 shows the optimal result of each user that can be achieved.

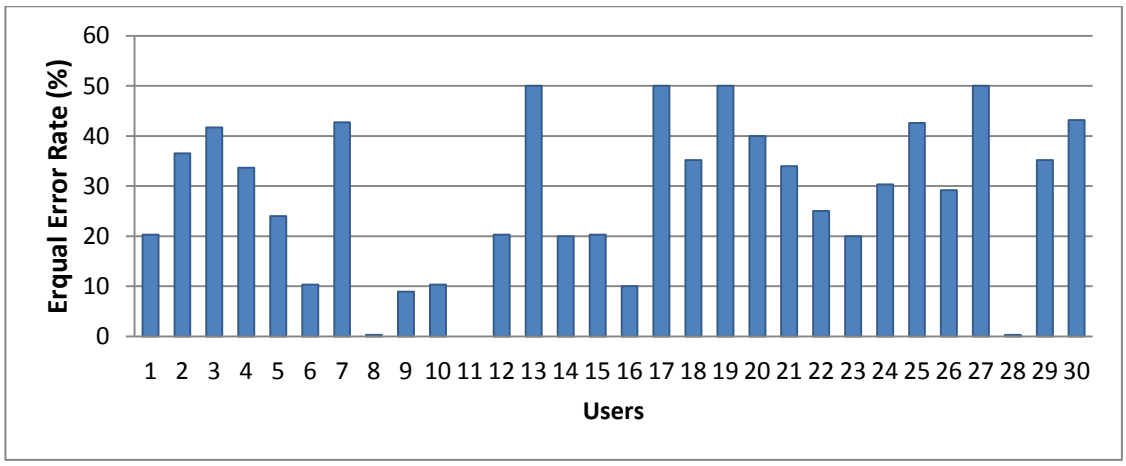


Figure 6.4: The optimal result for each user (KNN)

The results show that the overall performance was better than the successful network result of 3%, achieving an EER 27%. An analysis of individual error rates show that *User 11* achieved the best EER of 0% and *Users 13, 17, 19 and 27* obtained the worst EER of 50%. However, 50% of users obtained an EER of more than 30%. This suggests that this neural network technique might not be suitable to classify some users.

B. Radial Basis Function Networks (RBF)

An RBF network is amongst the most user configurable neural network topologies with only two network variables the performance goal and spread. To perform classification, a large number of iterative tests was required with both variables were changed independent of each other in order to achieve the optimal performance of the technique. The classification results that can be achieved by using the most successful network configuration are illustrated in Table 6.3.

Average			Worst Case				Best Case			
FAR (%)	FRR (%)	EER (%)	User	FAR (%)	FRR (%)	EER (%)	User	FAR (%)	FRR (%)	EER (%)
18.0	17.0	17.7	27	53.0	60.0	56.5	11,28	0.0	0.0	0.0

Table 6.3: The classification results by using the most successful *RBF* network

As can be seen from Table 6.3, a user can be classified using an RBF neural network with a good degree of classification performance with average an EER of 17.7%. This result was obtained by utilising the network configuration of the performance goal with 1, and spread value 0.5. Analysing through the user's performance shows a broad range of individual results, as illustrated in Figure 6.5, with *User 10 and User 11* achieving the lowest EER of 0%. In contrast, the *User 27* archived the worst EER of 56%. However, 53% of users have an EER of less than 20% with the same network configuration.

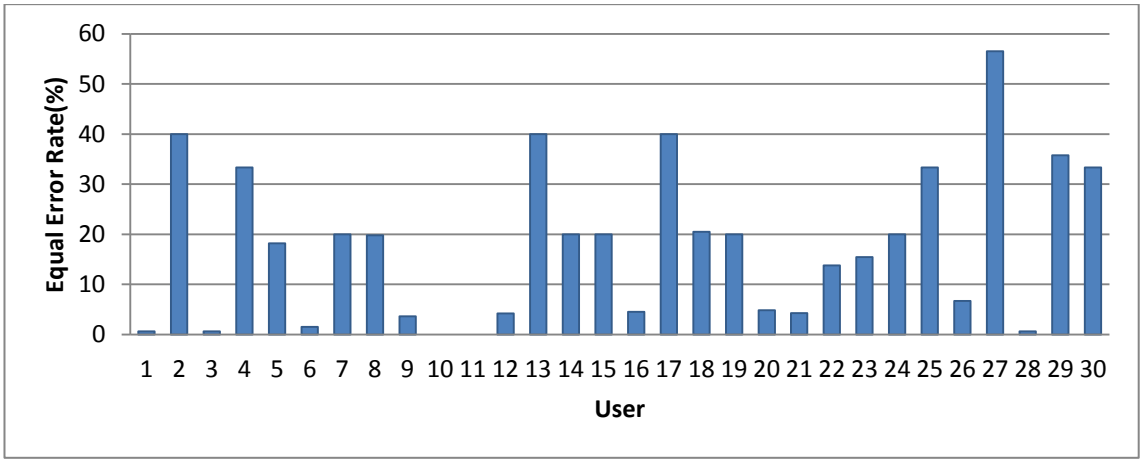


Figure 6.5: Individual results by using the best RBF network configuration

Figure 6.6 illustrated the best optimal results that each user can achieve. An analysis of individual user error rate showed that 20% of users achieved an EER of 0% and *User 27* obtained the worst EER of 37.1%. On average, the performance obtained with an EER of 10.4%, improving 7% compare to the most successful network result. In addition, 86% of users obtained an EER less than 20% and further 53% of users achieved an EER of less than 10%.

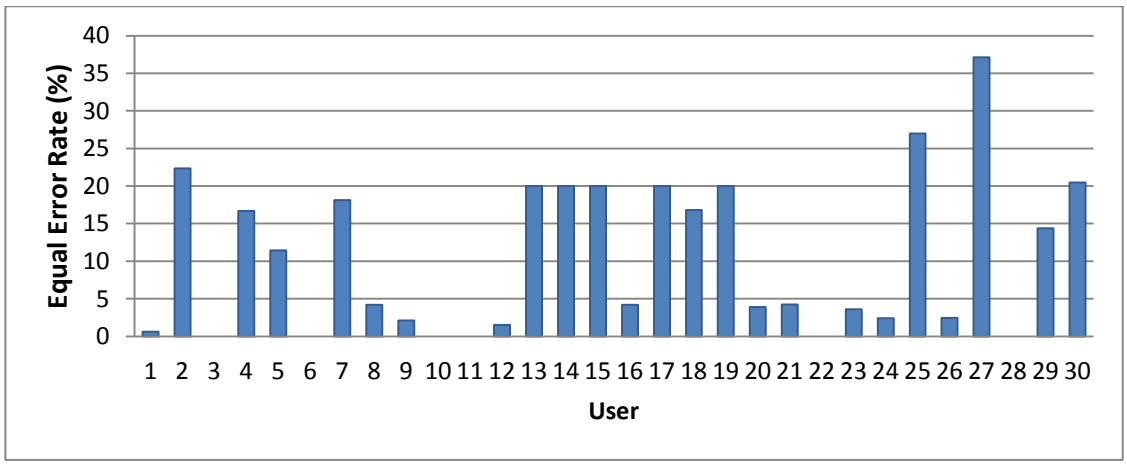


Figure 6.6: The optimal result for each user (RBF)

During the RBF investigation, it was noticed that by using the same goal value but various changes in the spread value gave varied in the individual network performance rate with 2-3 % different in an overall EER. However, the different configurations with goal value do appear to have a large effect on the overall performance. Furthermore, based upon the most successful network result it is noticed that the RBF has a large spread of performance between best and worst of 56%. Therefore, care needs to be taken as RBF is very sensitive to network changes.

C. Feed Forward Multi-layered Perceptron Network (FF-MLP)

The Feed Forward Multi-layered Perceptron Network is one of the most widely employed AI techniques utilised in pattern classification. With more network configuration variables available, the FF MLP neural network is a more complex classifier when compared to the RBF neural network. To perform classification, a large number of iterative tests were required to achieve the most optimal performance. From previous research (Stanczyk and Krzysztof, 2007; Selman and Husagic-Selman, 2011), two variables were identified as performing well for the classification problem.

- Transfer Function - Hyperbolic tangent sigmoid function
- Training Algorithm - Gradient descent with momentum and adaptive learning rate

For this study, the number of network layers was configured with 2 hidden layers. With two hidden layers the network can classify any data since they can always be represented as a sum or difference of some such simplexes classified by the second hidden layer (Selman and Husagic-Selman, 2011). Apart from the number of layers, the number of neurons per layer variables was modified. It should be noted that when the number of neurons is unnecessarily high the network easily learns but poorly generalise on new data (Nilsson, 1965). On the other hand, when there are too few hidden neurons the network may never learn the relationships amongst the input data. Since there is no precise indicator how many neurons should be used in the construction of a network, the FF-MLP network was initially configured with 2 neurons and this number will be increased in order to measure network performance variations with the view of achieving the most optimal performance rates.

The findings from this experiment are illustrated in Table 6.4. Using the FF MLP neural network to discriminate users resulted in a good degree of classification performance, with an overall EER of 13%. This result was achieved by employing the most successful network configuration of 4 neurons in the first hidden layer and a single output neuron with 2000 epochs.

Average			**Worst Case**				**Best Case**			
FAR (%)	**FRR (%)**	**EER (%)**	**User**	**FAR (%)**	**FRR (%)**	**EER (%)**	**User**	**FAR (%)**	**FRR (%)**	**EER (%)**
12.5	12.3	12.8	27	38.5	40.0	40.0	11	0.0	0.0	0.0

Table 6.4: The classification results by using the most successful *FF-MLP* network

An analysis of the individual performances highlights the differences in the EER between individual users, as illustrated in Figure 6.7, with *User 11* achieving the lowest EER of 0%. On the other hand, the worst individual result showed a fairy high EER 40 % by *User 27*; however, this was only exhibited by a single individual, with 67% of users having an EER less than 20%.

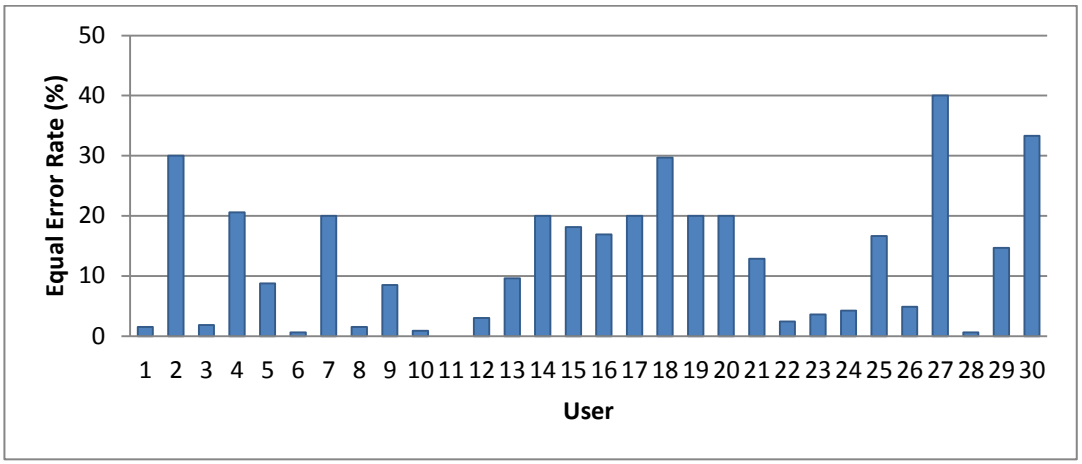


Figure 6.7: Individual results by using the best FF-MLP network

Through an analysis across spread settings, individual users obtained the best EER from different network configurations. An analysis of individual user's performance has identified *User 11 and 28* achieving the lowest EER of 0% as illustrated in Figure 6.8. The worst individual result was obtained by *User 30* with an EER of 30.9%. On average, the performance achieved an EER of 10.8%, improving 2% from the most successful network result. Furthermore, 80% of users obtained an EER of less than 20% with further 53% users achieving an EER of less than 10%.

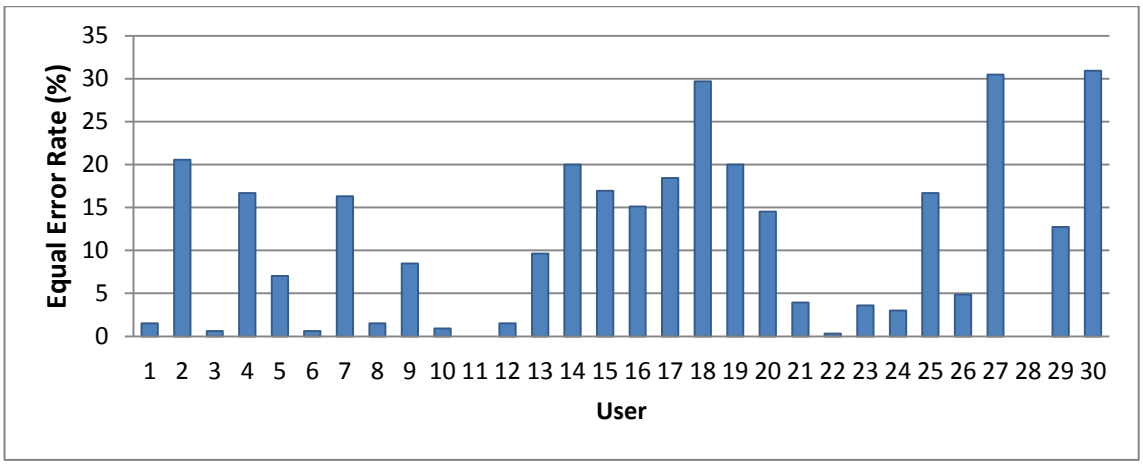


Figure 6.8: The optimal result for each user (FF-MLP)

Compared to KNN and RBF neural networks, the FF MLP neural network gave the best overall performance. Based upon the most successful neural network result, the classifier has a smallest spread of performance between best and worst case with a spread of only 40%. It could be suggested that this network provides a more robust and stable network topology. However, this technique is the most intensive of the three techniques as it is required to perform a large number of training epochs.

Additional analysis showed that an individual user's performance varied with the different neural network techniques and configurations, with some users performing better than others. It would be of interest to include these results as it could illustrate the best possible classification achieved by a neural network. The result could provide the meaning of how well authentication could take place with individual optimal network configuration.

The best results achieved by each user in any of the neural network techniques are illustrated in Figure 6.9. The results showed a wide range of EER with 20% of users achieving an EER of 0% and *User 27* obtaining the worst result with EER of more than

30%. However, the overall performance gave the best result with an EER 8.9%, the lowest error rates in this investigation as illustrated in Figure 6.10.

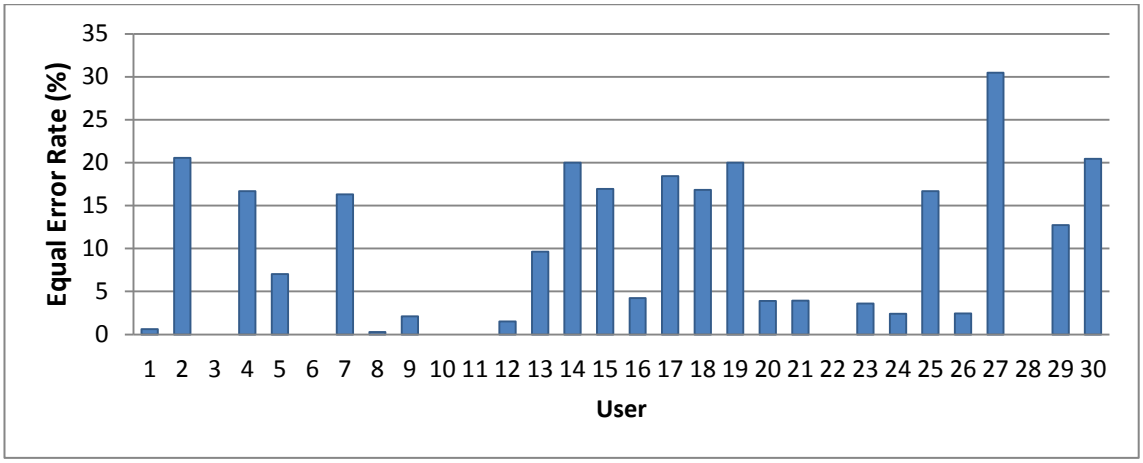


Figure 6.9: Optimal EER from the best case neural network

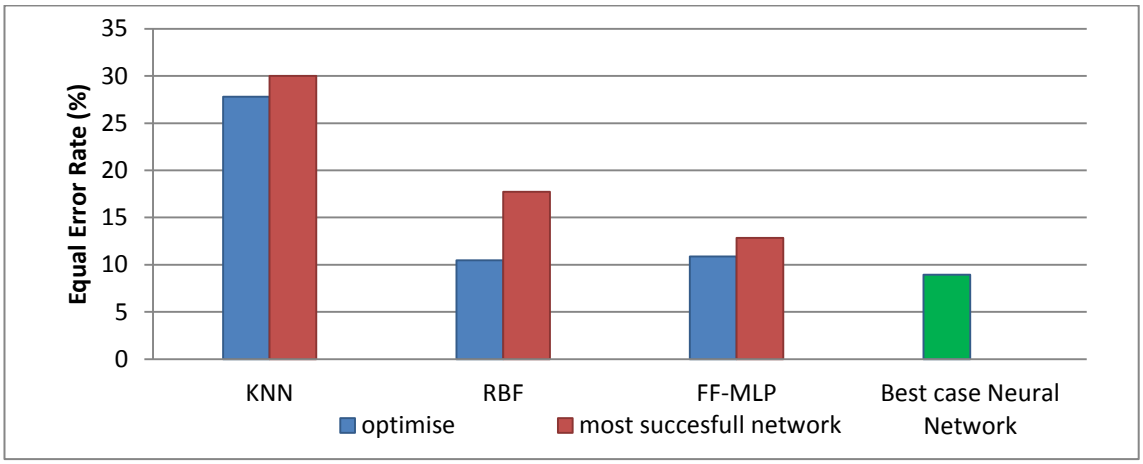


Figure 6.10: Performance comparisons of all techniques

From an analysis of the classification algorithms, it is clear that *User 27* experienced an EER more than 30% in every classifier. The reason could be due to the user's input varies too much from input to input for the linguistic profiling technique to prove useful. Therefore, it is important that any authentication system that implements a linguistic profiling technique must consider the small number of users that will experience high rate – a common conclusion from the analysis of behavioural-based biometrics.

Based on the above three sets of results, the study demonstrated that the user can be discriminated by using a linguistic profiling technique with a good level performance. The majority of all users' performance was within the bounds of an experimental-based behavioural biometric. However, individual user achieved optimal EER from different network configuration and classifiers. Therefore it is imperative to select the most optimal classifier for a particular user.

6.2.2 Keystroke Dynamic

This study looked into the use of a keystroke dynamic technique within the text-based multi-modal approach in order to increase the reliability of the authentication system. Therefore, this experiment focuses upon the ability to verify users by the way they type a text message. In previous research, Clarke and Furnell (2007) study sought to evaluate the performance of keystrokes based on the way a user types numerical and text messages. Building upon the findings of their study, the hold-time was used as discriminative characteristics and a feed-forward MLP neural network configuration was utilised as a classifier in this study. Further study on the feasibility of authenticating users using inter-keystroke latency was developed. A brief detail and the results of the experiments are presented in the following section.

6.2.2.1 Dataset

This experiment employed data provided by Clarke and Furnell (2006). This is due to this dataset containing keystroke dynamic data based on SMS text messages. More importantly, this database could be used to perform a dynamic based approach which is the most appropriate approach that can be used in almost any form of text messages. A total 30 participants were obtained and they entered a total of thirty text messages each, ten text messages over three sessions. The text messages were a

collection of quotes, lines from movies and typical SMS messages, where the only proviso was to ensure enough of the characters were repeated to enable classification. The length of the sequences varied, with an average length of 14 words per text message. A list of the messages can be found in Table 6.5.

#	Text Message
1	the quick brown fox jumped over the lazy cow
2	a lie gets halfway around the world before the truth has a chance to get its pants on
3	hi john cannot make lunch will phone you later
4	love cures people both the ones who give it and the ones who receive it
5	in a meeting at present will conference call you later
6	everything is funny as long as it is happening to somebody else
7	lack of will power has caused more failure than lack of intelligence or ability
8	fancy a couple of drinks tonight down the local
9	it is a rock tommy it does not have any vulnerable spots
10	i will meet you in town by the bus station at one
11	i love the smell of napalm in the morning
12	master yoda says i should be mindful of the future
13	a father is a guy who has snapshots in his wallet where his money used to be
14	a man knows when he is growing old because he begins to look like his father
15	all that stands between the graduate and the top of the ladder is the ladder
16	if computers get too powerful we can organise them into committees
17	a diplomat is a man who always remembers a women birthday but never remembers her age
18	i refuse to admit that i am more than fifty two even if that makes my children illegitimate
19	all of life great lessons present themselves again and again until mastered
20	do you want the job done right or do you want it done fast
21	a good marriage would be between a blind wife and a deaf husband
22	a man is a success if he gets up in the morning and gets to bed at night and in between he does what he wants to do
23	to educate a man in mind and not in morals is too educate a menace to society
24	advice is what we ask for when we already know the answer but wish we did not
25	immaturity is the incapacity to use ones intelligence without the guidance of another
26	a positive attitude may not solve all your problems but it will annoy enough people to make it worth the effort
27	a great obstacle to happiness is expecting too much happiness
28	consistency requires you to be as ignorant today as you were a year ago
29	adults are just children who earn money
30	children are natural mimics who act like their parents despite every effort to teach them good manners

Table 6.5: Text Message Dataset

6.2.2.2 Procedure

Due to keystroke dynamics utilising different biometric characteristics from previous studies, the investigation required the generation of new program scripts that had the ability to extract keystroke information from text input data and remove the outliers (a

standard pre-processing procedure for keystroke dynamics studies (Umphress and Williams, 1985; Joyce and Gupta, 1990; Obidat and Sadoun, 1997; Clarke and Furnell, 2006). However, the overall procedure of data splitting, classification and evaluation functions remained the same. The following scripts were generated to perform data extraction and outlier removal:

- Data Extraction function – Extracts the keystroke timing information from a number of text files and return input data for each of the thirty users.
- Outlier Removal function – Removes outlines from the input data based upon 'individual users' mean and standard deviation. An outlier being defined as an input whose value is outside three standard deviations of a user's mean. The figure of three standard deviations is based upon Gaussian or Normal distribution where 99% of a user's samples should theoretically reside within three standard deviations of a user's mean.

The sequence for each function utilised in this experiment is illustrated in Figure 6.11

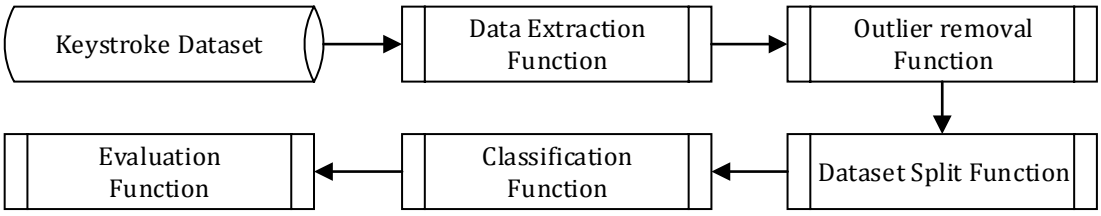


Figure 6.11: Keystroke dynamic system functions flow diagram

In the experiment, the following features were extracted from the dataset: inter-keystroke latency and hold time. According to the results from a number of experiments by Clarke and Furnell (2006), it was shown that using a hold time vector constructed from 5-character input provided the best classification performance. Therefore, the most popular recurring top five letters: E, T, A, O and N were being

utilised to create the hold-time dataset. Due to the type of data in this study was performed using a modified mobile phone handset (Nokia 5110), the hold time is somewhat different from the tradition definition of the word due to the keypad keys having to be pressed several times in order to type many of the characters (e.g. the letter 'b' requires the number 2 button to be pressed twice). Therefore, the hold time is defined as the time taken from when the first key press event has occurred until the final key press release. Further investigation to evaluate the feasibility of inter-key time has been conducted in this study. In this experiment, the latency between the release of a key and the pressing on the next key were extracted. The most recurrent top five pair of letters: 't–g', 'e–p' ,'e–m' , 'h–d' and 'a–m' were chosen to create the input vector. As a result, the final dataset contains 3510 hold-time data, 1080 inter-key time data and outliers were removed.

A number of analyses were undertaken using the FF-MLP neural network as it had demonstrated better performance in previous studies over other techniques (Bishop, 1995; Obaidat and Sadoun, 1997; Haykin, 1999; Cho *et al*, 2000; Clarke and Furnell, 2006). To perform classification, the FF MLP network parameter variables were modified following the previous successful study (Clarke, 2004) in order to achieve an optimal performance:

- Number of network layer – 2
- Number of neurons per layer - varied between 2 and 50
- Number of training epochs - 100

In the following experiment, user's data was divided into two datasets: 171 data samples were used for testing the classification's performance and the remaining data was utilised for training the neural network. The pattern classification tests were

performed with one user acting as the valid authorised user, whilst all the other users were acting as imposter and then rotating the role so that all users played the authorised user. Different network configurations were tested and the following results illustrate the most successful network configurations.

6.2.2.3 Performance of keystroke dynamics

The results of using individual keystroke characteristics to classify users are presented in Table 6.6. The results showed that the hold-time is the most effective measurement as it gave the lowest average EER of 20% with the best individual result achieving an EER of 8% by *User 14* and the worst result achieving an EER of 50% by *User 17* as illustrated in Figure 6.12. These results were achieved utilising a network configuration of 12 neurons in the first hidden layer and a single output neuron with 100 epochs. In contrast to the hold-time investigation, the inter-key characteristic provides little discriminative information to classify users with a fairly high average EER of 26%. These results were obtained using a network configuration of 12 neurons in the first hidden layer and a single output neuron with 100 epochs. However, there was the best case of a user achieving an EER of 3% by *User 13*, showing the ability to classify some users.

Classifier	**Average**			**Worst Case**				**Best Case**			
	FAR (%)	**FRR (%)**	**EER (%)**	**User**	**FAR (%)**	**FRR (%)**	**EER (%)**	**User**	**FAR (%)**	**FRR (%)**	**EER (%)**
Inter-Key	26.1	26.7	26.6	20	51.8	60.0	55.9	13	3.0	0	3.0
Hold–Time	20.4	21.4	20.8	14	41.6	60.0	50.7	17	0.6	0.0	0.6

Table 6.6: The classification results by using the most successful *FF-MLP* network

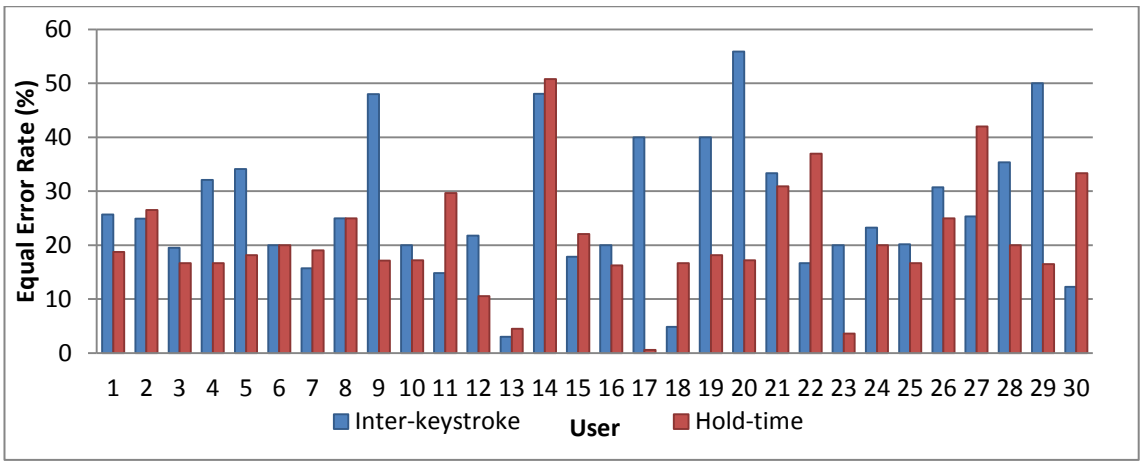


Figure 6.12: Individual results by using the best FF-MLP network

Combining or taking a user's best individual results for the FF-MLP configuration gave rise to an average EER of 18.7% and 16.5% for Inter-keystroke latency and hold-time respectively both of which are lower than previously obtained. An analysis of individual error rate of inter-keystroke latency characteristics showed that 70% of users achieved an EER of less than 20% as illustrated in Figure 6.13. For hold-time characteristics, 80% of users achieved an EER of less than 20%. The performance comparison of two characteristics is illustrated in Figure 6.14.

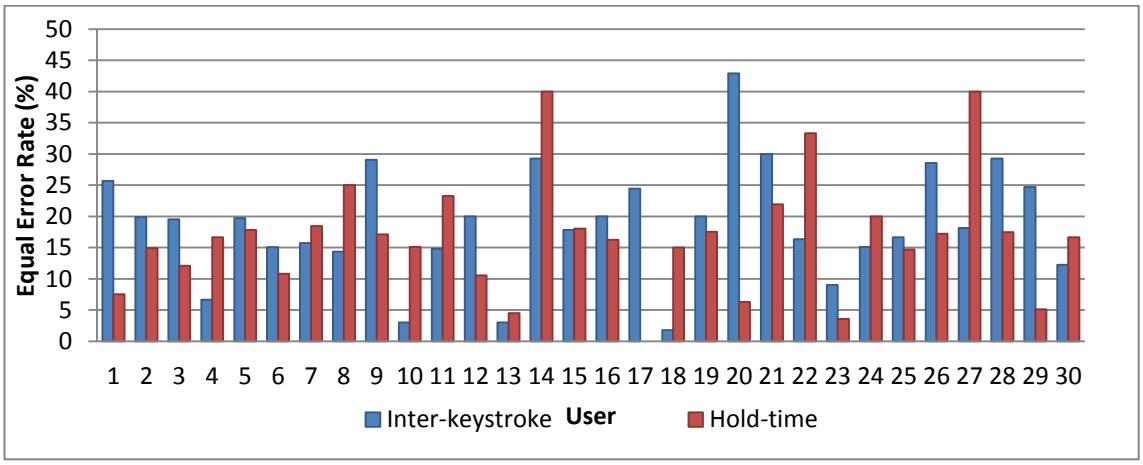


Figure 6.13: Individual optimal results for both keystroke chracteristics

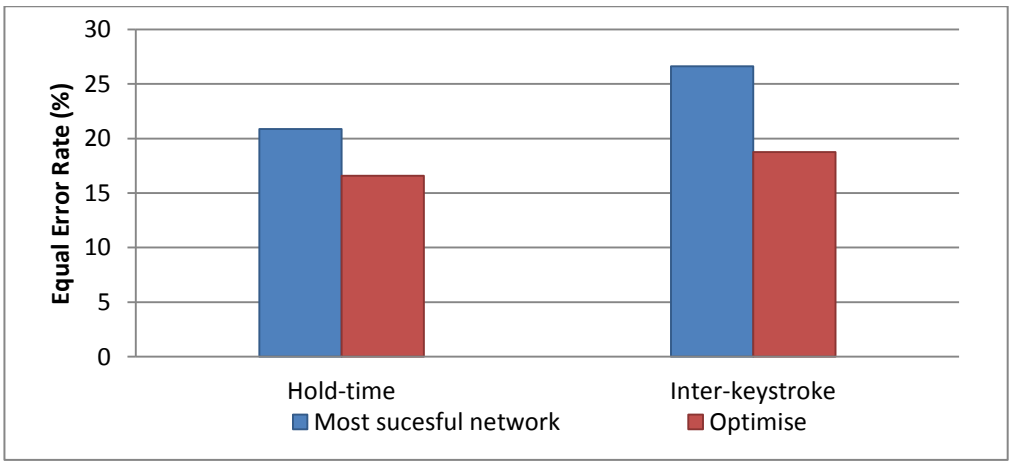


Figure 6.14: Performance comparisons for individual charateristics

The results showed that from the two traditionally used keystroke characteristics, the hold-time gave promising results in-line with previous studies undertaken (Clarke *et al*, 2004; Clarke and Furnell, 2006). As the hold-time in this study was defined by the first key press down until the last key release, this immediately increased the range of values available in the feature vector. Therefore, it arguably makes the classification process easier to discriminate between users. The study also showed that inter-keystroke characteristics did not perform very well in comparison. This may have been caused by the tactile interface. The mobile keyboard utilised in this study is very small, with a more restricted keystroke interface thus reduced distance between the keys (in comparison to other studies that have utilised full-sized QWERTY keyboards (Brown and Roger, 1993; Obaidat and Sadoun, 1997). In addition, the number of fingers utilised in typing are likely to be only one or two thumbs. Both of these factors restrict the typing dynamics, as the combination of the fingers in conjunction with the timing of the keystrokes and movement to achieve them, are reduced. This results in a smaller feature space for the inter-key characteristics to reside in and subsequently making it more difficult to distinguish between them.

The results presented however, must consider that this feasibility study was performed in controlled conditions, with users entering data repeatedly. In practice, it is possible that the variability of the user's input data could be larger as users might be walking or performing other tasks whilst they type. These factors would make classification more difficult. From this investigation, the hold time has proved useful as it avoided the problem of sampling and profiling the large number of digraph pair combinations that would have usually been required. However, authentication performance could be increased if the classification algorithm utilised a number of techniques to classify a user, capitalising on the specific content of the message. For instance, in a worst-case scenario the hold-time classification presented in this study could be used on messages with dynamic content utilising five of the most commonly regular used characteristics. For the best-case scenario, the hold-time characteristics could be used together with inter-key latency to perform authentication based on commonly recurring static-words such as "hello" or "c u later" in order to provide better classification of the user (Clarke and Furnell, 2006). Furthermore, user authentication can be performed by using either or both characteristics more than once within the same text message and the system responding on the combination of the results.

6.2.3 Behaviour profiling

This study will focus on verifying users by the way in which they utilise a text messaging application in order to provide authentication based on SMS communication usage. According to Li *et al* (2011) who conducted a number of experiments in order to investigate the performance of behaviour profiling based on general application, telephony and SMS applications usage. Therefore the SMS

application experiment in the Li *et al* (2011) study was reinvestigated. However, the number of participants in this experiment was increased from previous study. Based upon the finding result in their experiment, the classification algorithm utilised in this study is based upon an RBF neural network configuration.

6.2.3.1 Dataset

The study utilized a public database provided by the MIT Reality Mining project (Eagle *et al*, 2009). The data was collected by using automatic logging software preinstalled on participants' Nokia 6600 mobile phones. The MIT Reality Mining dataset contains a rich amount of information over a long period of time: 106 participants enrolled for the data collection process from September 2004 to June 2005; among these participants, 49 participants' text messaging activities were successfully logged. The experiment in this study utilised a subset of 30 participants whose text messaging activities occurred during the same period (from 24 /10/2004 - 21/11/2004) and their text message usage is summarised in Table 6.7. For example, *User 20* sent a text message to the anonymised telephone number 192 in cell 925 at 11:21 am on 17/11/2004. As not all participants started or finished the experiment at the same time, each user has a varied number of logs. In total, there were 275 unique telephone numbers being texted 1349 times. The maximum number of logs for any particular user is 149 and the minimum number is 8.

Number of Participants	30
Number of logs	1349
Number of unique telephone numbers	275
Information contained	Anonymised telephone number, date, time and location of texting

Table 6.7: Description of the final dataset of behaviour profiling

6.2.3.2 Procedure

The investigation required a new program script in order to extract a user's texting behaviour features record from the MIT Reality Mining dataset into the MATLAB programming environment and return each record entry for each individual user. However, methods of data normalisation, dataset splitting, classification and evaluation from previous studies are reused. The calling sequence for each function utilised in this experiment is illustrated in Figure 6.15:

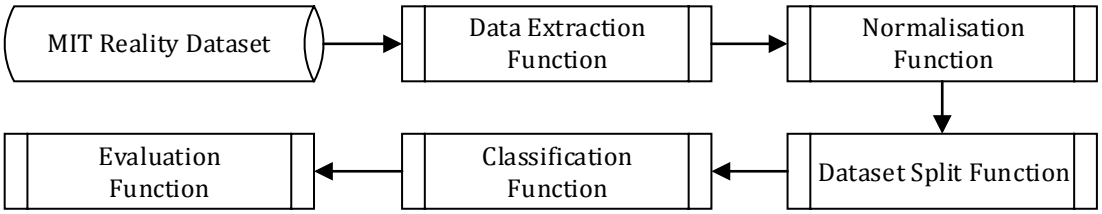


Figure 6.15: Behaviour profiling system functions flow diagram

Based upon prior research (Li *et al*, 2011), the following features were extracted: receiver's telephone number and location of texting. A number of analyses were undertaken, using a Radial Basis Function (RBF) neural network as it had performed the best in the prior study. From previous research (Li *et al*, 2011) two network parameter variables were modified as performing well to the classification problem:

- Spread – varied between 0.2 and 1.0
- Number of neurons – varied between 25 and 150

Similarly to the previous investigation, user's data was divided into two datasets: 171 of sample data were used for testing and the remaining was used for training user profile. The pattern classification tests were performed with one user acting as the valid authorised user, whilst all the other users were acting as imposters, with each user having the opportunity of acting as the authorised user.

6.2.3.3 Performance of Behaviour profiling

The most successful classification performance of using behaviour profiling to classify a user is shown in Table 6.8. The results illustrate that the text messaging application has significant potential to discriminate users with the overall performance EER of 8.7%. These results were achieved utilising a network configuration with a spread value of 0.5 and 150 neurons.

Average			Worst Case				Best Case			
FAR (%)	FRR (%)	EER (%)	User	FAR (%)	FRR (%)	EER (%)	User	FAR (%)	FRR (%)	EER (%)
9.2	8.4	9.2	25	64.2	50.0	50.0	1,7,8,14,20,22,26,27	0.0	0.0	0.0

Table 6.8: The classification results by using the most successful *RBF* network

Monitoring the individual error rates based on the most successful network configuration, the EER varies considerably between users as illustrated in Figure 6.16. The best result for an individual user was achieving an EER of 0% by 8 of the 30 users. Conversely, however, the worst result for individual performance was also very high with an EER of 50% by *User 25*. Further analysis illustrates that by utilising an identical classifier configuration, 24 of the users exhibit an EER of less than 20%.

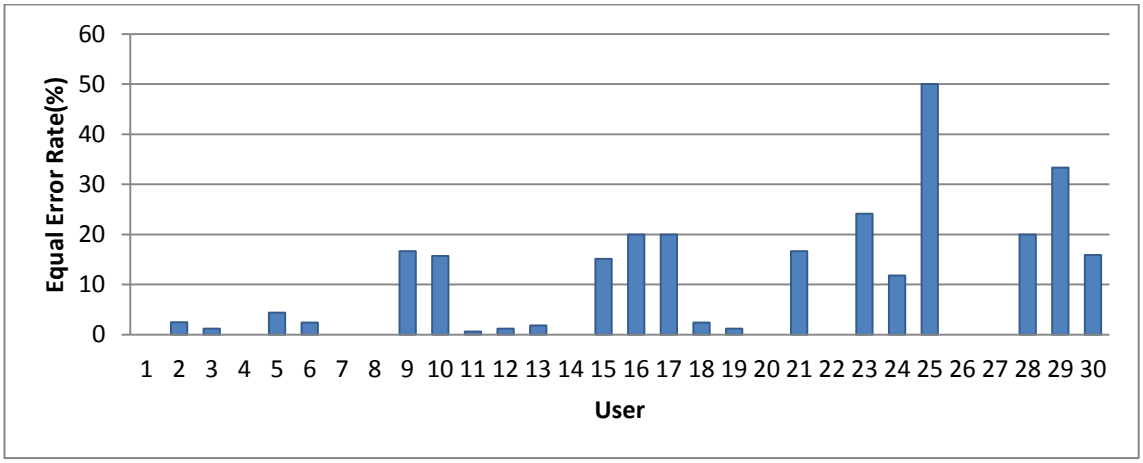


Figure 6.16: Individual results by using the best RBF network

An analysis of the individual error rate across all network configuration showed that individual users obtained the optimal results from different network performance. The best individual error rate for each user is illustrated in Figure 6.17. As can be seen, there has been significant improvement in the result with the overall performance achieving an EER of 3.7%. Analysis of individual results show that all the users achieved an EER of less than 20% with a further 83% of users achieving an EER of below 10%. In addition a total of 13 users obtained an EER of 0%.

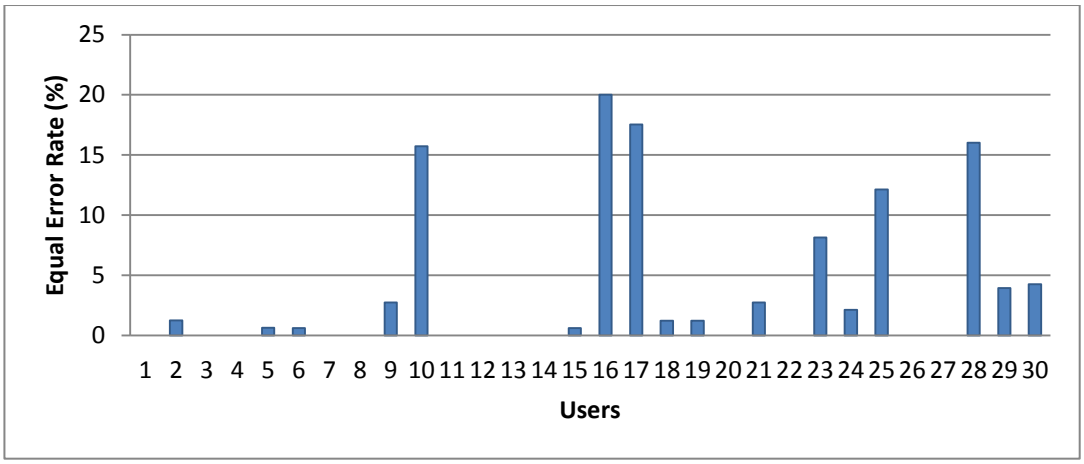


Figure 6.17: The optimal results for each user (RBF)

Through utilising only two features: (1) receiver's telephone number and (2) location of texting, users can be discriminated from each other. One of the reasons behind this could be that people only send text messages to very close contacts. Although the results suggested that the text messaging application usage can be utilised for classifying users with a good degree of performance, a minority of users achieved fairly high EERs. This is likely to be due to users tending to send messages to contacts and from different locations. As such, any authentication system that implements a behavioural profiling technique would also have to consider the small number of users

that will experience high an error rate in order to ensure both security and user convenience factors required by the overall system are met.

6.3 Performance of text-based multi-modal biometrics

On the basis of the aforementioned experimental findings, it is evident that a number of users remain that are unable to be correctly authenticated to a reasonable degree. This problem can be potentially alleviated through the combination of two or more biometric approaches in order to enhance the overall performance of biometrics. It is key however to ensure such combinations operate in a constructive rather than destructive manner. The use of multi-modal biometric approaches provides the ability of authentication systems to rely upon more than one biometric technique. For instance, those users for whom linguistic profiling is not a viable approach, keystroke dynamics or behavioural profiling could still provide with an adequate level of security. The use of more than one biometric technique also increases reliability due to the presence of multiple independent pieces of evidence, reducing the opportunity of forgery and circumvention. As a result, this study will utilise a novel combination of behaviour profiling, keystroke dynamics and linguistic profiling in order to evaluate the effectiveness of providing multi-modal authentication for text-based communications (such as texting, email, social networking and twitter).

In multi-modal biometrics, a requirement exists to combine some data, for instance, the classification results from multiple algorithms in order to enhance the overall performance of the system. From a literature review undertaken (and incorporated in Chapter 4), the combination or fusion method can occur effectively at any point within the biometric system: feature level, matching score level and decision level. Of all the fusion approaches, matching-level fusion is the most widely used technique in multi-

modal biometric systems (Ross *et al*, 2006). An additional advantage of fusion at this level is that a common fusion method can be utilised to create a reliable system and existing biometric systems do not need to be modified. This also offers up the opportunity for a highly modular and adaptable system. Individual and specialised classifiers can be utilised for the individual biometric modalities, with the fusion occurring immediately afterwards. Therefore, to enhance the overall performance of multi-modal biometrics, match-level fusion of the aforementioned behavioural biometric techniques was explored. Although many fusion techniques have been utilised in previous studies, this study will focus on the techniques that are well established in the literature (Snelick et al, 2003; Ross and Govindarajan, 2005; Jain et al, 2005). As a result, two different fusion methods: simple sum and weight average were utilised.

6.3.1 Dataset

To the best knowledge of the author, there is no multi-modal database available where the above three biometrics modalities are measured within the same individual. This is not an uncommon issue within this area of research. A standard practice employed is to combine biometric modalities from different datasets and create a virtual person. However, this is only justified in approaches that can be demonstrated to be independent. In this case, it is considered that the use of linguistic profiling, keystroke dynamics and behaviour profiling are independent and therefore their database can be combined. From an experimental perspective, these modalities belong to the same person, with *Virtual User A* being the combination of *User 1* from the linguistic dataset; *User 1* from the keystroke database and *User 1* from the behavioural profiling dataset. As a result, each individual user from the linguistic

profiling database was associated with an individual of the keystroke and behavioural profiling database, in order to create a virtual subject. Combining the users and samples in this manner, results in a virtual dataset whose size is equal to the smallest of the three individual modalities. Therefore, a final database consisted of 30 users, each user having his SMS messages, keystroke and text messaging activity logs data was created and utilised in this experiment.

6.3.2 Procedure

To evaluate the performance of the multi-modal approach, linguistic profiling, keystroke dynamics and behaviour profiling results were combined. The investigation of the study was developed using MATLAB. Additional program scripts were designed and created to investigate two fusion techniques (simple sum and weight average). The analytical process utilised in the experiment is illustrated in Figure 6.19. One significant challenge typically encountered with fusion-based approaches is the manipulation of the output from the individual biometric modality classifiers. However, as each of the three techniques was designed utilising the same family of classifiers, the output of the three approaches are already in the same form – enabling their direct use within the fusion algorithm.

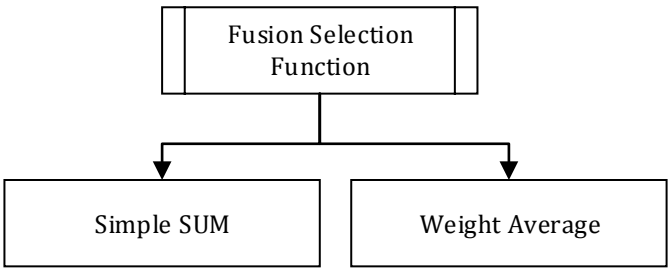


Figure 6.18: The two fusion methods being employed

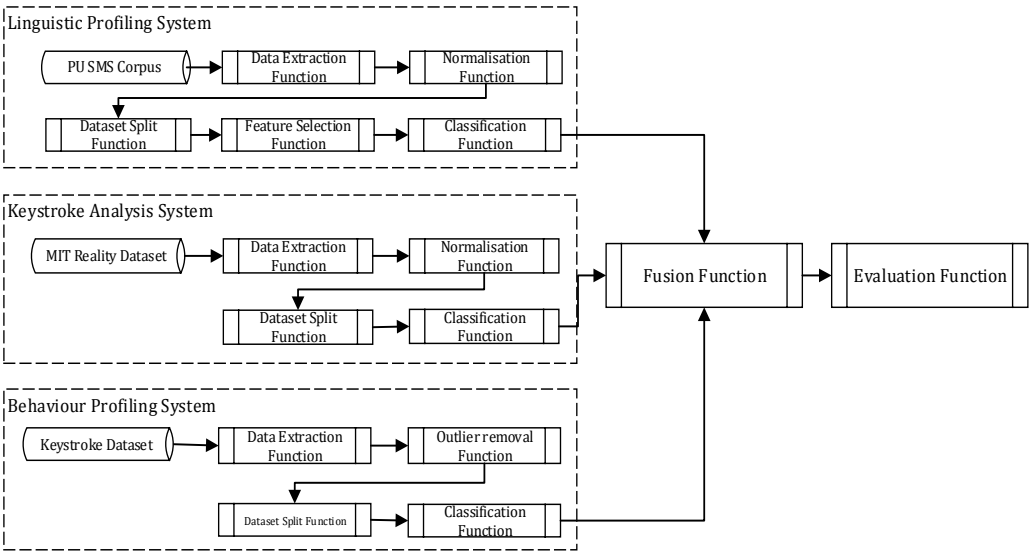


Figure 6.19: Multi-modal system functions flow diagram

To evaluate the experiment using the simple sum technique, the raw scores of each individual biometric system were simply added and rescaled into [0, 1] as shown below:

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

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 their weights are inversely proportional to the corresponding errors; the weights for techniques with a low EER are higher than those of techniques with a high EER.

$$Weight\ average = \frac{\sum_{i=1}^{N}(1-EER_i)}{\sum_{i=1}^{N} EER_i} \quad (2)$$

Where: i = the number of biometric system

N = the total number of biometric system

Similarly to previous studies, each user's data was divided into two datasets: 2/3 of the data was used to generate a user profile and the remaining 1/3 was used to evaluate the classification's performance. However, in order to combine the classification results from multiple biometrics, the size of each result is required to be exactly the same. Since the sizes of database utilised in this experiment varied between each technique, the 1/3 of testing data was calculated from the linguistic profiling database which has the smallest dataset and the remaining of data from each dataset was used as training data. The pattern classification tests were performed with one user acting as the valid authorised user, whilst all the other users were acting as imposter.

6.3.3 Performance of fusion approach

The aforementioned results of individual classification performance were combined using a simple sum and weight average approach. The results of all permutations are shown in Table 6.9 and performance comparisons of all techniques are illustrated in Figure 6.20.

Classifier	Average			Worst Case			Best Case		
	FAR (%)	FRR (%)	EER (%)	FAR (%)	FRR (%)	EER (%)	FAR (%)	FRR (%)	EER (%)
Fusion by sum									
LP+BP	5.3	4.4	5.5	27.9	33.3	30.6	0.0	0.0	0.0
KA+BP	6.7	5.1	6.2	19.9	20.0	20.0	0.0	0.0	0.0
LP+KA	11.7	10.2	11.2	50.0	40.0	45.0	0.0	0.0	0.0
BP+LP+KA	4.6	3.4	4.4	16.3	20.0	18.1	0.0	0.0	0.0
Fusion by weight average									
LP+BP	3.6	3.0	3.6	21.1	20.0	20.0	0.0	0.0	0.0
KA+BP	5.1	5.4	5.3	20.5	20.0	20.2	0.0	0.0	0.0
LP+KA	9.2	6.8	8.5	49.4	40.0	44.7	0.0	0.0	0.0
BP+LP+KA	3.4	2.8	3.3	18.7	20.0	19.3	0.0	0.0	0.0

Table 6.9: The classification results of fusion experiments

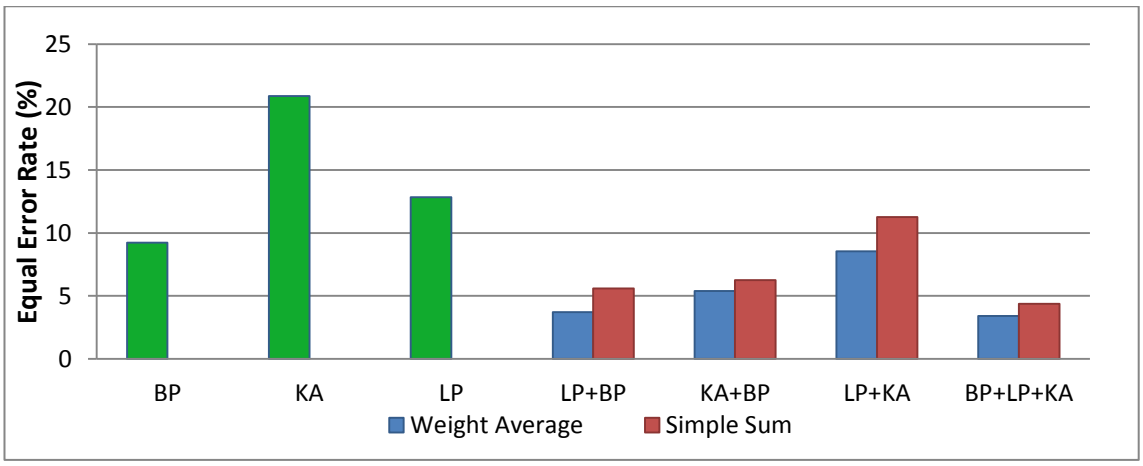


Figure 6.20: Overall performances of each technique

As can be seen from Table 6.9, the results clearly show that the use of both fusion methods enhances the performance significantly over the single-modal linguistic profiling, keystroke dynamics or behavioural profiling. Using weight average fusion of all biometric techniques gave the best performance with an overall EER 3.4%, as illustrated in Figure 6.20.

Further analysis based upon the individual error rate of each user using all biometric techniques proved very positive as all users obtained an EER of less than 20%. By using Weight Average fusion technique, 25 users achieved an EER of less than 5%.

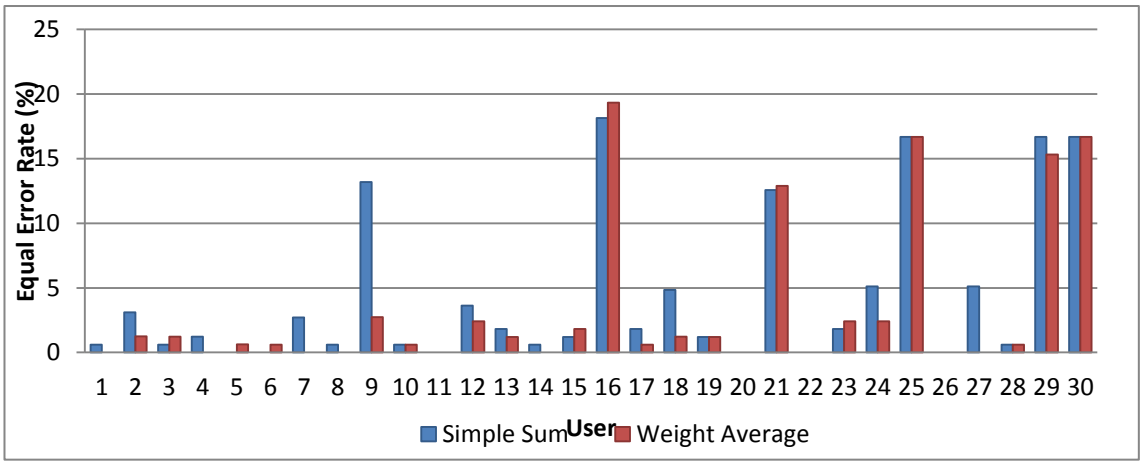


Figure 6.21: Individual results by using combination of three techniques

It would be expected that using a combination of three techniques would outperform using a combination of two techniques because the score of the sample is increased; however, this was not the case. As it was noticed from the results some users (i.e. *Users 10, 12 and 13)* obtained the best EER by using a combination of two techniques. This could be caused by the fact that the additional biometric technique gave a poor result and degraded the performance of the multi-modal system. Therefore, it is essential to ensure that the best optimal result is utilised in the multi-modal authentication system.

Monitoring the individual error rates, a user achieved the optimal performance by utilising different fusion methods. This indicates that the method chosen for fusion has an impact upon the resulting performance. An analysis of the individual error rate taken from the best performing fusion techniques, as illustrated in Figure 6.22, highlights the EER improving as the overall performance achieving an average EER 2.2% with the majority of users obtaining an EER of less than 5% and only 3 users obtaining an EER more than 10% (but less than 20%).

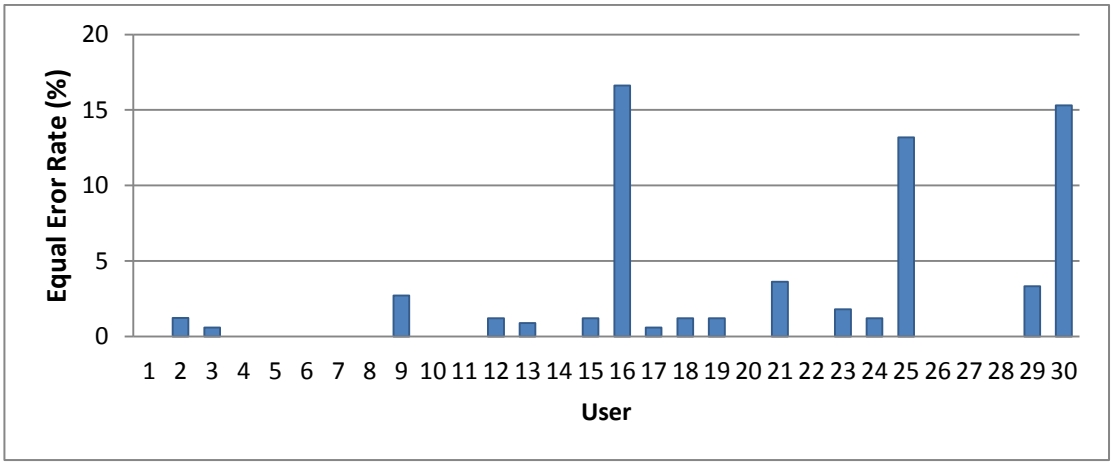


Figure 6.22: Optimal result for each user

Based upon the above results, the study demonstrated that a text-based multi-modal biometric approach is efficient for user authentication due to its low error rate. Individual user obtained the best error rate by using different combination techniques and fusion methods. As such it is essential that the authentication mechanism utilises the most appropriate combination technique and fusion method for particular user.

6.4 Conclusion

The study 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 an individual biometric technique can be used to discriminate users with relatively low error rates for a good proportion of participants. However, there are some users that experience fairly high error rates for each technique. To improve the performance of classification, multi-modal biometrics was investigated. In the fusion experiment, two fusion methods at matching score level were applied: Simple Sum and Weight Average. The results of the study show that multi-modal biometric systems outperform single modal biometric systems. By using the score-level fusion approach, the majority obtained an EER within the bounds of an experimental-based behavioural biometric with a number of users achieving an EER 0%. Based upon the success of these experimental results, the next chapter will focus upon designing a multi-modal biometric authentication architecture that could accommodate linguistic profiling, keystroke dynamics and behaviour profiling techniques in order to authenticate users in a transparent and continuous manner.

7 A Novel Framework for Multi-Modal Biometrics on Mobile Devices

Having established a successful basis for using linguistic profiling and multi-modal biometrics to classify users, it is imperative to consider the operational aspects that a practical system would require – particularly in designing appropriate mechanisms to take the raw authentication results and provide intelligent decisions that ensure the system is both secure and usable. This Chapter proposes a novel security framework that can enable linguistic profiling incorporation with other text-based biometric techniques in order to provide a transparent and continuous authentication mechanism for mobile devices. A detailed description of the system processes and procedures that enable such flexibility are presented.

7.1 Introduction

To increase security on mobile devices, authentication must be performed continuously, but with a view to maximising a user's convenience. Derived from the research undertaken, a successful authentication mechanism for mobile devices must work in the following manner:

- To increase the authentication security beyond that offered by the secret-knowledge based technique.
- To provide transparent non-intrusive authentication for the user (rather than intrusive) to maximise user's convenience.
- To provide continuous verification of the user, ensuring that the protection can be maintained throughout the duration of the device usage.

- To provide an authentication architecture that automatically works on all mobile devices regardless of hardware configuration, processing capability and network connectivity.
- To provide authentication architecture that can operate in three modes: standalone, cloud-based, within a pre-existing TAS.

These objectives have been achieved by utilising the combination of engines and processes within the novel Multi-modal Biometric framework. The framework employs a combination of secret knowledge and multiple transparent authentication techniques. Although intrusive techniques are used to ensure user's legitimacy, the majority of the authorised users will experience a transparent authentication phase. The result is an advanced authentication system that can provide transparent and continuous authentication to the user with minimum inconvenience.

7.2 A Novel Multi-modal Biometrics Framework

The proposed framework primarily verifies a user's identity based on their writing a text message. The system implements a multi-modal biometric topology utilising three transparent biometric techniques: linguistic profiling, keystroke dynamics and behaviour profiling. The framework is designed based upon a hybrid strategy, utilising a combination of multi-sample, multi-algorithmic and multi-modal approaches. By utilising the hybrid multi-modal biometric technique, the framework is able to make authentication decisions on a wider set of information. As a result, the framework is capable of providing continuous and non-intrusive user authentication with high security. Furthermore, the framework is highly modular, can utilise a wide-range of standardised biometrics and is able to take advantage of the different hardware configurations of mobile devices. Additionally, the framework is designed to operate in

both standalone and distributed modes in order to allow the framework to be useful for both non-wireless and wireless devices. In a standalone mode, the device performs all of the operations by itself. In a distributed mode, all of the system components were suggested to be stored in the server. As a client to server, a mobile device is used to provide input samples and the intrusive response interface (e.g. locking down the device, intrusive authentication to the user). The framework can also become a component for a TAS system. However, this thesis emphasises the standalone mode so that the completed framework architecture can be drawn. The proposed multi-modal biometric authentication framework is shown in Figure 7.1 and the mechanisms of the hybrid biometric system are shown in Figure 7.2.

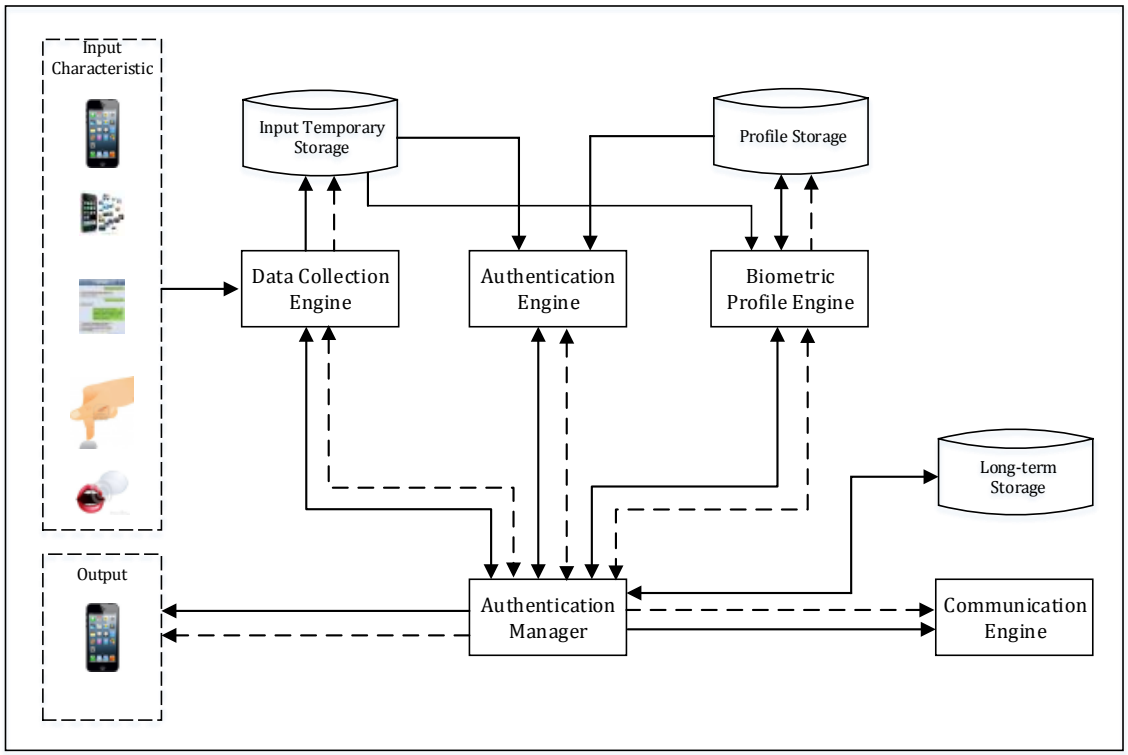


Figure 7.1: A Novel Multi-modal Biometric Framework

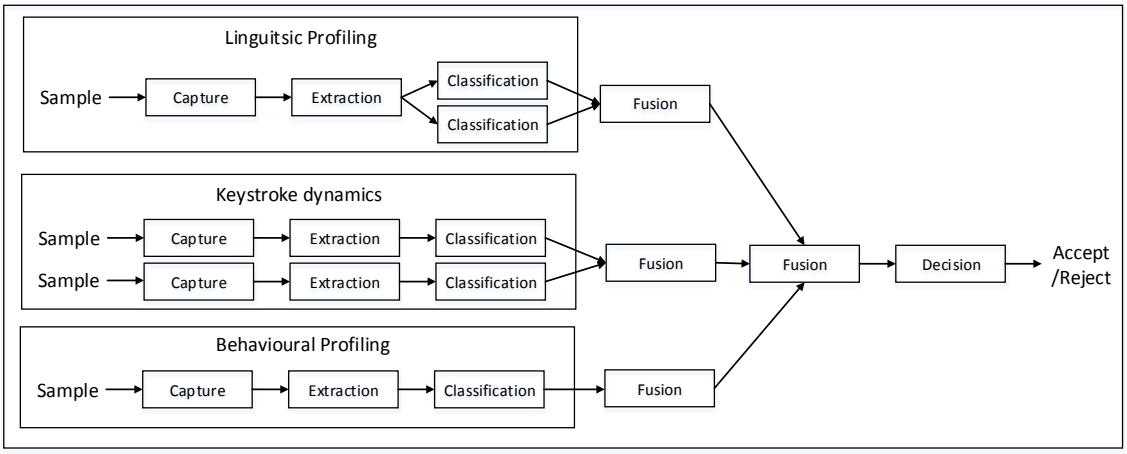


Figure 7.2: Hybrid biometric system

When a user utilises a text-based application on a mobile device, information about the user's typing, message writing style and the application usage are automatically collected by the Data Collection Engine and transformed into various biometric input samples. The Biometric Profile Engine utilises input text samples to generate various biometric templates that can be used in the classification process. An Authentication Engine compares the input samples with the biometric templates to determine the legitimacy of the user. Once the verification process is completed, the verification result is appropriately processed by the Authentication Manager. However, this will depend upon the mode in which the framework is operating. When the framework operates in standalone mode or cloud based mode, the Authentication Manager makes the decision by itself and, if appropriate, responds to the user with a request to authenticate. When the framework operates with a TAS, the Authentication Manager forwards the verification result to a TAS and the corresponding security mechanism makes any final decision accordingly. A detailed description of this process and the different operational models is presented in the following sections.

7.3 Processing Engines

The framework consists of a number of key components, including the Data Collection Engine, Biometric Profile Engine, Authentication Engine and Communication Engine. These perform various tasks such as collecting biometric data, generating user profiles, verifying a user's identity and communicating between mobile device and the Authentication Server.

7.3.1 Data Collection Engine

The primary role of the Data Collection Engine is to capture a user's input text sample. When a text-based application is utilised by a user, it is highly likely that all three linguistic profiling, keystroke dynamics and behavioural profiling agents can capture their biometric characteristics at the same time (i.e. synchronous mode). However, it is not possible to ensure that all of the biometric characteristics will be present. For example, keystroke dynamics can employ a pseudo dynamic approach that utilises the most frequently recurring top five letters (E, T, A, O, N). Unfortunately, if the text message does not contain all of these letters, there would be insufficient data to provide a keystroke dynamics sample. In this scenario, the system would have two of the three samples required for authentication. Conversely, it is also possible that the same word would appear more than once within the text and these words can be used as biometric samples. In this scenario, keystroke dynamics could utilise a multi-sample approach in order to provide stronger performance through the combination of matching scores from multiple samples of the same biometric.

The Data Collection Engine, as illustrate in Figure 7.3, contains a number of intelligent Data Collection Agents in order to capture the input data, one for each type of

biometric technique: linguistic profiling, keystroke dynamics, behavioural profiling and a system monitor. The process of sample acquisition is performed in a synchronous mode (i.e. at the same time). Apart from the system monitor agent, each of the agents continuously captures and logs their input samples in the background of a mobile device OS independently. The system agent captures the information of user activities in order to provide information on which services and applications the user is accessing so that the Authentication Manager can ensure the user has the correct authentication level required to access them.

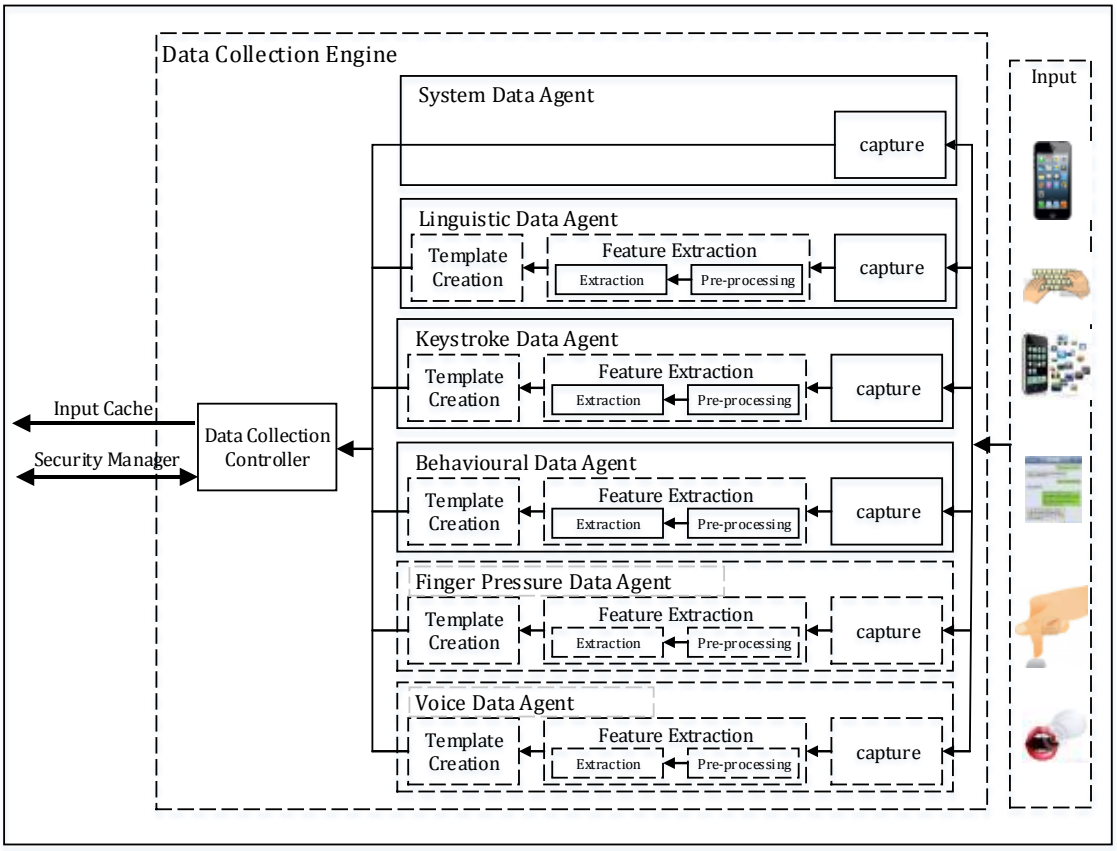


Figure 7.3: Data Collection Engine

Once the agents have captured a user's input data, the Data Collection Agent then proceeds to the next phase, pre-processing the input data and extracting all necessary biometric features from the input data. Dependent upon the individual biometric technique, the pre-processing task and further processing steps will be performed

differently. For linguistic profiling, the pre-processing for an SMS message involves splitting a text message into words. The pre-processing task of input data used by the keystroke dynamics technique involves calculating the duration time, performing outlier removal and normalisation of the timing vector. For the behaviour profiling technique, the pre-processing task of application usage data involves indexing the telephone number, and normalisation of the timing vector. The feature extraction process then extracts all potential features from the processed data and converts this data into feature vector that contains the concentrated biometric characteristics. The Template Creation process then transforms the feature vector into a sample template in a standard format. In the enrolment process, the sample template is used for creating a template, whereas during a verification process, the sample template is used for comparison with a user's profile. The pre-processing, feature extraction and template creation stages are optional in the client within the cloud-based system. However, it is recommended that the size of the template file that has to be sent across the network must be optimised to minimum in order to reduce the network traffic.

Various studies have shown that the performance of a biometric system is directly related to the quality of the biometric template and biometric samples (NIST, 2006; Hicklin and Khanna, 2006). It is therefore essential that the Data Collection Agent includes an ability to check the quality of the sample in order to maximise captured sample quality. For example, the Data Collection Agent for the linguistic profiling technique can use a predefined message length scaling threshold to check the quality of input text sample. The Data Collection Agent for the keystroke dynamics technique could have the ability to check the quality of the input sample by checking the duration

time information for indications that it is far higher than normal – suggesting the user was interrupted during data entry. However, how to measure quality is the subject of considerable discussion and research within the design of biometric modalities, rather than a wider architectural issue.

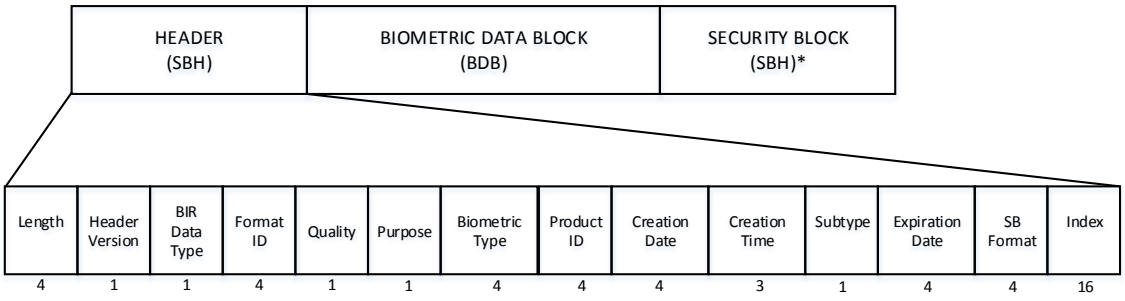
As the framework utilised more than one biometric technique to authenticate the user, it is essential to ensure that various biometric modalities from different vendors can work together and biometric templates can be communicated among the biometric components. As a result, the framework has specifically been designed to be compatible with international standards as follows:

- Data Interchange Standard
- Data Structure Standard
- Technical Interface Standard

The data interchange format standard called ISO/IEC 19794-1(2011) is utilised in order to provide the mechanism for structuring the biometric data into a meaningful form and representation of formats for the interchange of biometric data. The Data Interchange Format standard allows three forms of biometric data: raw data, intermediate data and the feature data that can be used by the matching stage directly. The use of standardized data interchange formats allows the biometric components to extract the biometric information required.

The data structure standard refers to the Common Biometric Exchange Formats Framework (CBEFF) and is utilised in order to define a data structure for the exchange of biometric data within the biometric system in a common way. The CBEFF specifies a

basic structure called Biometric Information Record (BIR) for exchanging biometric data. The structure of BIR consists of three parts, as illustrated in Figure 7.4 The Standard Biometric Header (SBH) provides information about data type and other properties of the Biometric Data Block (BDB) and security options. The meta-information stored on SBH allows the use of templates across different systems. The BDB contains the actual biometric data in the defined format. The Security Block (SB) provides information about algorithms used to secure the record. A BIR may have one or more BDBs so that multiple biometric samples can be packaged together. Through the use of a common data standard, the components within and between biometric systems are able to communicate using standardised records.



Source: Clarke, 2011

Figure 7.4: Biometric Information Record (BIR)

The Technical Interface standards called ISO/IEC 19784-1(2006) is utilised to define an Application Programming Interface (API) and controls interactions between biometric components. This enables the components to communicate, query and execute commands between each other. By utilising these standards, the biometric components within the framework are fully interoperable and as such the framework is able to provide more robust composite authentication and platform independence.

The biometric template is a representation of the unique characteristics of an individual person and as such it contains privacy-sensitive information. Therefore it is essential to ensure that the biometric information is protected from unauthorised user access. In order to secure the biometric data, the input text message utilised by the linguistic profiling technique will be transformed into an unreadable format such as hashed text message using a hash function.

After completing the template creation tasks, the Data Collection Agents will send the biometric template to the Data Collection Controller to store in the Input Temporary Storage. The template information consists of application name, types of captured input data, data and time of template creation, the quality of the sample and the location of the template. The actual size of the Input Temporary Storage is dependent on the type of input-data being collected. The general structure of the Input Temporary Storage consists of several tables, one for the master list of input samples and one for each operational mode in each biometric technique. A master table or Input Temporary database contains Sample ID, date and time that the input sample was captured, capturing application name that input sample captured from, technique, mode of operation, quality of the sample, and information on the data location of the biometric samples. Table 7.1 illustrates an example of the Input Temporary Database with a number of input sample records.

Sample ID	Date	Time	App Name	Technique	Mode	Sample Quality	Data Location
1	01/09/12	08.11	SMS	Keystroke	Dynamic:5	0.98	\\input_cache\\keystroke_Dynamic table
2	01/09/12	08.11	SMS	Keystroke	Dynamic:5	0.95	\\input_cache\\keystroke_Dynamic table
3	01/09/12	08.12	SMS	Linguistic	SMS	0.95	\\input_cache\\linguistic_SMS table
4	01/09/12	08.12	SMS	Behaviour	-	0.94	\\input_cache\\behaviour_SMS Application
5	01/09/12	09.00	Email	Keystroke	Dynamic:5	0.98	\\input_cache\\keystroke_Dynamic table
6	01/09/12	09.00	Email	Linguistic	Email	0.95	\\input_cache\\linguistic_SMS table
7	01/09/12	09.00	Email	Behaviour	-	0.94	\\input_cache\\behaviour_Email Application
:	:	:	:	:	:		:

Table 7.1: Input Temporary Database

7.3.2 Biometric Profile Engine

The Biometric Profile Engine's primarily role is to generate the various biometric profile templates that subsequently will be used by the Authentication Engine. This is achieved by utilising a combination of the user's historical data and a number of template generation algorithms. The Biometric Profile Engine contains a number of profile agents: one for each biometric technique and a multi-biometric profile agent shown in Figure 7.5. In a generic template creation process, the agent takes biometric samples into a template generation algorithm and outputs a unique template. The Profile Controller then takes both the biometric samples and biometric template and stores them within the Profile Storage element. The template is used by the Authentication Engine in the authentication process and the sample will be used again in re-generating or re-training the biometric template.

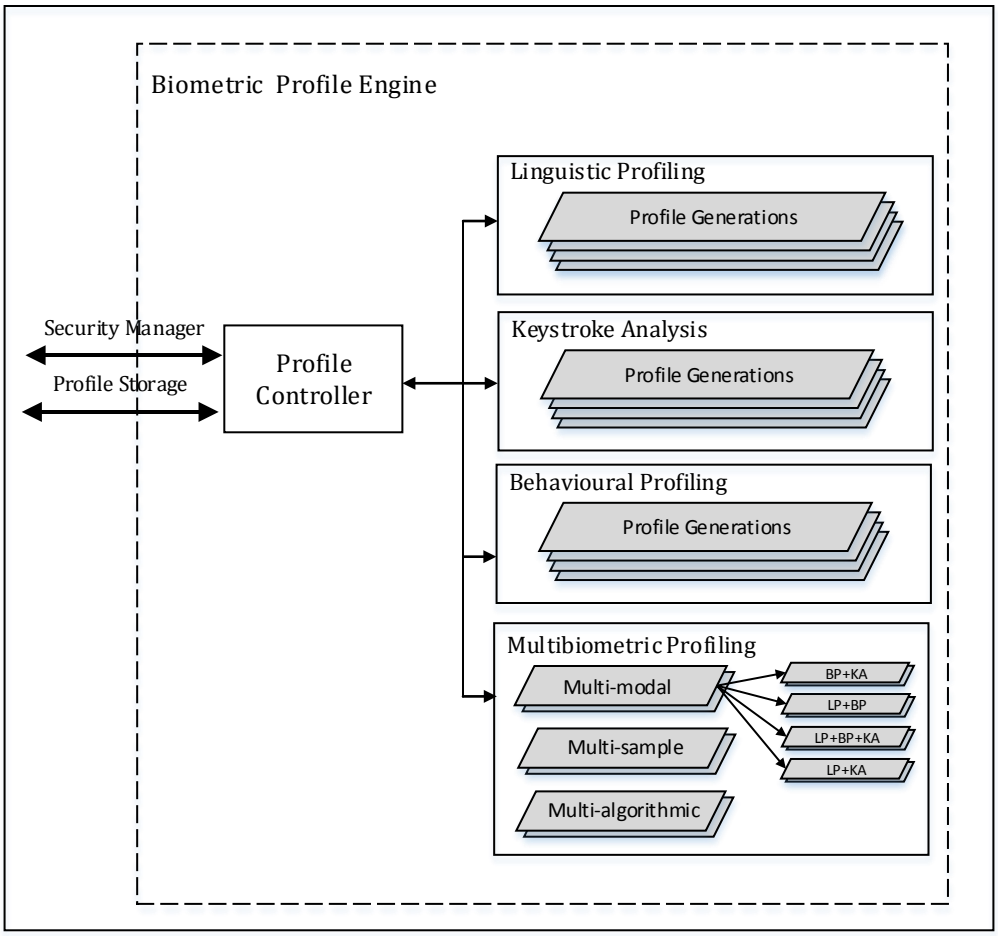


Figure 7.5: Biometric Profile Engines

The initial profile template generation process can be a complex task for any biometric technique as it would be difficult to request that the user provides all input data required to generate the templates at the devices registration stage. Also due to the nature of behavioural biometric techniques, they tend to change over time, the framework is designed to collect a user's input data over a period of time rather than gathering all information at the device registration stage. From the results and observations obtained through experiments in Chapter 6, a linguistic profiling system can provide a reasonable level of performance by utilising profiling data containing 10 SMS messages. Therefore, it suggests the framework should collect a minimum of 10 SMS messages after the initial device registration stage. During the enrolment period, the framework will automatically gather a user's text-based input by utilising its Data

Collection Engine in the background. The framework will not, initially, be able to provide any biometric security mechanism for the device due to the incomplete profile templates. Therefore, device security has to rely on other security methods available on the device such as PIN or Password authentication. However, if the framework operates with the TAS, it is possible that some biometric templates can be generated during the devices registration. For example, several copies of a user's facial image can be captured while the user is setting the cognitive responses and can be subsequently used to generate the template for use by facial recognition techniques.

The agent generates template profiles by utilising the user's input data and a series of template generation algorithms. An individual technique can utilise various algorithms to classify a user from different types of input such as the linguistic profiling techniques use of SMS or email messages or the keystroke dynamics use of static and dynamic input samples. Therefore the number of templates to generate is dependent upon the number of operational modes in each technique.

During the profiling template creation process, the user's input data is divided into two groups - one for training and one for testing. The profile agent utilises training data and an imposter's data for training the neural network. By testing the generated template with the testing data, the EER of the biometric system and the optimal threshold can be obtained. When the initial template generation task is completed, all of the generated templates and samples will be stored in Profile Storage. The storage contains several tables including the Biometric Profile Template and a series of tables that contain sample data from the authorised user for each biometric technique. The Biometric Profile Template table, as illustrated in Table 7.2 contains a list of biometric templates that have been generated, including re-train dates and the location of the

template file. This table will be accessed by the Biometric Profile Engine and Authentication Engine in order to re-train algorithms and perform authentication requests.

ID	Technique	Mode	Retrain Date	EER	Threshold	Template Location
1	keystroke	Dynamic:5	01/09/12	18%	0.75	\\ template\keystroke_s4
2	linguistic	SMS	30/08/12	14%	0.78	\\ template\linguistic_sms
3	linguistic	Email	29/08/12	10%	0.79	\\ template\linguistic_email
4	behaviour	SMS	27/08/12	8%	0.81	\\ template\behaviour_SMS
5	BP+KA		01/09/12	5.3%	0.83	\\ template\multimodal_BPKA
6	LP+BP+KA		01/09/12	1.7%	0.90	\\ template\multimodal_All
7	Multi-sample KA	Dynamic:5	01/09/12	10%	0.80	\\ template\multisample_KA
:	:	:	:	:	:	

Table 7.2: Biometric Profile Template

The imposter data is used together with the authorised user data in the training process for teaching the network the difference in input characteristics. As a result, the network is trained with input data belonging both to the authorised user and an imposter. The imposter's data can be automatically generated using traditional statistical tools by the framework itself (Jain et al., 2000). However, the quality of the artificial imposter data is highly dependent upon the authorised user's data and the generation method utilised. In the cloud-based topology, the imposter data can be obtained by randomly selecting sample data from other users in the database. Furthermore, it is possible to select the sample data from any other users that closely imitate the authorised user's input data so that the network is taught with the most difficult classification, as the imposter and authorised user data are very similar. In this way, the performance of the network can be increased.

As can be seen from Table 7.2, the Biometric Profile Template also contains threshold data. As demonstrated in Chapter 3, the threshold level determines the level of security provided by a biometric system (versus the convenience to the user). Setting

the threshold too high causes the system to exhibit a higher level of security but could also result in frequent rejection of the authorised user making the framework difficult to use. Lowering the threshold, the system allows more imposters to gain access but is more convenient for the authorised user. In practice, setting a threshold is a critical issue. Setting a predefined threshold (i.e. 0.85) might work well for one user but not the others. To this end, the framework implements a dynamic scaling threshold setting which allows users the flexibility to set the threshold based on their preference.

Security Level	Threshold level
5	High Security
4	Secure
3	Normal
2	Convenient
1	High convenience

Table 7.3: Security level

The user is able to select the level of security among five stages: high security, secure, normal, convenient and high convenience represented by a number 1, 2, 3, 4 and 5 on the scale respectively. The normal level will be used as the reference point for the threshold setting of the overall system. At the normal level, the system threshold for an individual user is set at the optimum point of the EER for the individual and multi-modal biometric techniques (not the system EER). When selecting a higher or lower security level, the threshold will be increased or decreased by x % (where x>0, i.e. 5) of the EER values for all techniques. During the device configuration, a user is required to determine which level of security they want for their devices and the normal setting is set by default. Each security level option is accompanied with descriptions to illustrate what these options mean, in order to aid the user in understanding the terms.

All the biometric samples used in the template generating process will be stored in the Sample Template Database within Profile Storage. The database consists of a series of

tables. One master table is present for each biometric technique that exists on the mobile device as well as one for each sub-category of a technique. The master table contains the location of the sample templates or sub-tables. For example, the master table for the linguistic profiling technique contains the file location of each biometric sub-category (as illustrated in Table 7.4) and a table that represents each mode containing the date and time of utilisation, status of input sample (whether the input sample has been used in the template creation process (1) or is a new input sample (2)) and a location of message data (see Table 7.5).

ID	Date	Time	Mode	Data Location
1	01/09/12	08.10	SMS	\\profile\linguistic\SMS Table
2	01/09/12	08.12	Email	\\profile\linguisic\email Table
3	01/09/12	10.25	Facebook	\\profile\linguisic\Facebook Table
:	:	:	:	:

Table 7.4: A Master Table for Linguistic Profiling

ID	Date	Time	Status	Message
1	30/08/12	21.30	1	0001 3454 0006
2	30/08/12	22.56	1	788A 751A 696C
3	01/09/12	08.12	1	DFCD 696C 3ABC
4	01/09/12	08.10	2	0001 7558 0003
5	01/09/12	10.25	2	76B1 8F8D 696C
:	:	:	:	:

Table 7.5: Linguistic Profiling: SMS Table

Index	Word
0001	Hello
0002	lol
0003	are
0004	No
0005	told
0006	2morow
0007	you
:	:

Table 7.6: Hashed Table

As mentioned in Chapter 3 (Section 3.3.2), behavioural biometric characteristics tend to change over time and under various situations so that the lifetime of each sample

will have an impact upon the template and the performance of the authentication algorithm. To maintain biometric templates with high level of quality, the templates must be re-generated using the most recent input samples. All the biometric samples stored in the Sample Template Profile can be utilised in a template re-generated process again. However, the age of the sample can result in high intra-template variability. As illustrated in Figure 7.6, the more recently captured samples should be included in the template profile while older samples (i.e. sample is out of date) should be removed from the profile otherwise natural variation between samples over time will result in an increasingly high FAR rate. The most common technique used to solve this problem is the periodical update of samples and re-generation of templates (i.e. 5 days). However, the time interval in which the template should be updated depends upon the user's behaviour. In the unusual case where no new samples have been added to the user's profile and all existing samples are out of date, the user is requested to perform enrolment again.

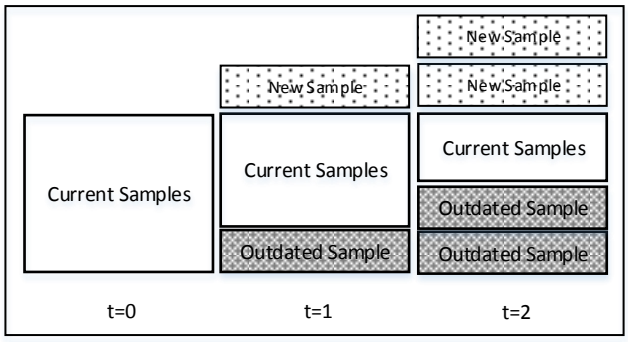


Figure 7.6: Sample Update Mechanism

In order to hold up-to-date sample data, all new successful input samples will be stored in the Sample Template Database within Profile Storage to be used in the next template re-generation phase. This approach requires the minimum user interaction and also avoids the re-enrolment process. When the template update requirement is

met, the Template Creation Agent will replace the oldest data sample with the new sample. In this way, the template will always be generated with an accurate profile containing the recent sample data. Once templates are generated, the Biometric Profile Template table will be updated accrodingly with the latest template and EER status of the system. The maximum time period that sample data remains valid will be limited by the individual user preference and by the storage capability of their mobile devices.

7.3.3 Authentication Engine

The main functionality of the Authentication Engine is to perform the user authentication process. The Authentication Engine has the ability to perform authentication for every permutation of inputs (that an enrolment template exists for) in order to ensure that the authentication can be performed even if all of the three biometrics samples are not presented. As a result, the engine is required to operate a total of 7 types of biometric system, with each type having the flexibility to incorporate multi-instance, multi-sample and multi-algorithmic:

- 3 uni-modal systems: linguistic profiling, keystroke dynamics and behaviour profiling
- A multi-modal system based upon linguistic profiling and keystroke dynamics
- A multi-modal system based upon behaviour profiling and keystroke dynamics
- A multi-modal system based upon linguistic profiling and behaviour profiling
- A multi-modal system based upon all three biometric systems

In order to perform authentication, the Authentication Engine contains an Authentication Controller and a number of Authentication Agents - one for each

biometric technique and one for the multi-biometric approach. When a verification process is required by the Authentication Manager, the Authentication Engine will perform the authentication by taking the input data from Input Temporary Database and the corresponding biometric template from Profile Storage, as illustrated in Figure 7.7.

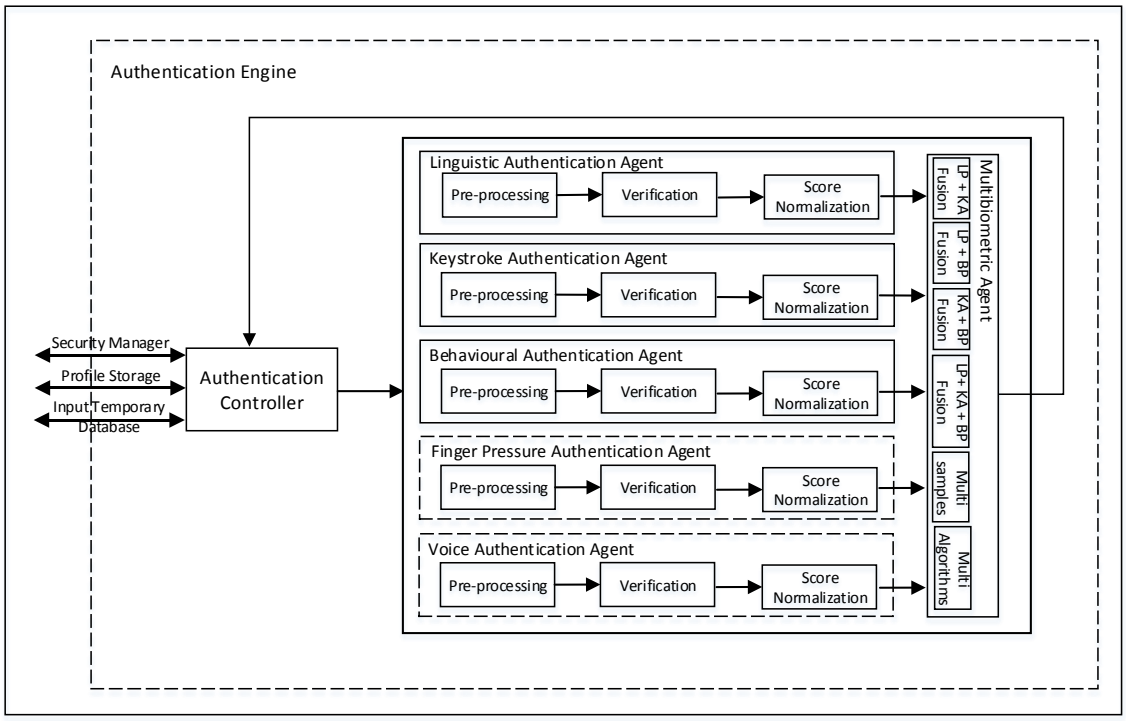


Figure 7.7: Authentication Engine

The Authentication Agent calculates a matching value fundamentally by comparing the similarity between the sample and biometric profile template resulting in a matching score. The results is then sent to the Authentication Controller to compare with the predefined threshold: if the result is less than the threshold, the sample (s) will be assumed to be valid; if the result exceeds the threshold, the sample (s) will be classified as invalid. The authentication decision is then reported to Authentication Manager.

Once the verification has been processed, in where the verification result indicates the sample(s) came from authorised user, the sample(s) are moved from the Input Temporary Database to the Sample Template Database within the Profile Storage repository to be used for profile generating; otherwise the input sample will be considered to be from an imposter and is deleted from the Input Temporary Storage Database. A multi-biometric authentication technique may produce a verification result that accepts the samples as coming from the authorised user even though the samples from one individual technique might nominally be rejected as coming from an imposter. Since the overall decision was that the samples come from the authorised user, the failed samples are deemed to be, in fact, from an authorised user and incorrectly failed. As such, these samples are added to the profile and are not deleted. In this way, the template re-training process can produce a more accurate profile that could provide better performance. However, all of the input samples will be deleted from the Input Temporary Storage Database if the overall verification result indicated that the samples are from an imposter.

7.3.4 Communication Engine

The Communication Engine provides a communication interface between the host and the client. Where the cloud-based model is in use, the mobile device is responsible for capturing a biometric sample of the user and sending it to the server (hosted in the cloud), the communication engine works as a bridge between the capture of the sample input and the framework. The role of the engine is to transfer information based upon 5 categories as illustrated in Figure 7.8.

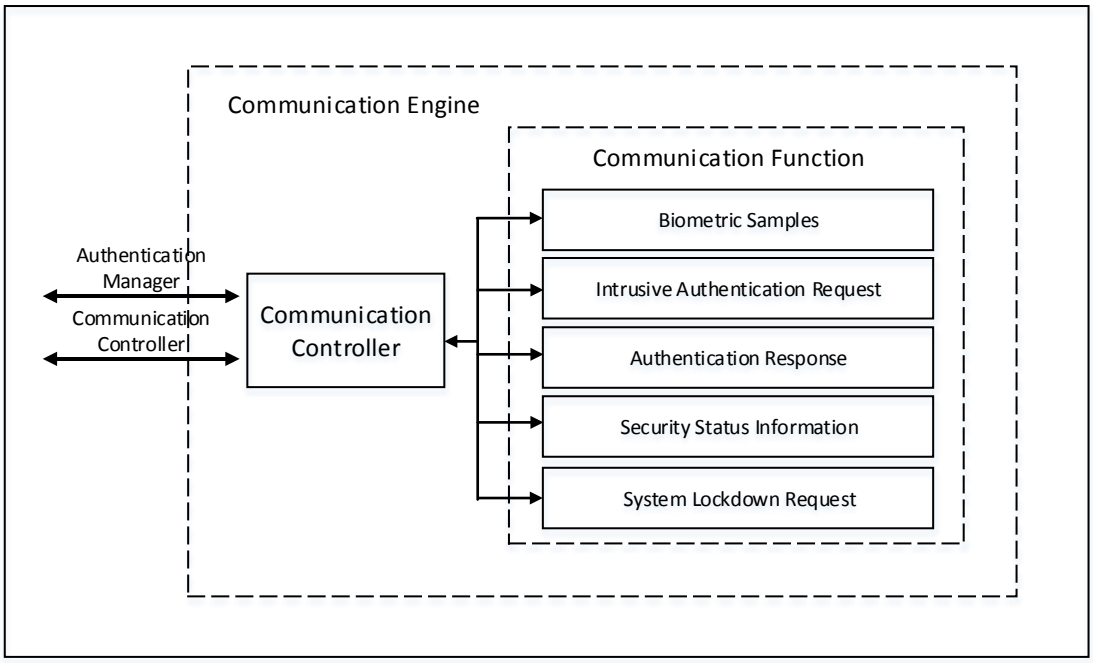


Figure 7.8: Communications Engine

The Communication Engine operates on both the host and client sides. From the client side, the Communication Engine is responsible for transfer of the biometric sample obtained from the Data Collection Engine for processing. The biometric samples will include text messages and numerical feature vectors. On the host side, the Communication Engine would be responsible for sending out the authentication result and the security status that enables the user to monitor the protection provided to the device. Furthermore, in cases where less computationally complex functions such as PIN or password authentications are performed within client side, the Communication Engine on the host side would be responsible for sending an intrusive authentication request. Additionally, in cases where the Authentication Manager locks down the device, the Communication Engine will send the lockdown code to an appropriate destination.

7.4 System Components

This section describes the remaining components of the framework including the Security Status level, Application Security Requirement and Authentication Asset database. All of these components provide information to assist the Authentication manager to maintain device security.

7.4.1 Security Status Element

The Security Status Element has two main functions: providing security information to the user and calculating the Security Status level (SS). The security information includes the status of the security based on the SS level and the authentication results (whether failed or passed) of previous authentication requests. This information provides a guideline to the user about how their device is utilised and therefore helps them to identify possible misuse.

The SS level is a numerical value in the range of 0 and +5 with 0 indicating a low security level and +5 indicating a protected security level[4]. When the device is initially switched on, the SS is set to the security level of 0. This SS level is a continuous measure increasing and decreasing over time and user usage sessions. The SS changes depending upon the performance of the authentication techniques as illustrated in Table 7.7.

Performance of Authentication (EER)	Increment/Decrement Value
0-5%	2
5-10%	1.5
10-15%	1
05-20%	0.5

Table 7.7: System Security Level Change

[4] The boundaries defined on the numerical scale are only provided as a suggestion. In practice, these values maybe redefined.

The time that has elapsed between authentication requests also affects the SS level. When a device with a high SS level is not used for a period of time, the framework will automatically decrease the SS level accordingly. In this way, the opportunities of an imposter accessing more sensitive information could be significantly reduced. The actual period is set depending upon individual user requirements. After each defined period of misuse the SS level decreases until the normal security level of 0 is reached.

The function of time provides benefit to mobile users in two aspects: convenience and security. For the frequent users, their SS level should remain at a high level most of the time because their SS level is updated regularly preventing the level dropping to 0. These users will experience more convenient access to information or applications that require a high level of security. For infrequent users, their SS level will degrade to 0 before they next utilise their mobile devices. These mobile devices are not left with a high SS level for a long time preventing abuse of valuable information. Although an opportunity for misuse could occur during the degrading period the time function will help to minimise it.

7.4.2 Application Security Requirement

According to Ledermuller and Clarke (2011), each mobile application category has a unique risk level dependent upon several criteria such as their connection types, how much information they are associated with, the nature of the information etc. Therefore the security requirement can be applied for each individual application based upon their risk level: the higher risk level an application has, the higher security should be required. To this end, each mobile application and service is given a security requirement level depending upon their risk level. As illustrated in Table 7.8, the Application Security Requirement Table is used to determine what the security level of

each application and service is. Applications or services that are associated with private information or expensive services would require a high level of security whereas the normal application would requires a low level of security. In order for a user to access any of the services listed they must have the SS level greater or equal to that specified. If not, the user will be required to authenticate themselves using an intrusive authentication technique in order to proceed with the service. If they are unable to obtain the security level required, the service will remain inaccessible to the user. The information contained within Application Security Requirement Table is used by the Authentication Manager to decide whether a user can access the application. Determining which services or what information the user is accessing at any particular moment is achieved by the Authentication Manager, via the System Monitor in the Data Collection Engine.

Applications/Services	Security Requirement Level
Bank Account	5
Email	4
Web browser	2
Maps &Navigation	0
Utilities	0
Photos	5
Contacts Detail	5
Voice Call	3
:	:

Table 7.8: Application Security Requirement Table

When the framework is installed on to the mobile device(s) each applications security requirement data will be created. All applications stored on the mobile device is listed in the table. The level of security for each application/service will be set to the normal level by default. However, the user will be recommended to consider the table and assign a higher security level to sensitive and expensive applications/services. The application security requirement data will be automatically updated after the user

installs a new application onto the mobile device. In order to prevent an imposter from setting a low level of security for important applications/services, the user is required to pass a secret knowledge based authentication (i.e. PIN or Password) in order to access this information.

7.4.3 Long-Term Database

The authentication mechanisms, including both biometric and secret knowledge based approaches, which are available for a particular mobile device is stored in the long-term database. This data is used by the Authentication Manager to determine which techniques are ready to be used (have a template generated) for authentication. This enables the Authentication Manager to decide upon which technique would be most appropriate for use for authentication. Although a large number of mobile devices are available, an authentication mechanism based upon text-based biometric techniques is possible to implement on all devices. The Authentication Asset would represent a list of all possible authentication techniques, as illustrated in Table 7.9. The table is populated as biometric samples are collected and templates are created.

ID	Technique	Mode	Template Create Date	EER	Library Location
1	PIN	-	29/08/12		\\library\PIN.dll
2	Cognitive	-	30/08/12		\\library\secret.dll
3	Linguistic	SMS	01/09/12	14%	\\library\LP.dll
4	Linguistic	Email	03/09/12	10%	\\library\LP.dll
5	Behaviour	SMS		8%	\\library\BP.dll
6	Behaviour	Email	03/09/12	8%	\\library\BP.dll
7	Keystroke	Static:4	-		\\library\KA.dll
8	Keystroke	Dynamic:5	30/08/12	18%	\\library\KA.dll
9	LP+KA	-	01/09/12	9.2%	\\library\LPKA.dll
10	LP+BP	-	01/09/12	3.6%	\\library\LPBP.dll
11	BP+KA	-	01/09/12	5.3%	\\library\BPKA.dll
12	LP+KA+BP	-	01/09/12	1.7%	\\library\All.dll
13	Linguistic	Multi-sample	-		\\library\LP.dll
:	:	:	:	:	:

Table 7.9: Authentication Assets table

Since the framework is designed using hybrid biometrics based on three techniques, the number of techniques is fixed to all possible combinations of these three techniques. However the table is able to grow in terms of the mode of operation as the framework captures more input data. The table also tracks a template generation date which indicates whether a valid template has been created and the location of the library file which contains the details of pre-processing, template generation and authentication for each authentication mechanism.

7.5 Authentication Manager

The Authentication Manager is the central controller of the framework and provides the "intelligence". The role of the Authentication Manager includes:

- Continuously authenticating the user's identity through their activities.
- Calculating and maintaining the Security Status level.
- Periodically requesting profile generation and re-generation.
- Making authentication requests and performing subsequent actions based on the authentication result.
- Updating and maintaining the Authentication Assets.

However, the key task of the Authentication Manager is to monitor the security level and make authentication decisions when the user requests access to an application. The Authentication Manager utilises two processing algorithms: the SS Level Automatic Update Algorithm and the Process Algorithms, in order to manage the balance between the security of the mobile device and user convenience. These processes have been designed based upon well-known study conducted by Clarke and Furnell (2007).

The Authentication Manager utilised the SS Level Automatic Update Algorithm in order to periodically update the SS level based on the user's input samples available, as illustrated in Figure 7.9.

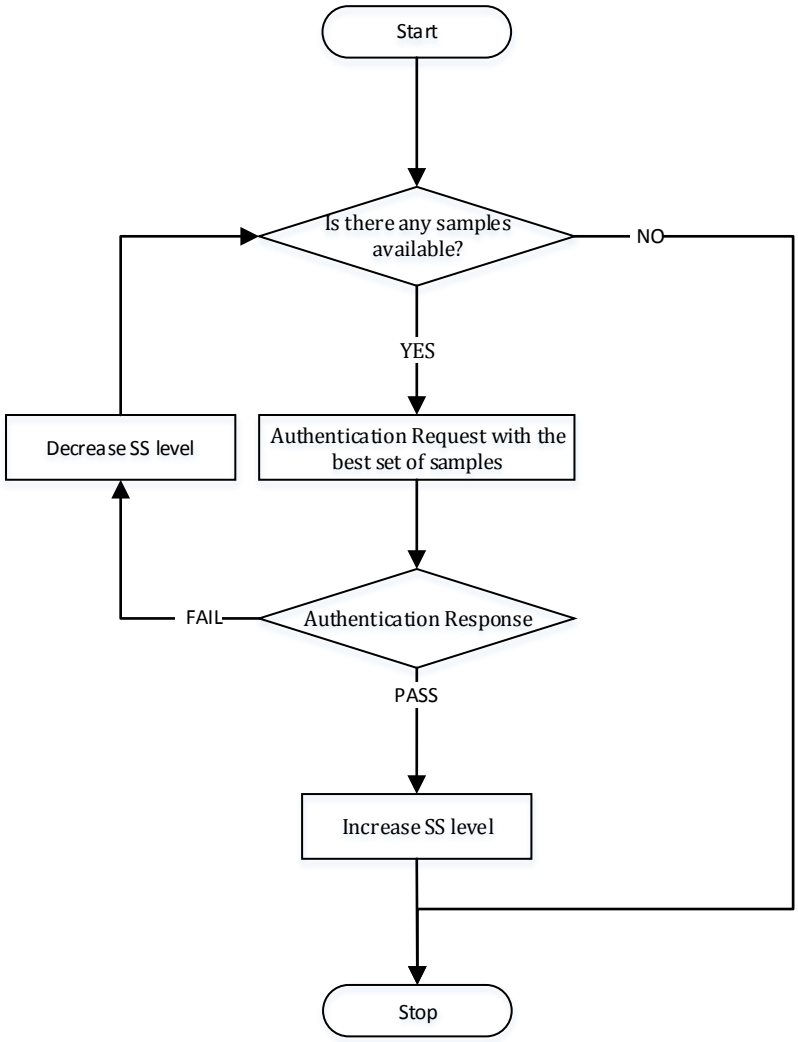


Figure 7.9: Authentication Manager: Automatic Update SS Level Algorithm

In general, the Authentication Manager periodically sends an authentication request to the Authentication Engine in order to update the SS level. The time interval in which the authentication should be requested is depends upon the user's preference (i.e. every 5 minutes). First of all, the Authentication Manager requires the Authentication Engine to perform authentication using the best set of the user's input samples from the last x minutes (i.e. 5 minutes) employing the strongest authentication technique.

In a case where no user's input data is presented, the Authentication Manager maintains the SS level at its latest updated value. However, if the Authentication Engine response with a pass then the Authentication Manager updates the SS level and go back to monitoring mode. If not, the Authentication Manager decrease the SS level and sends an authentication request again by using the next best set of user's input samples. The Authentication Manager will try three times to send an authentication request, every time with the next best available sample. The Authentication Manager updates the SS level based upon the authentication result. In a case where the updated SS level is less than 0, the Authentication Manager set the SS level back to 0. In this case the user will not be able to access high value application(s); only the applications that do not required security could be accessed. The process gives bias toward the user as they are given three non-intrusive chances to authenticate correctly and no intrusive authentication requests. This enables the system to minimise inconvenience to its user.

Should the user attempt to access applications that require a SS level greater than the current SS level, the Authentication Manager will utilise the Process Algorithm to check the legitimacy of the user. The flow of the Process Algorithm is illustrated in Figure 7.10

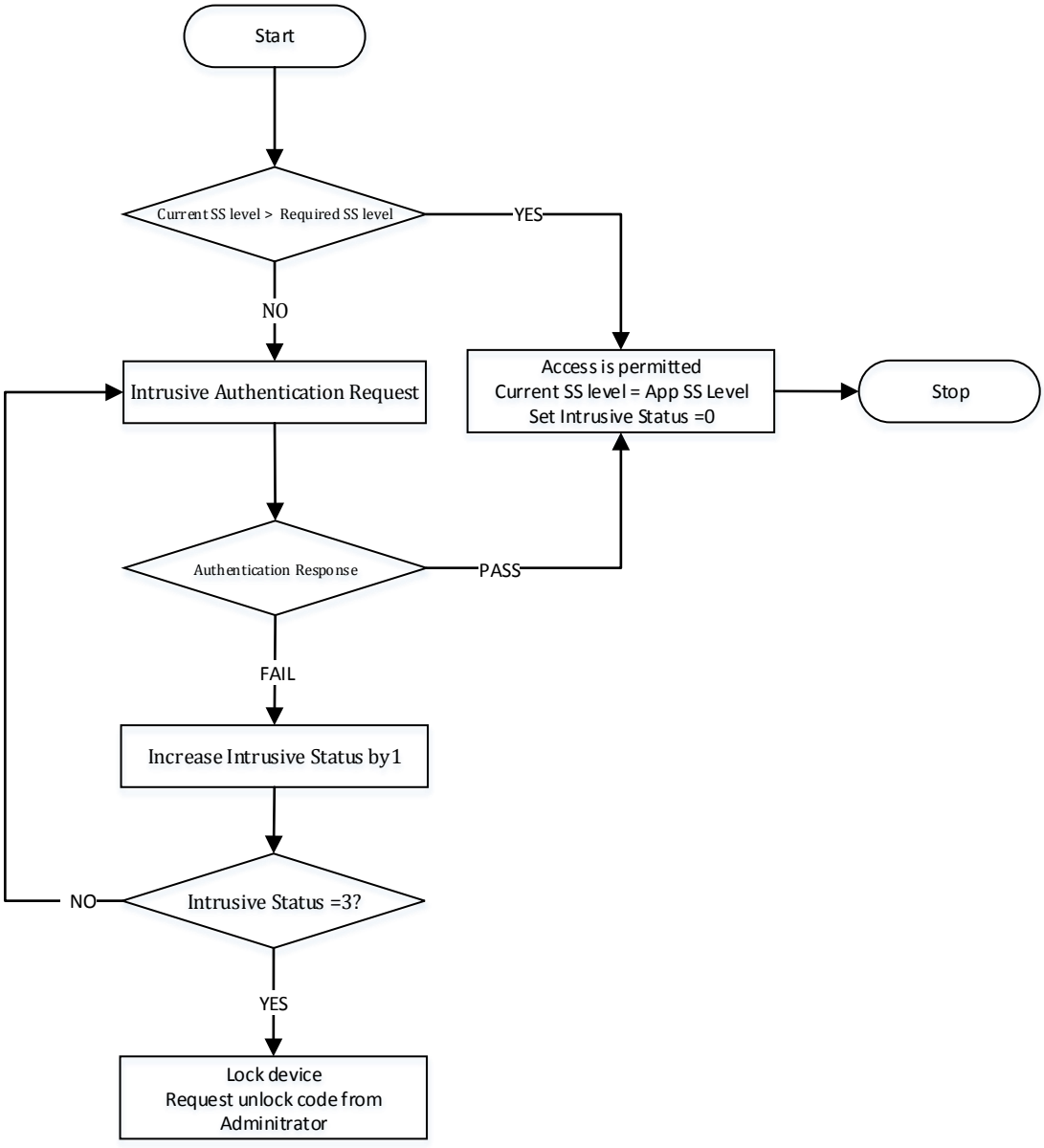


Figure 7.10: Authentication Manager: Process Algorithm

First of all, the Authentication Manager examines the up-to-date SS level and the SS level that the application requires and, if the user does not have the required SS level, the Authentication Manager will intrusively request the user to enter their PIN, password or cognitive question. If the user passes this, they will be grated to access the application and the SS level is set to the level required for that application. If the user enters an incorrect PIN number, the Authentication Manager sends an intrusive authentication request to the user again and the Intrusive Status will be increased by 1.

The system gives the user three chances to authenticate themselves. When the Intrusive Status equals 3, the device will be locked down and only the administrator security password can unlock it. When this code is entered, the Authentication Manager will set both the SS level and Intrusive Status to 0 and the user will be able to access the device once again.

In order to ensure that the devices do not continue to have a high SS level, the Authentication Manager will begin to decrease the SS level when no input sample is presented (i.e. the device is not used for a period). This can help to protect the sensitive information being access by an imposter.

It is envisaged that if the system is working correctly, the SS level should be high enough to permit automatic access for legitimate users and their experience of intrusive security challenges will be minimised. Also, the system should provide a low security level when an imposter is using the device and, eventually, the system should block the device if the imposter tries to access even low security level application(s).

The remaining duties of the Authentication Manager involved communication controls such as sending a template generation and retraining to the Biometric Profile Engine in order to update the biometric profile and subsequently updating the Authentication Assets table.

7.6 Evaluation of the Framework

This section evaluates the performance of the framework providing an illustration of the general system behaviour and performance. Due to the lack of public datasets, the effectiveness of the proposed framework was evaluated through a simulation approach using the MATLAB environment. The objective of this evaluation is to

examine the effectiveness of the framework in achieving both security and user convenience. The simulation process involves implementing a virtual user and applying the aforementioned proposed security system (i.e. the Process Algorithm) as illustrated in Figure 7.9 and Figure 7.10.

The simulation was evaluated by using a combination of intrusive and non-intrusive authentication techniques. The non-intrusive biometric techniques used in the experiment in Chapter6 were employed in the simulation. The performance of each technique was obtained from the results of Chapter 6**,** as illustrated in Table 7.10.

Technique ID	Biometric Technique	Mode	EER (%)
1	Linguistic Profiling	SMS	12.8
2	Keystroke Dynamics	Dynamic:5	20.8
3	Behaviour Profiling	Text message	9.2
4	LP+KA	Multi-modal	8.5
5	KA+BP	Multi-modal	5.3
6	LP+BP	Multi-modal	3.6
7	KA+LP+BP	Multi-modal	3.3
8	Intrusive Authentication		3.0

Table 7.10: Biometric Performance Rates

In order to create a simulation environment, a number of user interactions needed to be established. These are based upon 1. An authentication sample being captured due to a user interaction and 2. An application or service request being made. When these happen and the frequency was effectively randomized (to simulate a user use of the mobile device). The total number of interaction however was based upon the survey result (Locket, 2013) and showed that on average the mobile phone user check their devices 110 times a 12 hour day. In order to ensure the results are a fair reflection of user activity, the simulation was repeated with differing levels of user usage – to demonstrate how the model would perform under low medium and high usage conditions. The use of the mobile device is simulated using a flow of timeslot. Each

time slot represents a minute in real life time. In each time slot, it is assumed that, with a certain probability, a user can do one or both of the two actions available: enter an input sample or access an application. To simulate different type of users, the probability of these actions occurring will be changed. For the input sample action, a simulated user can enter 7 different types of input sample. Each type of input sample has the same probability of occurring 0.14 (being 1 in 7). The authentication result of each sample was simulated using probability based on the EER of each technique. For example, the probability of a false rejection of the authorised users input sample for linguistic profiling was 0.128.

Similarly, a simulated user can access different levels of applications with each application requiring a different level of security. In this simulation the most protected application associated with private information will require a security level of 5 and the normal application will require a security level of 0. Each type of application has the same probability of being accessing 0.16. The security system was simulated using the Authentication Manager. In this simulation, if the mobile device is not used for 10 minutes the system will start decreasing the SS level by 0.05 for every following minute, until the system is used again. The device is considered used if an input sample is presented to the system, or if the simulated user tries to access any application. Furthermore, every 10 minutes the SS level of the system is updated. This means that the system will take the samples it has and will try to authenticate the user, based on the EER of the best sample in the last 10 minutes. The SS level will be increased or decreased based on the success or failure of the authentication attempt, using the values shown Table 7.11.

Technique ID	Biometric Technique	Increment/Decrement value
1	Linguistic Profiling	1
2	Keystroke Dynamics	0.5
3	Behaviour Profiling	1.5
4	LP+KA	1.5
5	KA+BP	1.5
6	LP+BP	2
7	KA+LP+BP	2

Table 7.11: System Security Level Change in the Simulation System

An analysis of the process algorithm was conducted using several different scenarios with various configuration settings in order to examine the relationship between level of security and usability. Several factors such as the frequency of mobile usage and verification time will be modified to understand what effect they have upon the systems performance. The performance is presented in terms of the security level and convenience that the system can provide. The performance result of the simulation will be presented for both the authorised user and imposter perspectives. It is expected that the level of security provided to the authorised user will be high enough to permit automatic access and their experience of intrusive authentication will be minimised. On the other hand, the system must prevent the imposter using high value application(s) by decreasing the SS level to 0 and providing a number of intrusive authentications until the device is locked down. The simulation results for all scenarios are presented in section 7.6.1 below.

7.6.1 Simulation Results for Authorised User

To evaluate the performance of the security system provided to an Authorised user, three different usage levels (infrequent, moderate and frequent) will be investigated. The probability of a moderate user providing an input sample or accessing an application will set to 0.15 in order to simulate the usage of 110 applications in 12

hours. The configuration for each type of mobile user is illustrated in Table 7.12. The simulation simulated the use of the mobile phone for 12 hours or 720 minutes.

Types of User	Probability of input sample	Probability of access application
Infrequent	0.05	0.05
Moderate	0.15	0.15
Frequent	0.50	0.50

Table 7.12: Configuration for different types of user

In order to examine the role of the verification time, the configuration of time period between authentications will vary from 2 minutes, 3 minutes and 10 minutes. The result for all scenarios is represented using the average of running the simulation 10 times. The simulation results for an infrequent, moderate and frequent authorised user are presented in Table 7.13, Table 7.14 and Table 7.15 accordingly.

	Verification time 2 minutes		Verification time 3 minutes		Verification time 10 minutes	
Application level	#Application Request	#Intrusive Request	#Application Request	#Intrusive Request	#Application Request	#Intrusive Request
5	6.6	5.5	5.9	4.5	7.2	4.2
4	5.8	0.3	7.7	0.8	5.1	0.4
3	6.6	0.2	6.4	0.5	7.3	0.3
2	6.0	0.4	6.3	0.1	5.9	0.2
1	6.7	0.0	4.7	0.1	6.0	0.0
0	5.6	0.0	6.5	0.0	6.7	0.0
Average total	37.3	6.4	37.5	6.0	38.2	5.1

Table 7.13: Simulation Results for Infrequent User

	Verification time 2 minutes		Verification time 3 minutes		Verification time 10 minutes	
Application level	#Application Request	#Intrusive Request	#Application Request	#Intrusive Request	#Application Request	#Intrusive Request
5	17.8	9.2	16.8	7.4	16.2	1.5
4	19.4	1.0	19.3	1.0	17.9	0.5
3	18.0	0.3	19.3	0.6	20.0	0.2
2	17.3	0.0	17.9	0.3	16.8	0.3
1	18.3	0.0	18.3	0.2	19.5	0.2
0	17.4	0.0	17.8	0.0	19.3	0.0
Average total	108.2	10.5	109.4	9.5	109.7	2.7

Table 7.14: Simulation Results for Moderate User

	Verification time 2 minutes		Verification time 3 minutes		Verification time 10 minutes	
Application level	**#Application Request**	**#Intrusive Request**	**#Application Request**	**#Intrusive Request**	**#Application Request**	**#Intrusive Request**
5	58.5	5.5	58.4	5.1	60.0	1.50
4	60.7	1.7	62.9	1.4	60.1	0.50
3	59.2	0.4	61.9	0.5	61.3	0.30
2	58.9	0.2	58.3	0.1	59.6	0.10
1	58.1	0.2	57.3	0.2	55.2	0.00
0	59.1	0.0	59.5	0.0	57.6	0.00
Average total	354.5	8.0	358.3	7.3	353.8	2.40

Table 7.15: Simulation Results for Frequent User

Based upon simulation results, it demonstrates that the security system can provide high security while minimizing user inconvenience. Analysing the proportion between intrusive authentication request and application access permits an insight into how often the user experiences an intrusive authentication request when they access the application. Ideally, this proportion would be zero meaning that the user will not be required perform intrusive authentication request when they access an application. By calculating an average of the proportion across thee verification times, the results show that the infrequent user has been, on average, asked to authenticate 15% of the time that they accessed an application while the moderate and frequent user had averages of 2.5% and 1.6% respectively. The reason why the infrequent users experience of an intrusive request is greater than a frequent user is because the system decreases the SS level after the user does not use the mobile device for a period of time preventing abuse of high value applications. Therefore, when they wanted to access a high value application, they were required to perform an intrusive authentication again. An example of this case is showed in Figure 7.11. The example simulation simulated the device usage of an infrequent user and the time period of user authentication is every 10 minutes. The green triangle represents the user accessing an application. The number on the triangle indicates the security level of the application. The circles are the input samples used in the simulation and the y

coordinates specify the types of the input samples (1 linguistic profiling, 2 keystroke dynamics, 3 behaviour profiling, 4 linguistic profiling and keystroke dynamics, 5 keystroke dynamics and behaviour profiling, 6 linguistic profiling and behaviour profiling, 7 all of three techniques). The circles with the red border are the samples that were used in the authentication process when the authentication passed. The red crosses represent an intrusive authentication request made to the user. As can be seen from the graph, the user had experience an intrusive authentication request three times in this simulation. The first request occurred because the user wanted to access a high security level application (level 3) and the remaining requests happened because the security system starts decreasing the SS level after the mobile device was not activated for 10 minutes. Therefore, when the user tried to access a level 5 application again, they were requested to authenticate themselves using the intrusive authentication procedure.

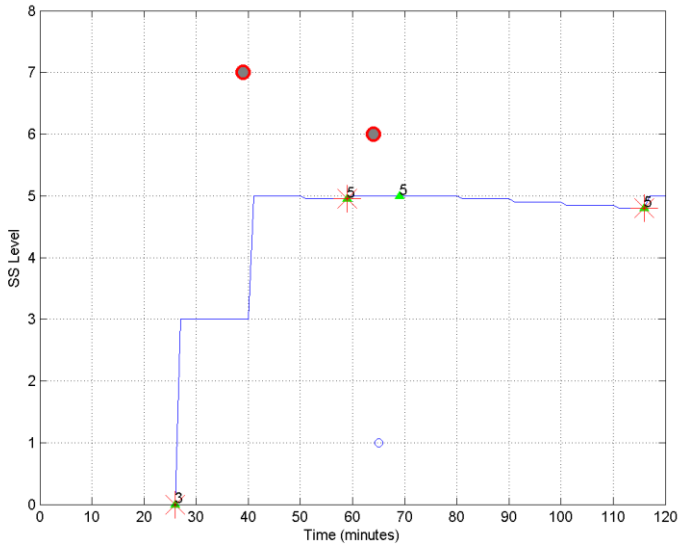


Figure 7.11: The use of mobile device simulation for infrequent user

The frequent user had the lowest experience of intrusive authentication request compare to other types of mobile user. However, the time period of authenticating

user does play an important role in providing convenience. Figure 7.12 and Figure 7.13 show examples of the device usage simulation for a frequent user with verification time of 2 and 10 minutes. The circles with a red border and yellow cross on top are the input samples used in the authentication process when the authentication failed.

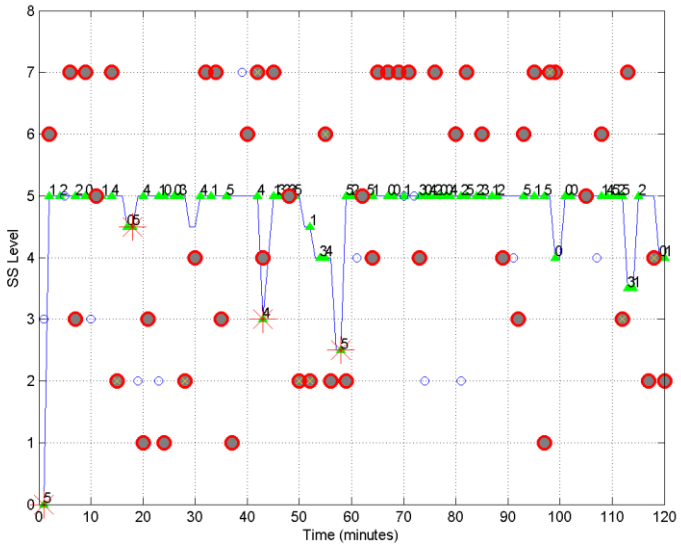


Figure 7.12: The use of mobile device simulation using verification time 2 minutes.

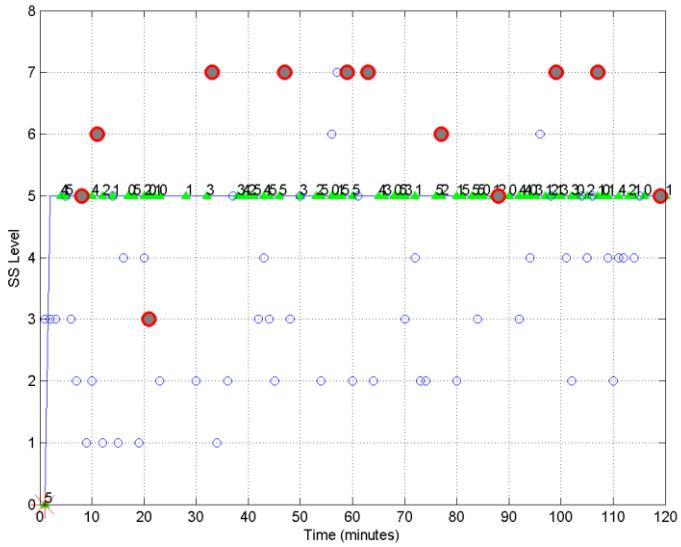


Figure 7.13: The use of mobile device simulation using verification time 10 minutes.

As can be seen from the result of frequent user, utilising the verification time of 2 or 3 minutes provides more number of intrusive authentication requests than verifying

user every 10 minutes. This is due to the decrease in the SS level based on the failed authentication result and the limited choice of authentication samples available to the authentication process. By using a slightly longer period between authentications of the user, the system has more chance of selecting the best input sample and authentication technique to keep the SS level high when authenticating the user.

7.6.2 Simulation Results for Imposter User

In order to examine the ability of the system security to prevent an imposter from using the mobile device, two scenarios were simulated: an imposter using a mobile device at the initial state (SS =0) and the imposter using a mobile device at the high level of security (SS=5). The imposter was simulated as a frequent user with the probability of accessing an application at 0.50 and the probability of entering an input sample at 0.50. The imposter has probability of being incorrectly identified as the authorised user when using an intrusive authentication of 0.03. The probability of being incorrectly identified as the authorised user using non-intrusive biometric techniques are identified based upon the EER of each authentication technique as demonstrated in Table 7.10. The results for both scenarios are illustrated in Table 7.16 and Table 7.17, respectively.

	Verification time 2 minutes		**Verification time 3 minutes**		**Verification time 10 minutes**	
Application level	**#Application Request**	**#Intrusive Request**	**#Application Request**	**#Intrusive Request**	**#Application Request**	**#Intrusive Request**
5	0.3	0.3	0.3	0.3	0.6	0.4
4	0.3	0.3	0.3	0.3	0.3	0.1
3	0.0	0.0	0.1	0.1	0.4	0.4
2	0.3	0.3	0.1	0.1	0.4	0.2
1	0.3	0.2	0.4	0.4	0.3	0.1
0	0.3	0.0	0.1	0.0	0.4	0.0
Average Total	0.1	1.1	1.3	1.2	2.4	1.2
Average Time use	2.3 minutes		2.6 minutes		5.0 minutes	

Table 7.16: Simulation results for imposter user start using device at SS= 0

The simulation results of the first scenario, showed that the security system works extremely well, blocking the imposter from using the mobile device after a few minutes. The majority of imposters never manage to access an application requiring a security level of more than 0. The reasons for this is that when the imposter tried to access an application that required a security level greater than 0, the system requested the imposter to authenticate themselves using an intrusive technique three times. Although, there was a case were the imposter passed the intrusive technique (the probability is 0.8732(0.03*0.97+0.03*0.97+0.97*0.03)) and the security system granted them permission to access the application. However, the security system decreased the SS level continuously as the imposter failed to authenticate themselves using non-intrusive authentication techniques. In this case, the imposter will be challenged again by an intrusive authentication request. An example of this case is shown in Figure 7.14. The red square indicates the user failed intrusive authentication three times and the system was blocked.

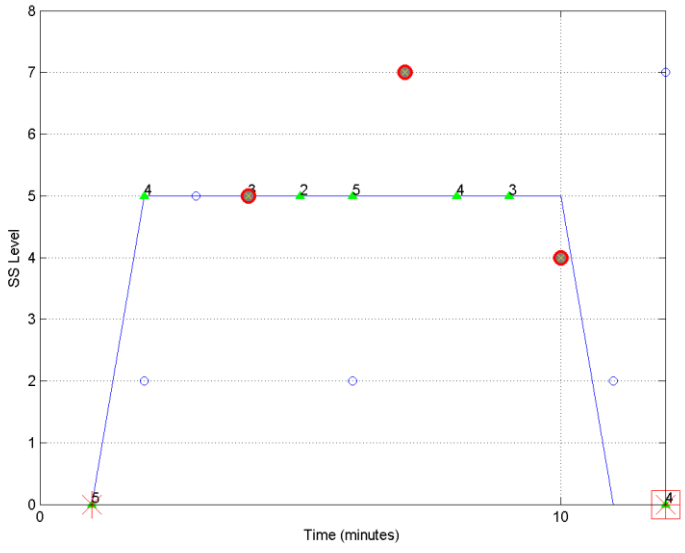


Figure 7.14: Simulation of imposter user start using device at SS= 0

The second scenario was simulated to examine the performance of the security system in preventing an imposter from accessing a high value application when the mobile device was left with SS level 5. As can be seen from Table 7.17, the simulation results showed that the security system works very well to prevent device misused by an imposter.

	Verification time 2 minutes		**Verification time 3 minutes**		**Verification time 10 minutes**	
Application level	**#Application Request**	**#Intrusive Request**	**#Application Request**	**#Intrusive Request**	**#Application Request**	**#Intrusive Request**
5	0.3	0.2	1.1	0.8	1.0	0.2
4	0.5	0.5	0.4	0.1	1.8	0.6
3	0.3	0.2	0.3	0.0	1.5	0.0
2	0.7	0.1	0.7	0.1	1.3	0.1
1	0.6	0.0	0.6	0.0	1.9	0.1
0	1.0	0.0	0.9	0.0	1.0	0.0
Average Total	2.8	1.0	4.0	1.0	8.5	1.0
Average Time use	8.2 minutes		8.2 minutes		14.7 minutes	

Table 7.17: Simulation results for imposter user start using device at SS= 5

Although the imposter begins with a big advantage in accessing the high security application, the system decreased the SS level as soon as the authentication request failed. However, the SS level will be decreased depending upon the input samples. An example of the simulation of an imposter using a mobile device with a high SS level is showed in Figure 7.15. Furthermore, by regularly authenticating the user, the system protected the mobile devices from misuse within a shorter period of time than when using periods of time between user authentications.

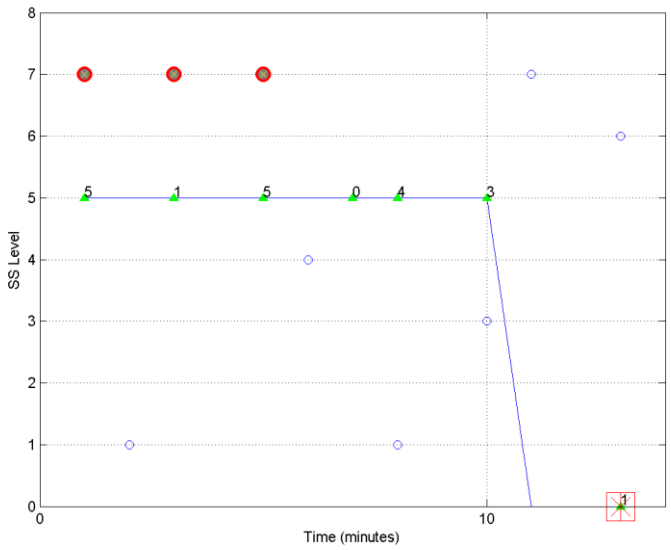


Figure 7.15: Simulation of imposter user start using device at SS= 5

Based upon the simulation results from both authorised and imposter scenarios, it can be seen that the verification time does play an important role of providing security and user convenience. By regularly authenticating the user, the user will suffer more intrusive authentication requests but the system will be able to recognise an imposter in a relatively short period of time. On the other hand, the users will find the device more convenient to use with longer time periods between user authentications but the system will take longer to recognise an imposter and lock down the system.

Although the time period of decreasing SS level was not examined, it is expected that this could have an impact on the system. The infrequent user will experience less challenges from the intrusive authentication technique when the time period of the degradation function gets longer. However, the imposter will have more chance of accessing a high level application in cases where the device was initially left with a high level SS. In this simulation, a linear function is used to decrease the SS level but it is suggested that the function for degrading the SS level should be implemented using an exponential function as it decrease slowly at first and then more rapidly.

7.7 Conclusion

In this chapter, a novel multi-modal authentication framework which provides transparent and continuous protection for mobile devices has been designed and the detail of the system components and functionalities were also described. The system is designed using multi-model biometric techniques without any additional hardware. The users can benefit from the framework in terms of both device security and convenience of use. The framework is able to generate up-to-date user profiles by utilising a combination of auto-update and dynamic profiling techniques permitting more accurate authentication results to be obtained. By setting various security requirement levels for different applications/services based upon their risk, the framework is capable of controlling the impact on each application/service. The user is able to set the level of security for their device depending upon their requirements using a sliding threshold scale. The mechanism implements the SS function to monitor the level of system security, when the SS level is high, all mobile applications can be accessed; however when the SS level is low, only unprotected applications or services can be utilised. The simulation results clearly showed that the proposed authentication framework is able to provide continuous and transparent authentication to protect mobile devices.

8 Conclusions and Future Work

This chapter concludes the thesis by summarising the achievements of the research projects together with discussing the limitations of the research. The chapter then highlights the future research directions within the mobile device security area.

8.1 Achievements of the research

Overall, the research has achieved all the objectives initially set out in Chapter 1, with a series of experimental studies and simulations undertaken for the development of a multi-modal biometric technique. The full achievements are:

1. A full investigation of the impact of mobile devices in society (Chapter 2). By reviewing the popularity and development of mobile devices, the sensitive information that they possess, the numerous security threats that lead to misuse and the lack of protection that is offered, the need for an alternative and stronger authentication for mobile devices that can provide continuous protection is identified.
2. A comprehensive review of biometric authentication techniques (Chapter 3). By presenting a good background for biometric authentication mechanism and focusing base on those that are applicable to provide security on mobile devices. Only a subset of biometric techniques was found to be suitable to provide continuous authentication. Of those available, linguistic profiling, keystroke dynamics and behaviour profiling represented the most interesting proposal given the most widely used of text based communication on mobile device and the transparent fashion of obtaining the biometric samples. By utilise these techniques, authentication of the user can take place transparently

enabling users to be authenticated numerous times without convenience, as biometric sample are captured during a user's normal interaction with the mobile device.

3. A thorough study on the current state of the art in text-based behavioural biometric techniques (Chapter 4). Contribution of previous studies on linguistic profiling defined a scope and requirement for studying the feasibility of linguistic profiling for mobile devices. By highlighting the weakness of individual behavioural biometric techniques, the need to develop a multi-modal biometric system is determined to ensure that the security can be provide within a wide range of user interaction with mobile devices.
4. An experimental examination and identification of text-based features based on SMS texting messaging (Chapter 5). A feasibility study of using linguistic profiling to authenticate users was conducted on real SMS messages from the NUS and PU SMS corpuses. By analysing a fairly large linguistic features extracted from SMS messages using descriptive statistical method, a number of linguistic features can provide discriminative information for a successful classification. A series of preliminary tests of profiling techniques were then evaluated by employing an advanced neural network classification technique for a preliminary study and the study demonstrated the most effective solution (i.e. the dynamic approach) to create user profile.
5. A deep study into authenticating users based on individual and multi-modal biometric techniques (Chapter 6). The individual biometric techniques: linguistic profiling, keystroke dynamics and behaviour profiling techniques were investigated using neural network techniques. All of studies concluded successfully with promising results when compared to typical results achieved

by other biometric techniques. However, from an analysis of the expected performances of individual biometric techniques, it was evident that no single authentication technique would sufficient. Therefore, the use of multiple biometric techniques was then evaluated using fusion of classification results approach and the result demonstrated the improving of the performance. The use of multi-modal biometric technique would permit transparent authentication of users while they compose text messages

6. A complete novel mobile authentication framework by utilising multi-modal biometric techniques (Chapter 7). The proposed architecture continuously verifies user based on their composing, typing and utilising text message on mobile device. The framework utilised the SS level in order to maintain security. The SS level changes depending upon the verification result. By adding the SS level to the authentication process, the system is no longer based on a pass or fails response to the identity of a user, but a possibility level indicating the confidence the system has in the identity of the user. With high level system security, automatic access will be granted to users; while, with low level system security, users would have to verify themselves through intrusive authentication approach before they can access the device. By evaluating the framework through the simulation approach with a number of scenarios from both authorised and imposter perspectives, the results clearly showed that this mechanism is capable of providing a continuous and transparent confidence measure for the identity of the user.

A number of papers related to the research project have been published and presented in referees conferences (the papers are demonstrated in Appendix (G)).

Therefore, it is deemed that the research has made positive contributions to the mobile security domain and especially in the field of biometric verification.

8.2 Limitations of the research project

Although the objectives of the research project have been met, a number of limitations associated with the project can be identified. The key limitations of the research are summarised below:

1. A number of practical limitations exist with linguistic profiling study. The limited number of participants and input SMS data prevented a more thorough evaluation of the technique. In the studies presented as part of this body of work, participants sent a total of at least 15 messages, thereby limiting the scope of the topic discussion. However, in reality, mobile users communicate with a plethora of different contacts about a wide ranging number of topics; therefore the simulation experiment may not necessarily reflect the variation of data input that could be observed in practice.
2. The three biometric techniques (linguistic profiling, keystroke dynamics and behavioural profiling) incorporated into the experiments presented in this thesis utilised separate databases. Therefore, the captured data across the three aforementioned databases was not from the same user and therefore did not present a true reflection of real time mobile usage. However, from a statistical perspective, the data captured remained valid for the purposes of validating the proof-of-concept.
3. In order to maximise number of testing data for fusion technique at matching-level in multi-modal biometric experiment, the sample data contained a very limited number of testing data from linguistic profiling technique. As a result,

the performance of multi-modal biometric was not examined by utilising a larger testing data in order to obtain the true performance of multi-modal biometric with high confidence level.

8.3 Suggestions & Scope for Future Work

This research project has improved the domain of authentication for mobile devices. However, there are a number of areas of future work that could be carried out to advance upon this research and within the area of authentication on mobile devices. The details of suggestion are listed below:

1. Development of a universal data collection software package. This would allow the proposed framework to be deployed on real mobile devices. It is imperative that the data collection software is embedded to monitor and capture data in a transparent manner without any substantial degradation of mobile system performance.
2. Deployment a multi-modal biometrics framework prototype on real mobile devices. This permits a comprehensive evaluation of multi-modal biometric technique on live users interactions to be proceeded and real participant feedback to be corrected.
3. Further research and development of robust linguistic profiling techniques that can be trained on text from one domain and applied to texts of another domain. For example, it is possible to have e-mail messages for training and a SMS message for testing. Then it is envisaged that, the performance of linguistic profiling technique can be improved and the authentication can be continuously performed in a wider text domain.

4. Further research in identifying positive linguistic features. As demonstrated in literature review, a number of writing style and different types of features can be extracted from text and successful used for classification. It is also evidenced by the Chapter5 experiments that positive linguistic features can improve the classification result while other features could downgrade the classification result and also consume more computation power. By examining the uniqueness of writing style from a wider feature set such as all possible linguistic features including semantic features, the performance of linguistic profiling technique could be improved. Therefore, the process of identifying positive features is mission critical for a wider deployment of linguistic profiling technique.

8.4 The Future of Verification for Mobile Devices

With the rapid and on-going development of cellular network technology and mobile devices, more and more people rely on their devices to complete personal and business tasks on a daily basis. They use mobile devices to make telephone call, check email, surf internet, store private information, to name but a mere fraction of the functionality and application available. Employees in all industries are increasingly bringing their mobile devices in to the workplace and using them to access corporate network, company email, business applications and other highly sensitive information. As more mobile services are being developed and mobile hardware capability is being enhanced, such a trend will continue to exist. It is clearly that the need to prevent information from being accessed by illegitimate or unauthorized users is required.

Although many mechanisms currently exist for authenticating users, this research has emphasised the need for a robust and reliable security mechanism which should operate in a continuous and transparent manner to offer both the security and user convenience. To this end, this research has designed and developed a novel authentication framework utilised multi-modal biometric techniques which capable of providing continuous and user friendly verification of the user to offer enhanced security.

To conclude, verifying a mobile user's identity will be crucial in the near future due to the financial services the mobile device provides and the sensitive information it carries. It is envisaged these services and information could become the main motivation towards device misuse. In order to provide adequate protection on mobile devices, mechanisms will have to utilise multiple security techniques and operate in a continuous and user friendly fashion.

In conclusion, authentication on mobile devices will remain an increasingly important consideration for users due to the functionality of services and sensitive information on devices makes them desirable targets for misuse. The ability to perform authentication continuously and conveniently will be fundamental to the successful deployment of future mechanisms.

References

1. Abate, A. F., Nappi, M., Riccio, D. and Sabatino, G. (2007) “2D and 3D face reocgnition: A survey”, Pattern Recognition Letters, vol.28, pp 1885-1906
2. ABI Research (2012) “Smart, the Next Wave of Bluetooth”, available at: http://www.abiresearch.com/research/product/1013429-smart-the-next-wave-of-bluetooth/
3. Akhtar, Z., Fumera, G., Marcialis, G.L., and Roli, F. (2012) “Evaluation of multimodal biometric score fusion rules under spoof attacks”, Proceedings of the 5th International Conference on Biometrics, pp.402-407
4. Aol tech (2011) “Orange and Barclaycard launch 'Quick Tap' NFC mobile payments in the UK”, available at: http://www.engadget.com/2011/05/20/orange-and-barclaycard-launchquick-tap-nfc-mobile-payments-in/
5. Aupy, A. and Clarke, N.L. (2005) “User Authentication by service utilisation profiling”, Proceedings of ISOneWorld 2005, Las Vegas.
6. AussiePete (2008) “A Language Guide for Foreigners”, available at: http://www.aussiepete.com/2008/05/singlish-language-guide-for-foreigners.html
7. AxxonSoft (2011) “Face Recognition”, available at: http://www.axxonsoft.com/integrated_security_solutions/face_recognition/index.php?phrase_id=3032106
8. Bartholomew, D. (2008) “The Rhythm of Identity Management”, available at: http://www.baselinemag.com/c/a/Security/The-Rhythm-of-Identity-Management/
9. Basit, A. and Javed M. (2007) “Localization of iris in gray scale images using intensity gradient”, Optics and Lasers in Engineering, vol.45, no. 12,pp.1107-114
10. BBC news (2009) “Mobile phone ID fraud increases”, available at: http://news.bbc.co.uk/newsbeat/hi/technology/newsid_8242000/8242709.stm
11. Berg insight (2010) “GPS AND MOBILE HANDSETS”, available at:

http://www.berginsight.com/ReportPDF/Summary/bi-gps4-sum.pdf

12. Berg Insight (2011) “Mobile Money in Emerging Markets”, available at: http://www.berginsight.com/ReportPDF/ProductSheet/bi-mm1-ps.pdf
13. Bhattacharyya, D., Ranjan, R., Alisherov, A.F. and Choi, M. (2009) “Biometric Authentication: A Review”, International Journal of u- and e- Service, Science and Technology, vol.2, no. 3, September, 2009
14. Biggio, B., Akhtar, Z., Fumera, G., Marcialis, G.L. and Roli, F. (2012) “Security evaluation of biometric authentication systems under real spoofing attacks”, IET Biometrics, vol.1, no.1, pp.11–24
15. Biometric Newsportal (2011) “Retina biometrics”, available at: http://www.biometricnewsportal.com/retina_biometrics.asp
16. Bishop, M. (1995) “Neural Networks for Pattern Classification”, Oxford University Press, New York, 1995
17. Boyd, J.E. and Little, J.J. (2005) “Biometric gait recognition”, LCNS: Advanced studies in Biometrics: Summer School on Biometrics, pp.19-42
18. Boukerche, A. and Nitare, M.S.M.A. (2002) “Behavior-Based Intrusion Detection in Mobile Phone Systems”, Journal of Parallel and Distributed Computing, vol.62, pp.1476-1490
19. Bours, P. and Shrestha, R. (2010) “Eigensteps: A giant leap for gait recognition”, In 2nd International workshop on security and communication networks (IWSCN), Karlstad, Sweden.
20. Brown, M. and Rogers, J. (1993). “User Identification via Keystroke Characteristics of Typed Names using Neural Networks”, International Journal of Man-Machine Studies, vol.39, pp.999-1014
21. BTopenzone (2011) “What’s a Wi-Fi hotspot?”, available at: http://www.btopenzone.com/help/whats-a-hotspot/index.jsp

22. Buchoux, A. and Clarke N.L. (2008) “Deployment of Keystroke Analysis on a Smartphone”, Proceedings of the 6th Australian Information Security & Management Conference, 1-3 December, Perth, Australia

23. Burge, M. and Burger, W. (2000) “Ear biometrics in computer vision”, Proceedings of 15th International Conference on Pattern Recognition, vol.2, pp.822-826,

24. Burrows, J. (2007) “All the way through: testing for authorship in different frequency strata ”, Literary and Linguistic Computing, vol.22, no.1, pp.27-47

25. Buschkes, R., Kesdogan, D. and Reichl, P. (1998) “How to increase security in mobile networks by anomaly detection”, Proceedings of the 14th Annual Computer Security Applications Conference, pp.23-12

26. Business Wire, (2013) “Barclays Uses Nuance Voice Biometrics to Identify Customers by the sound of their voice”, available at: http://www.businesswire.com/news/home/20130508005400/en/Barclays-Nuance-Voice-Biometrics-Identify-Customers-Sound

27. Bustard, J.D. and Nixon, M.S. (2008) “Robust 2D ear registration and recognition based on sift point matching”, IEEE Conference on Biometrics: Theory, Applications and Systems (BTAS’08), Washington, DC, September 2008.

28. Bustard, J.D. and Nixon, M.S. (2010) “3D morphable model construction for robust ear and face recognition”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.2582-2589

29. Calix, K., Connors, M., Levy, D., Manzar, H., MCabe, G. and Westcott, S. (2008) “Stylometry for e-mail author identification and authentication”, Proceedings of CSIS Research Day, Pace University.

30. Campbell, J. P. (1997) “Speaker Recognition: A Tutorial”, Proceedings of the IEEE, vol.85, no.9, pp.1437-1462

31. Campisi, P., Maiorana, E., Bosco, M. L. and Neri, A. (2009) “User Authentication Using Keystroke Dynamics for Cellular Phones”, IET Signal Processing - Special Issue on Biometric Recognition, vol.3, no.4, pp.333-341

32. Castro, A., Sotoye, O., Torres, L., Truly, G., Monaco, V. and Stewart, J (2011) “A Stylometry System for Authenticating Students Taking Online Tests”, Proceedings of CSIS Research Day, Pace University.

33. Chen, H. and Bhanu, B. (2007) “Human Ear recognition in 3D”, IEEE transactions on pattern analysis and machine intelligence, vol.29, no.4, pp.718-737

34. Changa, T., Tsaib, C. and JLina, H. (2012) “A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices”, The Journal of Systems and Software, vol.85, pp.1157-1165

35. Coyotl-Morales, R., Villasenor Pineda, L., Montes-yGomez, M. and Rosso, P. (2006), “Authorship attribution using word sequences.” Proceeding of the 11th Iberoamerican Congress on Pattern Recognition, Cancun, Mexico: Springer, pp.844-53

36. ComputerWeekly, (2010) “Millions download suspicious Android wallpaper”, available at: http://www.computerweekly.com/Articles/2010/08/24/242169/Millions-downloadsuspicious- Android-wallpaper.htm

37. Credant (2008) “How a lost or stolen cell phone can lead to identity theft”, available at: http://www.helium.com/items/1161605-how-a-lost-or-stolen-cell-phone-can-lead-to-identity-theft

38. Credant (2009) “Phone Data makes 4.2 Million* Brits Vulnerable to ID Theft”, available at: http://www.credant.com/news-a-events/press-releases/337-phone-data-makes-42-million-brits-vulnerable-to-id-theft-.html

39. Cha, B.H. (2009) “Robust MC-CDMA-Based Fingerprinting Against Time-Varying Collusion Attacks”, IEEE Transactions on Information Forensics and Security, vol.4, no.3, pp.302-317

40. Chellappa, R., Wilson, C.L., and Sirohey, S. (1995) “Human and machine recognition of faces: a survey”, Proceeding of the IEEE, vol.83, no.5, pp.705-740

41. Chen, H. and Bhanu, B. (2007), “Haman Ear Recognition in 3D”, IEEE Transactions on pattern analysis and machine intelligence, vol.29, no.4, pp.718-737

42. Cho, S., Han, C., Han, D.H. and Kim, H.I. (2000) “Web-Based keystroke Dynamics Identity Verification Using Neural Network”, Journal of Organizational Computing and Electronic Commerce, vol.10, no.4, pp.295-307

43. Cisco IBSG (2012) “BYOD and Virtualization Top 10 Insights from Cisco IBSG Horizons Study”, available at: http://www.cisco.com/web/about/ac79/docs/re/IBSG_Horizons_BYOD_KeyInsights.pdf

44. Cisco IBSG (2009) “Data Leakage Worldwide: Common Risks and Mistakes Employees Make”, available at: http://www.cisco.com/en/US/solutions/collateral/ns170/ns896/ns895/white_paper_c11-499060.pdf

45. Clarke, N. Furnell, S.M. Reynolds, P.L. and Rodwell, P.M. (2002) “Advanced Subscriber Authentication Approaches For Third Generation Mobile Systems”, IEEE conference publication, Institute of Electrical Engineers, pp.319-323

46. Clarke, N. (2004) “Advanced User Authentication for Mobile Devices”, PhD thesis.

47. Clarke, N., Furnell, S., Lines, B., and Reynolds, P. (2004) “Application of Keystroke analysis to mobile text messaging”, Proceedings of the 3rd Security Conference, Las Vegas, USA

48. Clarke, N. and Furnell, S.M. (2005) “Authentication of users on mobile telephones – A survey of attitudes and practices”, Computer & Security, vol.24, no.7, pp.519-527

49. Clarke, N. and Furnell, S.M. (2006) “Authenticating Mobile Phone Users Using Keystroke Analysis”, International Journal of Information Security, ISSN: 1615-5262, pp.1-14

50. Clarke, N. and Mekala, A.R. (2007) “The application of signature recognition to transparent handwriting verification for mobile devices”, Information Management & Computer Security, vol.15, no.3, pp.214-225

51. Clarke, N., Karatzouni, S., and Furnell, S. (2008) “Transparent Facial Recognition for Mobile Devices”, Proceedings of the 7th Security Conference, Las Vegas, USA, 2nd-3rd June, 2008

52. Clarke, N. (2011) “Transparent User Authentication”, Springer, ISBN 978-0-85729-804-1

53. CFCA (2009) “Communications fraud control association (CFCA) announces results of worldwide telecom fraud survey”, available at: http://www.cfca.org /pdf/survey/2009%20Global%20Fraud%20Loss%20Survey-Press%20Release.pdf

54. Das, R. (2007) “Signature recognition: An introduction to signature recognition as a biometric technology”, Keesing Journal of Documents & Identity, no.24, pp.13-14.

55. Daugman, J. (1993) “High confidence recognition of persons by a test of statistical independence ”, IEEE transaction on PAMI, vol.15, no.11, pp.1148-1161

56. Daugman, J. (1997) “Face and gesture recognition: overview”, IEEE transactions on pattern analysis and machine intelligence, vol.19, no.7, pp.675-676

57. Dave, G., Chao, X. and Sriadibhatla, K. (2011) “Face Recognition in mobile Phones”, Department of Electrical Engineering Stanford university

58. Delac, K. and Grgic, M. (2004) “A survey of biometric recognition methods”, International Conference on Symposium Electronics in Marine, pp.184-193

59. Diederich , J., Kindermann, J., Leopold, E. and Paass G. (2003) Authorship Attribution with Support Vector Machines, Applied Intelligence, vol.19 n0.1-2, pp.109-123

60. Dimensional Research (2012), “The impact of mobile devices on information security: A survey of IT professionals”, available at: http://www.checkpoint.com/downloads/products/check-point-mobile-security-survey-report.pdf

61. Distimo (2010), “Our Presentation From Mobile World Congress 2010 – Mobile Application Stores State Of Play”, available at: http://blog.distimo.com/2010_ 02_our-presentation-from-mobile-world-congres-2010-mobile-application-stores-state-of-play/

62. Doddington, G. R., Przybocki, M. A., Martin, A.F. and Reynolds, D. A. (2000) “The NIST speaker recognition evaluation–overview, methodology, systems, results, perspective.”, Speech Communication, vol.31, no.2-3, pp.225-254

63. Doring, N. (2002) “Abbreviation and acronyms in SMS communication”, available at: http://www.nicola-doering.de/.

64. Du, Y. (2006) "Review of iris recognition: cameras, systems, and their applications", Sensor Review, vol.26, no.1, pp.66 – 69

65. Eagle, N., Pentland, A. and Lazer, D. (2009) “Inferring Social Network Structure using Mobile Phone Data”, Proceeding of National Academy of Sciences (PNAS), vol.106, pp.15274-1578

66. Ericsson (2012a) “Traffic and market report on the pulse of the networked society”, available at: http://www.ericsson.com/res/docs/2012/traffic_and_ market_report_june_2012.pdf

67. Ericsson (2012b) “Ericsson Mobility Report”, available at: http://www.ericsson.com/res/docs/2012/ericsson-mobility-report-november-2012.pdf

68. Etemad, K. and Chellappa, R. (1997) “Discriminant Analysis for Recognition of Human Face Images”, Journal of the Optical Society of America A, Vol. 14, No. 8, pp.1724-1733, August 1997.

69. Face-rec (2011) “Vendors”, available at: http://www.face-rec.org/vendors/

70. Farzin, H., Moghaddam, H.A. and Moin, M. S. (2008) “A novel retina identification system”, EURASIP Journal on Advances in Signal Processing,no.260835,pp.1-10

71. Forge, S. (2004) “Is fourth generation mobile nirvana or … nothing?”, info, vol.6, no.1, pp.12-23

72. Frost & Sullivan (2011) “Promised Market for NFC Effectively Commences in 2011 with Commercial Roll out within All verticals ”, available at: http://www.frost.com / prod/servlet/press-release.pag?docid=223107191

73. Fu, K.S., Pyung J.M. and Li, T.J. (1970) "Feature Selection in Pattern Recognition", Systems Science and Cybernetics, IEEE Transactions on Systems Science and Cybernetics, vol.6, no.1, pp.33-39

74. Furnell, S., Rodwell, P. and Reynolds, P. (2001). “A Conceptual Security Framework to Support Continuous Subscriber Authentication in Third Generation Networks”, Proceedings of Euromedia 2001.

75. Gamon, M. (2004) “Linguistic correlates of style: authorship classification with deep linguistic analysis features”, Proceeding of the 20th international conference on Computational Linguistics (COLING’04), Association for Computational Linguistics, Stroudsburg, PA, USA

76. Gartner (2010), “Gartner Outlines 10 Mobile technologies to watch in 2010 and 2011”, available at: http://www.gartner.com/newsroom/id/1328113

77. Gartner (2011), “Gartner Says Worldwide Mobile Application Store Revenue Forecast to Surpass $15 billion in 2011”, available at: http://www.gartner.com/ newsroom/id/1529214

78. Gartner (2012), “Gartner Identifies the Top 10 Strategic Technology Trends for 2013”, available at: http://www.gartner.com/newsroom/id/2209615

79. Goode Intelligence (2011) “Significant growth expected from mobile biometric security market” available at: http://www.goodeintelligence.com/media-centre/view/significant-growth-expected-from-mobile-phone-biometric-security-market/

80. Goodman, R., Hahn, M., Marella, M., Ojar, C. and Westcott, S. (2007) “The use of stylometry for email author identification: a feasibility study”, Proceedings of Student/Faculty Research Day, CSIS, Pace University, May 2007

81. Gosset, P. (1998) “ASPeCT: Fraud Detection Concepts: Final Report”, Doc Ref. AC095/VOD/W22/DS/P/18/1

82. Grady, M. (2012) “SMS usage remains strong in the US:6 billion SMS messages are sent each day”, available at: http://blogs.forrester.com/michael_ogrady/ 12-06-19-

sms_usage_remains_strong_in_the_us_6_billion_sms_messages_ are_sent_each_day?cm_mmc=RSS-_-MS-_-1710-_-blog_Michael%20O% 27Grady

83. Guyon, I. and Elisseeff, A. (2003) “An Introduction to Variable and Feature Selection”, Journal of Machine Learning Research, vol.3, pp.1157-1182

84. Halteren, V. H. (2004) “Linguistic Profiling for Author Recognition and Verification”, Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics (ACL 04), Association for Computational Linguistics, Morristown, NJ, USA,

85. Hall, J., Barbeau, M. and Kranakis, E. (2005) “Anomaly-based intrusion detection using mobility profiles of public transportation users”, the Proceeding of IEEE International Conference on Wireless And Mobile Computing, Networking And Communications, 2005 (WiMob'2005), vol.2, pp.17- 24, ISBN: 0-7803-9181-0

86. Halteren, H. (2004) “Linguistic Profiling for Author Recognition and Verification”, Proceeding of the 42nd annual meeting of the association for computational linguistics, pp.199-206

87. Haykin, S. (1999) “Neural networks: A Comprehensive Foundation (2nd edition)”, Prentice Hall, New Jersey

88. Helium (2008) “How a lost or stolen cell phone can lead to identity theft”, available at: http://www.helium.com/items/1161605-how-a-lost-or-stolen-cell-phone-can-lead-toidentity-theft.

89. Hirst, G. and Feiguina, O. (2007), “Bigrams of syntactic labels for authorship discrimination of short texts”, Literary and Linguistics Computing, vol.22, no.4, pp.405-417

90. Hocquet, S., Ramel, J. and Cardol, H. (2005) “Fusion of Methods for Keystroke Dynamic Authentication”, Proceedings of 4th IEEE Workshop on Automatic Identification Advanced Technologies, USA, pp.224-229

91. Hollingsworth, K.P., Bowyer, K.W. and Flynn, P.J. (2009) "The Best Bits in an Iris Code," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.31, no.6, pp.964-973

92. Hong, L. and Jain, A. (1999) “Multimodal biometric”, Biometrics: Personal Identification in Networked Society, Kluwer.

93. How, Y. and Lee, M. F. (2004) “NUS SMS Corpus”, available at: http://www.comp.nus.edu.sg/~rpnlpir/downloads/corpora/smsCorpus.

94. Hube, P.J. (1981) “Robust Statistics”, Wiley, New York.

95. Iannarelli, A. (1989), “Ear Identification”, Paramont Publishing.

96. IBG (2010) “How is biometric defined?” International Biometric Group, available at: http://www.biometricgroup.com/reports/biometric_definition.html

97. IBM (2011) “Securing mobile devices in the business environment”, available at: http://public.dhe.ibm.com/common/ssi/ecm/en/sew03027usen/SEW03027USEN.PDF

98. Indovina, M., Uludag, U., Snelick, R., Mink, A. and Jain, A.K. (2003) “Multimodal biometric Authentication Methods: A COTS Approach”, Proceeding of Workshop on Multimodal User Authentication, Santa Barbara CA, pp.99-106

99. Iqbal, F., Khan, L.A., Fung, B.C.M. and Debbabi, M. (2010) “E-mail authorship verification for forensic investigation”, Proceedings of the 25th ACM SIGAPP Symposium on Applied Computing (SAC), pp.1591–1598

100. ITU (2010) “The world in 2010 ICT Facts and Figures”, available at: http://www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf

101. ITU (2011) “The world in 2011: ICT Facts and Figures”, available at: http://www. itu.int/en/ITU-D/Statistics/Documents/facts/ ICTFactsFigures2011.pdf

102. Jain, A.K., Bolle, R. and Pankanti, S. (1999a) “Biometrics: Personal Identification in networked Society”, Norwell, MA:Kluwer.

103. Jain, A.K., Prabhakar, S. and Hong, L. (1999b) “A Multichannel Approach to Fingerprint Classification”, IEEE Transactions on pattern analysis and machine intelligence, vol.21, no.4, pp.348-382.

104. Jain, A.K., Duin, R.P.W. and Jianchang, M. (2000) “Statistical pattern recognition: a review”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.22, no.1, pp.4-37

105. Jain, A. K., Prabhakar, S. and Pankanti, S. (2002) “On the similarity of identical twin fingerprints”, Pattern Recognition, vol.35, Issue: 11, pp.2653-2663

106. Jain, A. K., Ross, A. and Prabhakar, S. (2004) “An Introduction to Biometric Recognition” , IEEE Transactions on circuits and systems for video technology, vol.14, no.1, pp.4-20

107. Jain, A.K., Flynn, P. and Ross, A. (2007) “Handbook of Biometrics”, New York, Springer-Verlag

108. Joachims, T. (1998) “Text Categorization with Support Vector Machines: Learning with Many Relevant Features”, Proceeding of the European Conference on Machine Learning, pp.137-142.

109. Jockers, M.L. and Witten, D.M. (2010) “A comparative study of machine learning methods for authorship attribution”, Literary and Linguistic Computing, vol.25, no.2, pp.215-223.

110. Joyce, R. and Gupta, G. (1990) “Identity Authentication Based on keystroke Latencies”, Communication of the ACM, vol.39, pp.168-176

111. Juniper Research (2010) “A World of Apps”, available at: http://www. juniperresearch.com/shop/products/whitepaper/pdf/MAS10_White%20Paper.pdf

112. Karatzouni, S., Clarke, N.L. and Furnell, S.M. (2007) “NICA design specification”, University of Plymouth, available at: http://www.cscan.org/nica/

113. Keselj, V., Peng, F., Cercone, N. and Thomas, C. (2003) “N-gram-based Author Profiles for Authorship Attribution”, Proceedings of the Conference Pacific Association for Computational Linguistics. Dalhousie University, Halifax, Nova Scotia, Canada

114. Kaspersky Lab (2011), “European Users Mobile Behaviour and Awareness of Mobile

Threats", available at: http://www.kaspersky.com/news?id=207576289

115. Kekre, H.B., Thepade, S.D., Jain, J. and Agrawal, N. (2011) "Iris recognition using texture features extracted from walshlet pyramid", Proceedings of International Conference and workshop on Emerging Trends in Technology (ICWET), pp.76–81.

116. Koppel, M. and Schler, J. (2003), "Exploiting Stylistic Idiosyncrasies fir Authorship Attribution", Proceeding of IJCAI'03 Workshop on Computational Approaches to Style Analysis and Synthesis, pp.69-72.

117. Koreman, J., Morris, A. C., Wu, D., Jassim, S., Sellahewa, H., Ehlers, J., Chollet, G., Aversano, G., Bredin, H., Garcia-salicetti, S., Allano, L., Van, B., and Dorizzi, B. (2006) "Multi-modal biometric authentication on the SecurePhone PDA", Proceeding of the second workshop on multimodal user authentication, Toulouse, France.

118. Kounoudes, A., Tsapatsoulis, N., Theodosiou, Z. and Milis, M. (2008) "POLYBIO: Multimodal Biometric Data Acquisition Platform and Security System", Biometrics and Identity Management, Springer Berlin Heidelberg, vol.5372, pp.216–227.

119. Kusakci, A. O. (2012) "Authorship attribution using committee machines with k-nearest neighbors rated voting", The 11th Symposium on Neural Network Applications in Electrical Engineering, pp.161-166.

120. Kurkovsky, S. and Syta, E. (2010) "Digital natives and mobile phones: A survey of practices and attitudes about privacy and security", Proceedings of the 2010 IEEE International Symposium on Technology and Society (ISTAS), pp.441-449.

121. Kumar, A., Wong, D., Shen, H. and Jain, A. K. (2003) "Personal Verification Using Palmprint and Hand Geometry biometric", Proceeding of the 4th International Conference Audio and Video-based Biometric Person Authentication, Springer Berlin Heidelberg, pp.668-678

122. Lam, L. and Suen, C. (1995) "Optimal combinations of pattern classifiers", Pattern Recognition Letters, vol.16,no.9, pp9.45-954

123. Lam, L., Huang, Y. and Suen, C. (1997) “Combination of Multiple Classifier Decisions for Optical Character Recognition”, Handbook of Character Recognition and Document Image Analysis, Chapter 3, World Scientific,pp.79-101

124. Ledger, TG.R., and Merriam, T.V.N. (1994) “Shakespeare, Fletcher, and the two noble Kinsmen”, Literary and Linguistic Computing, vol.9, 235-248

125. Lee, H. C. and Gaensslen R.E. (1991) “Advances in Fingerprint Technology”, New York: Elsevier

126. Leggett, J., Williams, G. and Usnick, M. (1991). “Dynamic Identity Verification via Keystroke Characteristics”, International Journal of Man-Machine Studies.

127. Li, B., Yuzhong, C. and Yu, S, (2002) “A Comparative Study on Automatic Categorization Methods for Chinese Search Engine”, Proceedings of the Eighth Joint International Computer Conference, Hangzhou: Zhejiang University Press, pp.117-120.

128. Li, B., Yu, S. and Lu, Q. (2003) “An Improved k-nearest neighbour Algorithm for Text Categorization”, Proceeding of the 20 International Conference on Computer Processing of Oriental Language, Shenyang, China

129. Li, F., Clarke, L.N., Papadaki M., and Dowland, P.S. (2010) “Behaviour Profiling on Mobile Devices”, International Conference on Emerging Security Technologies, 6-8 September, Canterbury, UK, pp77-82

130. Li, F., Clarke, L.N., Papadaki, M., and Dowland, P.S. (2011) “Behaviour Profiling for Transparent Authentication for Mobile Devices”, PhD thesis.

131. Lindoso, A., Entrena, L., López-Ongil, C. and Liu-Jimenez, J. (2005) "Correlation-Based Fingerprint Matching Using FPGAs", Proceedings of Field-Programmable Technology, pp.87-94.

132. Locket (2013) “Study Says We Unlock Our Phones a LOT Each Day”, available at: http://www.getlocket.com/press/

133. McAfee (2009), “Mobile Security report 2009”, available at: http://www.mcafee.com/uk/resources/reports/rp-mobile-security-2009.pdf

134. McAfee (2011), “Mobility and Security”, available at: http://www.mcafee.com/uk/resources/reports/rp-cylab-mobile-security.pdf

135. McAfee (2013), “McAfee Reveals Consumers Fail To Protect Their Mobile Devices”, available at: http://www.mcafee.com/us/about/news/2013/q1/ 20130224-01.aspx

136. Maddala, S., Tangellapally, S. R., Bartuněk, J. S. and Nilsson, M. (2011) “Implementation and evaluation of NIST Biometric Image Software for fingerprint recognition”, International Conference on Biosignals and Biorobotics (BRC), pp.1-5

137. Maiorana, E., Campisi, P., Fierrez, J., Ortega-Garcia, J. and Neri, A. (2010) “Cancelable templates for sequence-based biometrics with application to on-line signature recognition”, IEEE Transactions on System, Man and Cybernetics, Part A: Systems and Humans, vol.40, no.3, pp.525-538

138. Maiorana, E., Campisi, P., Gonzalez-Carballo, N. and Neri, A. (2011) “Keystroke dynamics authentication for mobile phones”, Proceedings of the 2011 ACM Symposium on Applied Computing (SAC’11), pp.21-26

139. Maltoni, D., Maio, D., Jain, A K. and Prabhakar, S. (2003) “Handbook of Fingerprint Recognition”, Springer-Verlag, New York

140. Mendenhall, T.C. (1887) “The characteristic curves of composition”, Science, vol.11, no.11, pp.237-249

141. Metropolitan Police Service (2011) “Safeguarding your mobile phone”, available at: http://www.met.police.uk/crimeprevention/phone.htm

142. Miyazawa, K., Ito, K., Aoki, T., Kobayashi, K. and Nakajima, H. (2008) “An Effective Approach for Iris Recognition Using Phase-Based Image Matching ”, IEEE transactions on Pattern Analysis and Machine Intelligence, vol.30, no.10, pp.1741-1756

143. Mohan, A., Baggili, M., and Rogers, M. K. (2010) “Authorship attribution of SMS messages using an N-grams approach”, available at: https://www.cerias.purdue.edu/assets/pdf/bibtex_archive/2010-11.pdf

144. Monrose, R. and Rubin, A. (1997). “Authentication via Keystroke Dynamics”,

Proceedings of the 4th ACM Conference on Computer and Communication Security, ACM, pp.48-56

145. Moreau, Y., Verrelst, H. and Vandewalle, J. (1997) “Detection of mobile phone fraud using supervised neural networks: A first prototype”, International Conference on Artificial Neural Networks Proceeding (ICANN’97), pp.1065-1070

146. Mosteller, F. and Wallace, D. (1964) “Inference and Disputed Authorship: The Federalist”, Series in Behaviour Science: Quantitative Methods ed. Addison-Wesley, Massachusetts.

147. Napier, R., Laverty, W., Mahar, D., Henderson, R., Hiron, M. and Wagner, M. (1995) “Keyboard User Verification: Toward an Accurate, Efficient and Ecological Valid Algorithm”, International Journal of Human-Computer Studies, vol.43, pp.213-222

148. National Mobile Phone Crime Unit (2013), available at: http://www.nmpcu.police.uk/

149. Newscientist (2007) “Cellphone firewall”, available at: http://www.newscientist.com/blog/invention/2007/04/cellphone-firewall.html

150. Newman, R. (2009) “Security and Access Control Using Biometric Technologies: Application, Technology and Management”, Course Technology Press, Boston, MA, United States

151. Nilsson, N.J. “Learning Machines: Foundations of Trainable Pattern-Classifying Systems”, New York: Mcffraw-Hill, 1965.

152. Obaidat, S. M. and Sadoun, B. (1997) “Verification of Computer User Using Keystroke Dynamics”, IEEE Transactions on Systems, Man and Cybernetics – Part B: Cybernetics, vol.27, no.2, pp.261-269

153. Ofcom (2012) “UK is now texting more than talking”, available at: http://media.ofcom.org.uk/2012/07/18/uk-is-now-texting-more-than-talking/

154. ONS (2011) “Nearly half of Internet users have accessed the Internet via a mobile phone”, available at: http://www.ons.gov.uk/ons/dcp29904_230811.pdf

155. O2 (2012), “Making calls has become fifth most frequent use for a Smartphone for newly-networked generation of users”, available at: http://news.o2.co.uk/?press-release=making-calls-has-become-fifth-most-frequent-use-for-a-smartphone-for-newly-networked-generation-of-users

156. Ord, T. and Furnell, S. (2000) “User Authentication for Keypad-Based Devices using Keystroke Analysis”, MSc Thesis, University of Plymouth, UK.

157. Ouamour, S. and Sayoud, H. (2012) “Authorship attribution of ancient texts written by ten arabic travelers using a SMO-SVM classifier”, International Conference on Communications and Information Technology (ICCIT), pp.44–47

158. Pflug, A. and Busch, C. (2012) “Ear Biometrics - A Survey of Detection, Feature Extraction and Recognition Methods”, IET Biometrics Journal, ISSN 2047-4938, vol.1, no.2, pp.114–129

159. Portio Research (2012) “Portio Research Mobile Factbook 2012”, available at: http://www.portioresearch.com/media/1797/Mobile%20Factbook%202012.pdf

160. ProtecStar (2009) “ProtectStar™ Mobile Firewall 1.0”, available at http://www.protectstar.com/index.php?y=60&x=73

161. Rafi, M. (2009) “SMS Text Analysis: Language, Gender and Current Practices”, available at: http://www.tesol-france.org/Documents/Colloque07/SMS%20Text%20Analysis%20Language%20Gender%20and%20Current%20Practice%20_1_.pdf

162. Rao, J.R., Rohatgi, P., Scherzer, H. and Tinguely, S. (2002) “Partitioning Attacks: Or How to Rapidly Clone Some GSM Cards”, IEEE Symposium on Security and Privacy, 2002, pp.31-41

163. Robinson, J.A., Liang, V.W., Chambers, J.A.M. and MacKenzie, C.L. (1998) “Computer user verification using login string keystroke dynamics”, Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on Systems, Man, and Cybernetics, vol.28, no.2, pp.236-241

164. Rodrigues, R.N., Kamat, N. and Govindaraju, V. (2010) “Evaluation of Biometric Spoofing in a Multimodal System”, International Conference on Biometrics: Theory Applications and Systems, pp.1-5

165. Rodwell, M., Furnell, M. and Reynolds, L. (2007) “A non-intrusive biometric authentication mechanism utilising physiological characteristics of human head”, Computer & Security, vol.26, no.7, pp.468-478

166. Ross, A., Jain, A. K. and Nandakumar, K. (2006) “Handbook of Multibiometrics”, Springer, berlin, Germany

167. Ross, A. and Abaza, A. (2011) “Human Ear Recognition”, Computer, vol.44, no.11, pp.79-81

168. Ross, A. and Govindarajan, R. (2005) “Feature Level Fusion Using Hand and Face Biometrics”, Proceeding of SPIE conference on biometric technology for human Identification II, vol.5779, pp.196-204

169. Samfat, D. and Molva, R. (1997) “IDAMN: an Intrusion Detection Architecture for Mobile Networks”, IEE journal on Selected Areas in Communications, vol.15, pp.1373-1380.

170. Sanderson, C. and Guenter, S. (2006), “Short text authorship attribution via sequence kernels, Markov chains and author unmasking: An investigation”, Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP), Sydney, Australia: Association for Computational Linguistics, pp.482-91

171. Selman, S. and Husagic-Selman, A. (2011) “Multilayered feedforward neural networks as a tool for distinction of the authors of texts”, Information, Communication and Automation Technologies (ICAT), pp.1-6

172. Smith, R. (2002) “Authentication. From Passwords to Public Keys”, Addison-Wesley

173. Snelick, R., Indovina, M., Yen, J. and Min, A. (2003) “Multimodal Biometrics: Issues in Design and Testing”, Proceedings of the 5th international conference on Multimodal interfaces, pp.68-72

174. Snelick, R., Uludag, U., Mink, A., Indovina, M., and Jain, A.K. (2005) “Large-Scale Evaluation of Multimodal Biometric Authentication Using State-of-the-Art Systems”, IEEE Transactions on pattern analysis and machine intelligence, vol.27, no.4 pp.450-455

175. Soltane, M., Doghmane, N. and Guersi, N. (2010) “Face and Speech Based Multi-Modal Biometric Authentication”, International Journal of Advanced Science and Technology, vol.21, no.8, pp.41–46

176. Stańczyk, U. and Krzysztof, A. C. (2007) “Machine learning approach to authorship attribution of literary texts”, International Journal of Applied Mathematics and Informatics, vol.1, no.4, pp.151-158

177. Stamatatos, E., Fakotakis, N. and Kokkinakis, G. (2001) “Computer-based authorship attribution without lexical measures”, Computers and the Humanities, vol.35, no.2, pp.193–214.

178. Stamatatos, E. (2007) “Author identification using imbalance and limited training texts”, Proceedings of the 18th International Conference on Database and Expert Systems Applications, Regensburg, Germany: IEEE Computer society, pp.237-241.

179. Stamatatos, E. (2009) “A survey of modern authorship attribution methods,” Journal of the American Society for Information Science and Technology, vol.60, no.3 pp.538–556.

180. Stolfo, S.J., Wei, F., Wenke, L., Prodromidis, A. and Chan, P.K. (2000) “Cost-based modelling for fraud and intrusion detection: results from the JAM project”, DARPA Information Survivability Conference and Exposition, pp.130-144

181. Sun, B., Chen, Z., Wang, R., Yu, F. and Leung, V.C.M. (2006) “Towards adaptive anomaly detection in cellular mobile networks”, Consumer Communications and Networking Conference, 2006 (CCNC 2006), vol.2, pp.666-670, ISBN: 1-4244-0085-6

182. Sun, B., Yu, F., Wu, K. and Leung, V. (2004) “Mobility-based anomaly detection in cellular mobile networks”, Proceedings of ACM wireless security (WiSe’ 04), Philadelphia, PA, pp.61-69

183. Sun, Z., Tan, T. and Qiu, X. (2006) “Graph matching iris imageblocks with local binary pattern”, International Conference on Biometrics (Springer LNCS 3832), pp.366-372

184. Symantec (2012) “State of Mobility Survey”, available at: http://www.symantec.com/content/en/us/about/media/pdfs/b-state_of_mobility_survey_2012.en-us.pdf

185. Tanviruzzaman, M., Ahamed, S., Hasan, C. and O’brien, C. (2009) “epet:When Cellular Phone Learns to Recognize its owner”, Proceeding of the 2009 2nd ACM workshop on Assurable and usable security configuration, pp.13-17

186. Times Newspapers (2007) “How quickly did you type that password?”, available at: http://technology.timesonline.co.uk/tol/news/tech_and_web/personal_tech/article1667057.ece,

187. Toshiba America Information Systems (2011) “Face recognition”, available at: http://us.toshiba.com/computers/research-center/technology-guides/face-recognition

188. Trend Micro (2009) “Trend Micro™ Mobile Security 5.0” available at: http://us.trendmicro.com/us/products/enterprise/mobile-security/index.html

189. Tsimboukakis, N. and Tambouratzis, G. (2010) “A comparative study on authorship attribution classification tasks using both neural network and statistical methods”, Neural Computing and Application, vol.19, no.4, pp.573–582

190. Turk, M.A. and Pentland, A.P. (1991) “Face Recognition Using Eigenfaces”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3-6 June 1991, Maui, Hawaii, USA, pp.586-591.

191. Turkoglu, F., Diri, B. and Amasyali, M.F. (2007) “Author Attribution of Turkish Texts by Feature Mining”, Third International Conference on Intelligent Computing (ICIC2007), Qingdao, China, LNCS, vol.4681, pp.1086-1093

192. UKBA (2011) “Using the iris recognition immigration system (IRIS)”, available at: http://www.ukba.homeoffice.gov.uk/travellingtotheuk/Enteringtheuk/usingiris/

193. Umphress, D. and Willams, G. (1985) “Identity verification through keyboard characteristics”, International Journal of Man-Machine Studies, vol.23, pp.263-273

194. De Vel, O., Anderson, A., Corney, M. and Mohay, G. (2001) “Mining E-mail Content for Author Identification Forensics”, ACM Sigmod Record, ACM New York, NY, USA, vol.30,no.4, pp.55–64,

195. Weinstein, E., Ho, P., Heisele, B., Poggio, T., Steele, K. and Agarwal, A. (2002) “Handheld face identification technology in a pervasive computing environment”, In Pervasive 2002, Zurich, Switzerland, pp.48–54

196. Wildes, R. (1997) “Iris recognition: an emerging biometric technology”, Proceeding of the IEEE, vol.85, no.9, pp.1348-1363

197. Wilson, C., Hicklin, A.R., Bone, M., Korves, H., Grother, P., Ulery, B., Micheals, R., Zoepfl, M., Otto, S. and Watson, C. (2004) “Fingerprint Vendor Technology Evaluation 2003: Summary of Results and Analysis Report.”, NIST Technical Report NISTIR 7123, National Institute of Standards and Technology

198. Wisegeek (2011) “How does a Retinal Scan Work?”, available at: http://www.wisegeek.com/how-does-a-retinal-scan-work.htm

199. Wright, J., Yang, A., Ganesh, A., Sastry, S. and Ma, Y. (2009) “Robust face recognition via sparse representation”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.30, no.2, pp.210-227

200. Wolpert, D.H. and Macready, W.G. (1997) "No Free Lunch Theorems for Optimization", IEEE Transactions on Evolutionary Computation 1, pp.67-82.

201. Woo, R., Park, A. and Hazen, T. (2006) “The MIT Mobile Device Speaker Verification Corpus: Data collection and preliminary experiments”, Proceeding of Odyssey, The Speaker & Language Recognition Workshop, San Juan, Puerto Rico, pp.1-6

202. Yampolskiy, R. and Govindaraju, V. (2008) “Behavioural biometrics: a survey and classification”, International journal of Biometrics, vol.1, no.1, pp.81-113

203. Yan, P. and Bowyer, K.W. (2007) “Biometric Recognition Using 3D Ear Shape”, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.29, no.8, pp.1297-1308.

204. Yang, Y. and Liu X. (1999) “A Re-examination of Text Categorization Methods”, Proceeding of 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp.42-49

205. Yang, J., Zhang, D. and Yang, J.Y. (2007) “Constructing PCA baseline algorithms to re-evaluate ICA-Based face-recognition performance”, IEEE Transaction Systems Man and Cybernetics, vol.37, no.4, pp.1015-1021.

206. Yuan, L. and Mu, Z (2007) “Ear Recognition based on 2D Images”, First IEEE International Conference on Biometrics: Theory, Applications, and Systems, pp.1-5.

207. Zheng, R., Li, J., Chen, H. and Huang, Z. (2006) “A Framework for Authorship Identification of Online Messages: Writing-Style Features and Classification Techniques”, Journal of the American Society for Information Science and Technology, vol.53, no.3, pp.378-393

208. Zhao, W., Chellappa, R., Phillips, P.J. and Rosenfeld, A. (2003) “Face recognition: a literature survey”, ACM Computing Surveys, December, 2003, vol.35, no.4, pp.399–458

Appendix A: SMS dataset

Appendix B: The SMS communication simulation scripts

Appendix C: The Neural Networks scripts

Appendix D: The simulation scripts

Appendix E: The preliminary study's experimental results

Appendix F: The simulation results

Appendix G: Publications

- **Text-based Active Authentication for Mobile Devices**

  Saevanee H., Clarke N. L., Furnell S. M., Accepted for Proceeding of 27th IFIP International Information Security and Privacy Conference (SEC 2014), Marrakech, Morocco, 2-4 June, 2014

- **Multi-Modal Behavioural Biometric Authentication for Mobile Devices**

  Saevanee H., Clarke N. L., Furnell S. M., Proceeding of 27th IFIP International Information Security and Privacy Conference (SEC 2012), Heraklion, Crete, Greece, 4-6 June, pp465-474, 2012

- **SMS Linguistic Profiling Authentication on Mobile Devices**

  Saevanee H., Clarke N. L., Furnell S. M., Proceedings of the 5th International Conference on Network and System Security (NSS 2011), p224-229, 2011

- **Behavioural Biometric Authentication for Mobile Devices**

  Saevanee H., Clarke N. L., Furnell S. M., Proceedings of the Collaborative European Research Conference (CERC2011), 14-15 January, Cork, Ireland, ISSN: 2220-4164, pp175-184 2011

Text-Based Active Authentication for Mobile Devices

Hataichanok Saevanee[1], Nathan Clarke[1,3], Steven Furnell[1,3] and Valerio Biscione[2]

[1]Centre for Security, Communications and Network Research,
[2]Centre for Robotics and Neural Systems,
Plymouth University, Plymouth, United Kingdom
[3]Security Research Institute, Edith Cowan University,
Perth, Western Australia
info@cscan.org

Abstract. As modern mobile devices are increasing in their capability and accessibility, they introduce additional demands in terms of security – particularly authentication. With the widely documented poor use of PINs, Active Authentication is designed to overcome the fundamental issue of usable and secure authentication through utilizing biometric-based techniques to continuously verify user identity. This paper proposes a novel text-based multimodal biometric approach utilizing linguistic analysis, keystroke dynamics and behavioral profiling. Experimental investigations show that users can be discriminated via their text-based entry, with an average Equal Error Rate (EER) of 3.3%. Based on these findings, a framework that is able to provide robust, continuous and transparent authentication is proposed. The framework is evaluated to examine the effectiveness of providing security and user convenience. The result showed that the framework is able to provide a 91% reduction in the number of intrusive authentication requests required for high security applications.

Keywords: Active authentication ·Transparent authentication · Continuous authentication · Multimodal · Biometric · Mobile devices

1 Introduction

Mobile devices are commonplace with over 6 billion subscribers worldwide [1]. With the rapid development of mobile network technology and the increasing popularity of mobile devices, modern devices are capable of providing a wide range of services and applications over multiple networks. The plethora of functionalities offered by the mobile device enables users to store increasing amounts of a wider variety of information from business to personal and sensitive data. A series of studies have highlighted the potential risk of mobile device misuse through the storing of personal information (e.g. home address), security credentials (e.g. PIN codes, user names and passwords) and business data (e.g. customer data) [2,3].

Although PIN or password authentication is available on most mobile devices, a survey conducted by [4] demonstrated that a third of mobile users do not protect their devices with this simple technique. Furthermore, the poor use of PIN or pass-

© Springer-Verlag Berlin Heidelberg 2011

word techniques when they are used is also widely documented in several studies [4,5]. A fundamental weakness of the PIN is that as a point-of-entry approach, once the user has been successfully authenticated, they obtain access to the system without having to re-authenticate. Several studies [6,7] proposed Active Authentication or transparent authentication to overcome the fundamental issue and more closely associate the authentication and access control decisions. There are a number of biometric techniques that have the potential to be used for authentication in a transparent and thus continuous fashion, such as keystroke dynamics, behavioral profiling, gait recognition, speaker verification and facial recognition. Unfortunately, research has demonstrated that using a single biometric may be inadequate for verification due to a variety of reasons, such as noise in the sample data, the unavailability of a sample at a given time and the underlying performance of the technique [8]. To overcome this limitation within traditional the point-of-entry domain, several researchers have proposed the use of multiple biometric modalities, which have demonstrated increased accuracy of verification [9,10,11].

This paper presents the findings of a research study exploring the application of multimodal biometric authentication in a transparent fashion to text-based entry. As users frequently use their mobile device to send SMS text messages (over 9.8 trillion in 2012), social network posts, emails and tweets, it was felt this medium provided a frequent opportunity to capture samples [12]. The focus upon text-based entry provides the possibility to apply keystroke dynamics, linguistic analysis and behavioral profiling. It is the aim of this paper to present the results of an exhaustive investigation into optimizing the recognition performance and an evaluation of the security processes required to maximize the security of the approach whilst minimizing user inconvenience. Section 2 presents the state of the art in behavioral biometrics that have been applied in the mobile domain. Section 3 describes the feasibility study of multimodal biometric. Based upon the results, a novel text-based multimodal framework that will provide the verification of a mobile user's identity in a continuous and transparent manner is proposed in Section 4 and then evaluated through simulation in Section 5. The paper concludes by highlighting the future direction of research in Section 6.

2 Text-based behavioral biometric for mobile devices

With the rapid evolution of mobile devices, utilizing biometrics on them has become a reality. Many mobile devices come equipped with a number of hardware components that are able to be used for capturing a variety of biometric traits, enabling several biometric approaches to be deployed – such as keystroke dynamics, behavioral profiling and voice recognition. For example, Apple has now incorporated TouchID, a fingerprint-based approach, and Google has Face Unlock for its Android Operating System [13,14]. To date, however, these are point-of-entry solutions that focus upon usability rather than security. Of interest in this research is the use of three behavioral biometric techniques: linguistic profiling, keystroke dynamics and behavioral profiling. It is hypothesized that the integration of these three techniques together offers the opportunity

to improve upon the usability through transparent capture, improve the overall recognition performance and mitigate the unavailability of samples at a given time.

Linguistic profiling is a behavioral biometric that identifies people based upon linguistic morphology. Previous studies have investigated the feasibility of linguistic profiling for several tasks such as text categorization, authorship identification and authorship verification. In the authorship verification domain, examples of writing from a single author are given to the system, which is then asked to confirm if the given texts were written by this author. According to previous studies [15], almost 1000 writing styles have been analyzed and both statistical and machine learning methods were used in the analytical process. Many studies have confirmed the good discriminating capability of linguistic features. Through using a machine learning method, the performance accuracies were in the range of 80%-100% [16,17]. However, there is no agreement on a best set of features for authorship verification and historically large volumes of text are required for the training dataset. The performance of linguistic profiling technique highly depends upon the combination of the selected features and classification models utilized.

Behavioral profiling aims to identify users based upon the way in which they interact with the services on their mobile device. Previous behavior-based studies have mainly focused upon the area of fraud detection. Research in mobile IDSs can be divided into two categories: call-based and mobility-based mechanisms. The former monitors user's calling behavior (e.g. start date of call and dial telephone number) that have been collected over a service provider's network during a period of time [18,19]. Based upon the theory that people have a predictable travelling pattern when they travel from one location to another, the mobility-based approach monitors a mobile user's location activities to detect abnormal behavior [20]. Through monitoring a user's calling or location activities, behavioral-based IDS can offer a high detection rate and ability to detect unforeseen attacks [18,19,20,21]. Depending upon application types, profiling techniques and classification approach, a study by [7] showed that behavioral profiling could be used for authentication on mobile devices with accuracies of between 87% and 98%.

Keystroke dynamics identifies a user based upon the typing pattern of a user, looking at characteristics of their interaction with a keyboard. Based upon previous studies, two main characteristics were identified: inter-key and hold time [24]. The inter-key is the duration between two successive keys. The hold-time represents the duration between the press down and releasing of a single key. Many studies have shown it is feasible to authenticate users successfully based upon usernames and passwords (i.e. in parallel with a typical Windows login request), with a commercial product on the market utilizing this technology [22, 23]. More recent studies [6, 24] investigated the possibility of using keystroke dynamics on mobile devices, showing the possibility of keystroke dynamic based authentication can be deployed in practice to provide an extra layer of security for mobile devices with an average accuracy of 87%.

Based upon the prior-art, these three techniques provide valuable discriminative information to permit identity authentication. All of the biometric traits of these three techniques can be captured during user interactions with a mobile device without a user explicit interaction to authenticate. In addition, no additional hardware is required to deploy these techniques. As a result, these approaches arguably provide a

cost effective and a non-intrusive solution for mobile handset authentication. Furthermore, a significant amount of prior research within the point-of-entry authentication domain [9,10,11] has concluded that using multiple biometric modalities can improve accuracy and reliability of single-modal systems. For example, using combination of fingerprint and face modality can achieve better performance than using single biometric, improving the accuracy of 2.3% at 0.1% FAR [25].

3 A feasibility study of text-based multimodal biometrics

Since no multimodal database availability where the above three biometric modalities are measured within the same individual, a standard practice employed within multi-biometrics is to combine the modalities from different datasets and create a virtual person [11]. The SMS corpus collected by the authors, a public mobile usage dataset provided by [26] and keystroke dataset provided by [24] were used in this experiment. An individual user from the linguistic profiling database was associated with an individual of keystroke and behavioral profiling database to create a virtual subject. As a result, a final database consisting of 30 users, each user having their SMS messages, keystroke and text messaging activity data was created and utilized in this experiment.

3.1 Experiment procedure

The experiments investigated the performance both of the individual techniques and their combination. To investigate the linguistic profiling's effectiveness; four types of linguistic features were examined: word profiling, lexical, syntactic and structural. The frequency distribution of a total 133 abbreviations and emotional words were used to create a user's word profiles, including 64 discriminating characteristics of every possible type of feature. To create a user profile, the t-test ranking measure was utilized to rank input features according to its discriminative capability. From the ranking list, features with a p value less than 0.05 were selected to create input vectors. The key to utilizing the t-test was to ensure a set of features that was as unique to the individual authorized user in comparison to the wider population. Therefore, the number of linguistic features required for discrimination will vary between users. Three different classification techniques: K-Nearest Neighbor (K-NN), the Radial Basis function (RBF) and Feed-Forward Multi-Layered Perceptron (FF-MLP) neural networks were utilized with differing network configurations - looking to optimum performance.

In the keystroke dynamics experiment, the hold time vector constructed from five letters: E, T, A, O and N were extracted. A number of analyses were undertaken using the FF-MLP neural network as it had demonstrated the better performance in previous studies over other techniques [24].

For the behavioral profiling technique, the following features were extracted: receiver's telephone number and location of texting. A number of analyses were undertaken, using a Radial Basis Function (RBF) neural network as it had performed the best in the prior study [7].

Table 1. Final dataset used in the experiments

	Training size	Testing size
Linguistic profiling	316	171
Keystroke dynamics	3339	171
Behavior profiling	1178	171

To perform the classification for the individual techniques, the dataset was divided into two groups: 171 data samples were used for the testing set and the remainder was used for training (as illustrated in Table 1). The pattern classification test was performed with one user acting as the valid user, while all others are acting as impostors (a standard procedure in this type of test) [6-8]. The Equal Error Rate (EER) was calculated to evaluate the system. The EER is the value where False Acceptance Rate (FAR) crosses the False Rejection Rate (FRR), and is typically used as a comparative measure within the biometric industry [28].

The multimodal experiment was conducted using all possible combination of three techniques. The results of each technique were combined at the matching-level - as each technique utilized different classifiers and a different range of outputs, the min-max score normalization method was applied to scale the results of each technique into the range between 0 and 1. Based upon prior research, two fusion methods were utilized: simple sum and matcher weighting [11], [29]. For the Simple Sum fusion, the raw score of each individual technique were simple added and rescaled into the 0 to 1 range. For the Matcher Weighting approach, weights are assigned to the individual matchers based on their individual EER. The weights are inversely proportional to the corresponding errors; the weights for less EER are higher than those of with a high EER.

3.2 Experiment results

The results of using individual biometrics and the multimodal approach are shown in Table 2. The results illustrated an average of all the users' EERs by using a single optimized neural network. The results showed that the individual techniques can be used to discriminate users with relatively low error rates for a good proportion of participants. However, further analysis showed that the individual user is able to achieve a better overall EER when each user is permitted to use a different network configuration. By using individually optimized network configurations for individual user, the overall performance was an EER of 8.9%. Behavioral profiling demonstrated the best individual performance using a single network configuration, with keystroke dynamics being the worst performer.

A further analysis of individual performances raises a number of interesting points. Foremost, that the best-case EERs are extremely good. However, it is noticeable that there are some users that experience very high error rates, reiterating the importance of multimodal approaches.

Table 2. Experiment results for text-based authentication

		Equal Error Rate (EER)%		
		Average	Best Case	Worst Case
Linguistic Profiling (LP)		12.8	0.0	40.0
Behavioral Profiling (BP)		9.2	0.0	50.0
Keystroke Dynamics (KA)		20.8	0.0	50.7
Fusion by Sum				
	Multimodal (LP+BP)	5.5	0.0	30.6
	Multimodal (KA+BP)	6.2	0.0	20.0
	Multimodal (LP+KA)	11.2	0.0	45.0
	All techniques	4.4	0.0	18.1
Fusion by Matcher Weighting				
	Multimodal (LP+BP)	3.6	0.0	20.0
	Multimodal (KA+BP)	5.3	0.0	20.2
	Multimodal (LP+KA)	8.5	0.0	44.7
	All techniques	3.3	0.0	19.3

As seen in the Table 2, both of the two fusion methods lead to better performance than any of the individual classifiers. Generally, the Matcher Weighting technique outperforms simple sum method. Whilst the results show that on average the use of more modalities leads to a better performance, this is not reflected within the individual user results. On occasions, it was noticed that users performed better when using two inputs (typically LP+BP) rather than three. Therefore in an operational environment case must be taken on selecting the most appropriate classifier. Examining the individual worst-case performance, it can be seen that the multimodal models have significantly improved upon the error rates – further supporting the use of multimodal approaches.

4 A novel framework for active authentication

The concept of Transparent Authentication System (TAS) on mobile devices was first proposed in 2002 [30]. The framework utilizes a mixture of biometric techniques to verify a mobile user's identity in a continuous and transparent manner. The framework is able to:

- to increase the authentication security beyond that offered by the password based approach;
- to provide transparent non-intrusive authentication for the user (rather than intrusive) to maximize user convenience;
- to provide continuous verification of the user, ensuring that the protection can be maintained throughout the duration of the device usage;
- to provide an authentication architecture that automatically works on all mobile devices regardless of hardware configuration, processing capability and network connectivity.

A number of process engines and a security manager have been devised to achieve these objectives (as demonstrate in Fig.1). A detailed description of these processes is presented in the following sections.

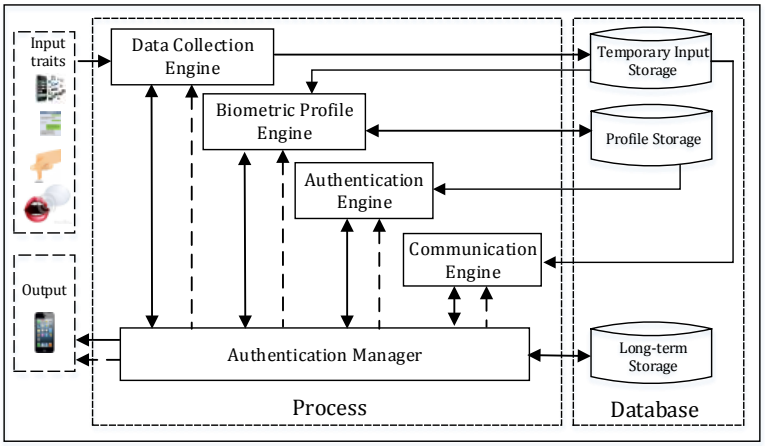


Fig. 1. Text-based multimodal framework

4.1 Processing engines

The primary role of the Data Collection Engine is to capture a user's input text. When a user utilizes a text-based application on the mobile, information about the user's typing, message writing style and the application usage are automatically collected by the Data Collection Engine and transformed into various biometric input samples. The captured input samples are then stored in the Temporary Input Storage to be used further in the authentication process by the authentication engine.

The main duty of the biometric profiling engine is to generate the various biometric profile templates by using the combination of the user's historical data and a number of template generation algorithms. The generated biometric templates will be stored in the Profile Storage and will be used in the verification process.

The main functionality of the Authentication Engine is to perform the user authentication process. The Authentication Engine has the ability to perform authentication for every permutation of inputs to ensure that authentication can be performed even if all of the three biometric samples are not presented (e.g. location may be not be able to be determined). When a verification process is required by the Authentication Manager, the Authentication Engine compares the input samples with the biometric templates to determine the legitimacy of the user. Once the verification process is completed, the verification result is appropriately processed by the Authentication Manager. If the verification result indicates the sample(s) came from authorized user, the sample(s) will be stored within the Profile Storage to be used for profile (re)generation; otherwise it will be deleted. A multibiometric authentication technique may produce a verification result that accepts the samples as coming from the authorized user even though the sample from one individual technique might be rejected as

coming from an imposter. Since the overall decision was that the sample comes from the authorized user, the failed samples are deemed to be, in fact, from the authorized user and incorrectly failed. As such, these samples are added to the profile and are not deleted. In this way, the template re-training process can produce a more accurate profile that could provide better performance. This process overcomes a fundamental issue with biometric template re-training and ensuring the correct inclusion of relevant samples.

The framework can operate in both standalone and distributed modes to allow the framework to be useful for non-wireless and wireless devices. If the framework operates in client-server mode, the communication engine works as a bridge between the capture device and the comprehensive framework. When the framework operates in a standalone mode and the device is locked down, the communication engine sends a code to the user which they can use to unlock their device.

4.2 Security manager

The Authentication Manager is the central controller of the framework and provides the "intelligence". The key task of the Authentication Manager is to monitor the security level and make authentication decisions when the user requests access to an application. It is the responsibility of the Authentication Manager to handle the security and user convenience trade-off. In order to achieve this, the Authentication Manager utilizes two processing algorithms: the System Security (SS) Level Automatic Update Algorithm and the Application Request Algorithm to manage the balance between the security of the mobile device and user convenience. These processes have been designed based upon a well-known study [24].

The SS level is a sliding numerical value in the range of 0 and +5 with 0 indicating a low security level and +5 indicates a high security level[1]. The SS level changes depending upon the outcomes of the authentication processes and the time that has elapsed between authentication requests. In this proposed framework, each application will have its own security level. The high value application will have a high security level and a normal application will have a low security level. This can be achieved either manually by the user or automatically by the system, using a database stored in the Long-term Storage. Prior research has investigated simple mechanisms by which these risk-based evaluations for applications can be made [30].

The Authentication Manager utilizes the SS Level Automatic Update Algorithm in order to periodically update the SS level based on the results of authentication decisions based upon the user's input samples. The Authentication Manager periodically sends an authentication request to the Authentication Engine in order to update the SS level. The time interval in which the authentication should be requested depends upon the user's preference (i.e. every 5 minutes). Initially, the Authentication Manager requires the Authentication Engine to perform authentication using the best set of the user's input samples (i.e. utilizes the classifier with the lowest EER that samples exist

[1] The boundaries defined on the numerical scale are only provided as a suggestion. In practice, these values maybe redefined.

for) from the last x minutes (i.e. 5 minutes). In a case where no user's input data is presented, the Authentication Manager maintains the SS level at its latest updated value. However, if the Authentication Engine responds with a pass then the Authentication Manager updates the SS level and subsequently reverts back to monitoring mode. If not, the Authentication Manager decreases the SS level and sends an authentication request again by using the next best set of user's input samples. The Authentication Manager will try three times to send an authentication request, every time with the next best available sample being employed. The Authentication Manager updates the SS level based upon the authentication result. The SS value is increased or decreased based on the type of sample used. For example, a sample using the keystroke dynamics technique will have an increment/decrement value of 0.5; a sample which contains both linguistic profiling and behavior profiling will have an increment/decrement value of 2. This numbers are based on the performance of the technique or combination of techniques. In scenarios where the updated SS level is less than 0, the Authentication Manager will set the SS level back to 0, meaning that the user will be able to access only the applications that do not required security. The process gives bias toward the user as they are given three non-intrusive chances to authenticate correctly and no intrusive authentication requests. This enables the system to minimize inconvenience to its user. Should the user attempt to access applications that require a SS level greater than the current SS level, the Authentication Manager will utilize the Application Request Algorithm to check the legitimacy of the user as shows in Fig.2.

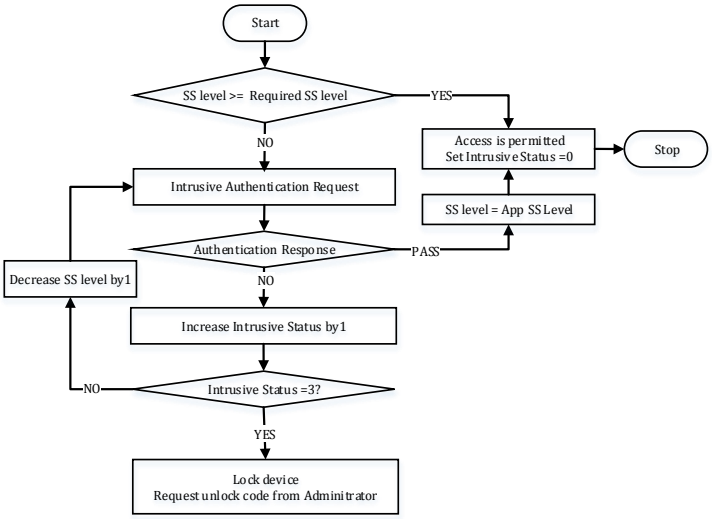


Fig. 2. Application Request Algorithm

The current SS level of the user is compared with the security level of the requested application. If the level is equal to or greater than the security level of the required application, the user can automatically access the application. Otherwise the user will be asked intrusively to authenticate. If the authentication response to this intrusive

request fails to pass, the device is locked. Otherwise, the level of the user will be updated to the security level of the requested application and access will be granted.

5 Evaluation

To examine the effectiveness of the framework in providing security and user convenience, the proposed framework was evaluated through a simulation. The simulation process involves implementing a virtual user and applying the SS Level Automatic Update Algorithm and the Application Request Algorithms.

To evaluate the performance of the security mechanisms to an authorized user, three different usage levels (infrequent, moderate and frequent) will be investigated - as the level of usage will have a direct impact on the availability of biometric samples and thus the capability of the system to maintain the security level. The use of the mobile device is simulated using a flow of timeslots. Each time slot can be seen as a minute in real life. Within each time slot the user can do one of two actions, or both: provide an input sample (thus simulating a text-based entry) or the use an app. Within each timeslot, the probability for the user to provide an input sample or accessing an application will set to 0.05, 0.15 and 0.50 in order to simulate an infrequent, moderate and frequent user respectively. There are 6 different types of application that can be chosen by the user (reflecting the possible security levels of an application from 0 to 5). Each type of application has the same probability of being accessed. Similarly, there are 7 different non-intrusive techniques (refer to Table 2). Given that within a time slot the user provides an input sample, each type of technique has the same probability of occurring.

All non-intrusive techniques are evaluated based upon the EER of each authentication technique as demonstrated in the experimental result section. This means that, when the system evaluates a sample, there is a probability (equal to the EER of the technique) that an authorized user will be rejected or an imposter will be authorized. With regards to the intrusive authentication requests, the probability of an authorized user and impostor being rejected and accepted respectively is set to 0.03. This approach to the methodology removes any bias and provides for a randomly generated dataset with a mix of samples, performances and application requests across three usage scenarios. To further remove any bias that would exist from a single run of the simulation, the simulation is repeated.

The security system will work as described in the Security Manager session. The SS will be updated every 10 minutes. If the mobile device is not used for 10 minutes consecutively, the SS will be decreased by 0.05 for every following minute, until the system is used again. The simulation simulated the use of the mobile phone for 12 hours or 720 minutes.

In order to examine the ability of the system security to prevent an imposter from using the mobile device, two scenarios were simulated: an imposter using a mobile device at the initial state (SS =0) and the imposter using a mobile device starting from a high level of security (SS=5). This can simulate an imposter taking control a mobile device which has just been used by the authorized user.

5.1 Simulation results

The result for all scenarios is represented using the average of running the simulation 10 times. The simulation results for an infrequent, moderate and frequent authorized user are presented in Table3.

Table 3. Simulation results for different types of authorized user

	Infrequent User		Moderate User		Frequent User	
App Level	**#App Request**	**#Intrusive Request**	**#App Request**	**# Intrusive Request**	**#App Request**	**# Intrusive Request**
5	7.2	4.2	16.2	1.5	60.0	1.50
4	5.1	0.4	17.9	0.5	60.1	0.50
3	7.3	0.3	20.0	0.2	61.3	0.30
2	5.9	0.2	16.8	0.3	59.6	0.10
1	6.0	0.0	19.5	0.2	55.2	0.00
0	6.7	0.0	19.3	0.0	57.6	0.00
Total	38.2	5.1	109.7	2.7	353.8	2.40

Based upon the simulation results, it can be shown that the security system can provide a high level of security whilst minimizing user inconvenience in all three scenarios. Analysing the proportion between intrusive authentication request and application access permits an insight into how often the user experiences an intrusive authentication request. Ideally, this proportion would be zero meaning that the user would not be required to perform an intrusive authentication request when they access an application. In our simulation these values are 13%, 2% and 0.6%, for the infrequent, moderate and frequent user respectively. The infrequent user experiences a higher intrusive request because it will probabilistically have fewer samples in the system and the system decreases the SS level if the device is not used for 10 consecutive minutes. Therefore, when this user want to access an application, it is more likely that its SS will not be sufficient to be granted immediate access. Throughout the complete 720 minute simulation the device was never incorrectly blocked for the authorized user. Further analysis of the results demonstrates for a level 5 app (which is arguably sensitive enough to warrant authentication of the user), this transparent approach results in a 97.5% reduction in intrusive authentication requests (for a frequent user).

The simulation results of the imposter scenarios showed that the security system blocks the imposter from using the mobile device after few minutes in both cases (as illustrated in Table 4). The reasons for this is that when the imposter tried to access an application that required a security level greater than 0, the system requested the imposter to authenticate themselves using an intrusive technique three times. There is a really small chance for the imposter to successfully authenticate, so after three requests the device will be blocked. As expected, the system will take more time to block the device if the imposter starts using the device when the SS is high.

Table 4. Simulation results for imposter user start using device at SS=0 and 5

	Device at SS= 0		Device at SS= 5	
App Level	**#App Request**	**# Intrusive Request**	**#App Request**	**# Intrusive Request**
5	0.6	0.4	1.0	0.2
4	0.3	0.1	1.8	0.6
3	0.4	0.4	1.5	0.0
2	0.4	0.2	1.3	0.1
1	0.3	0.1	1.9	0.1
0	0.4	0.0	1.0	0.0
Total	2.4	1.2	8.5	1.0
Time In Use	5.0 minutes		14.7 minutes	

5.2 Discussion

The simulations show how the proposed framework can provide a good compromise between improving the level of security provided and without increasing the user convenience. Indeed, it can be argued that user convenience under this model is also significantly improved over existing approaches. However, further investigations are required in order to better examine the values of the parameters. For example, it seems clear that the verification time does play an important role of providing security and user convenience. By regularly authenticating the user, the user will suffer more intrusive authentication requests but the system will be able to recognize an imposter in a relatively short period of time. On the other hand, users will find the device more convenient to use with longer time periods between user authentications but the system will take longer to recognize an imposter and lock down the system. In our simulation the verification time was 10 minutes. However, this may not be the optimum compromise between convenience and security.

Similarly, decreasing SS level was not examined, but it is expected to play a relevant role in the system. The infrequent user will experience less challenges from the intrusive authentication technique when the time period of the degradation function gets longer. However, the imposter will have more chance of accessing a high level application in cases where the device was initially left with a high level SS. In this simulation, a linear function is used to decrease the SS level but it is suggested that the function for degrading the SS level should be implemented using an exponential function as it decrease slowly at first and then more rapidly.

6 Conclusions & Future Work

The first part of this paper presented a feasibility study that demonstrated the ability of utilizing text-based entry to authenticate users. The use multimodal biometrics, specifically the combination of linguistic profiling, behavior profiling and keystroke dynamics showed an excellent level of recognition performance, validating the feasi-

bility that multimodal text-based has the ability to authenticate user on mobile devices.

The novel multimodal authentication framework subsequently presented to support text-based biometrics was designed to add additional security to a mobile handset, providing transparent and continuous authentication. The system is designed using a variety of single and multimodal biometric techniques without any additional hardware. The users can benefit from the framework in terms of both device security and convenience of use. By setting various security requirement levels for different applications/services based upon their risk, the framework is capable of controlling the impact on each application/service. The simulation results clearly showed that the proposed authentication framework is able to provide continuous and transparent authentication to protect mobile devices.

Future work will focus upon the development of a more representative and larger biometric corpus from which to further examine the level of recognition performance that can be achieved. To accompany this work, an operational prototype will also be developed to enable an end-user evaluation to be undertaken so that user acceptance and operational performance can be established.

7 References

1. Ericsson.: Traffic and market report on the pulse of the networked society, http://www.ericsson.com/res/docs/2012/traffic_and_ market_report_june_2012.pdf
2. Kaspersky Lab.: European Users Mobile Behaviour and Awareness of MobileThreats, http://www.kaspersky.com/news?id=207576289
3. Dimensional Research.: The impact of mobile devices on information security: A survey of IT professionals, http://www.checkpoint.com/downloads/products/check-point-mobile-security-survey-report.pdf
4. McAfee.: McAfee Reveals Consumers Fail To Protect Their Mobile Devices, http://www.mcafee.com/us/about/news/2013/q1/ 20130224-01.aspx
5. Clarke, N. and Furnell, S.M.: Authentication of users on mobile telephones – A survey of attitudes and practices, Computer & Security, vol.24, no. 7, pp519-527, (2005)
6. Karatzouni S., Clarke, N. and Furnell, M.: Utilising Biometric for transparent user authentication on mobile devices. In: 2nd Internet Technologies and Applications, pp.549-557 (2007)
7. Li, Fudong, Nathan Clarke, Maria Papadaki, and Paul Dowland.: Behaviour Profiling for Transparent Authentication for Mobile Devices. In *Proceedings of the 10th European Conference on Information Warfare (ECIW), Tallinn, Estonia*, pp. 307-314. (2011)
8. Sim, T., Zhang, S., Janakiraman, R., & Kumar, S.: Continuous verification using multimodal biometrics. In: Pattern Analysis and Machine Intelligence, vol 29, no 4, pp. 687-700. (2007)
9. Kittler, J.,Matas, J., Jonsson, K. and Ramos Sanchez, M. U.: Combining Evidence in Personal Identity Verification Systems. Pattern Recognition Letters, vol.18, pp.845-852 (1997)
10. Poh, N. and Korczak, J.: Hybrid Biometric Authentication System Using Face and Voice Features. Lecture Notes in Computer Science, vol.2091/2001, pp. 348-353 (2001)
11. Snelick, R., Uludag, U., Mink, A., Indovina, M., and Jain, A.K.: Large-Scale Evaluation of Multimodal Biometric Authentication Using State-of-the-Art Systems. IEEE Transactions on pattern analysis and machine intelligence, Vol.27, no. 4 pp.450-455, (2005)

12. Cmo Council, http://www.fastcompany.com/3010237/bottom-line/texting-is-the-new-email-does-your-company-do-it-right
13. ComputerWeekly,http://www.computerweekly.com/news/2240205200/Apple-adopts-hands-off-approach-to-iPhone-fingerprint-scanner
14. MIT Technology Review, http://www.technologyreview.com/news/425805/new-google-smart-phone-recognizes-your-face/
15. Rudman, J.: The state of authorship attribution studies: Some problems and solutions. Computers and the Humanities, 31, 351-365. (1998)
16. Halteren, V. H.: Linguistic Profiling for Author Recognition and Verification, In: 42nd Annual Meeting on Association for Computational Linguistics (ACL 04), Association for Computational Linguistics, Morristown, NJ, USA, (2004)
17. Zheng R., Li, J., Chen, H., and Huang Z.: A Framework for Authorship Identification of Online Messages: Writing-Style Features and Classification Techniques. Journal of the American Society for Information Science and Technology, vol. 53, pp. 378-393. (2006)
18. Boukerche, A., Nitare,M.S.M.A.: Behavior-based intrusion detection in mobile phone systems. J. Parallel Distrib. Comput. **62**(9),1476–1490 (2002)
19. Damopoulos, D. Menesidou, S. Kambourakis, Papadaki, M. Clarke, N. Gritzalis, S. Evaluation of Anomaly-Based IDS for Mobile Devices Using Machine Learning Classifiers, Security and Communication Networks, Vol. 5, No. 1, pp. 3-14, 2012, Wiley
20. Buschkes, R., Kesdogan, D., Reichl, P.: How to increase security in mobile networks by anomaly detection. In: Proceedings of the 14th Annual Computer Security Applications Conference, pp. 3–12 (1998)
21. Hall, J.,Barbeau,M., Kranakis, E.: Anomaly based intrusion detection using mobility profiles of public transportation users. In: Proceeding of IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, vol. 2, pp. 17–24 (2005)
22. Biopassword.: the keystroke dynamics approach, http://www.biopassword.com/bp2/welcome.asp.
23. Behaviosec, http://www.behaviosec.com/products/enterprise/
24. Clarke, N. and Furnell, S.M.: Authenticating Mobile Phone Users Using Keystroke Analysis, International Journal of Information Security, ISSN: 1615-5262, pp.1-14. (2006)
25. Indovina, M., Uludag, U., Snelick, R., Mink, A., & Jain, A.: Multimodal biometric authentication methods: a COTS approach. *Proc. MMUA*, 99-106. (2003)
26. Eagle, N., Pentland, A., Lazer, D.: inferring Social Network Structure using Mobile Phone Data, Proceeding of National Academy of Sciences (PNAS),vol.106, pp.15274-1578, (2009)
27. Ashbourne, J.: Biometric, Advanced identity verification. The complete guide. Springer, (2000)
28. Jain, A. K. Nandakumar, K. and Ross, A.: Score normalization in multimodal biometric systems. Pattern Recognition, vol. 38, no. 12, pp.2270-2285, Dec. (2005)
29. Clarke, N., Furnell, S.M. and Reynolds P.L.: Biometric Authenticating for Mobile Devices. In: 3rd Australian Information Warfare and Security Conference, Western Australia, (2002)
30. Ledermuller, T. and Clarke, N.L.: Risk assessment for mobile devices. In: 8th International Conference Privacy and Security in Digital Business, TrustBus, pp. 210–221 (2011)

Multi-Modal Behavioural Biometric Authentication for Mobile Devices

H. Saevanee[1], N.L Clarke[1,2] and S.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
e-mail: 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

adfa, p. 1, 2011.
© Springer-Verlag Berlin Heidelberg 2011

is commonly acknowledged as a reliable solution which provides enhanced authentication over the traditional password ("something you know") and token ("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 dis-

criminative 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 authentica-

tion 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.

7 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. and Furnell, S.M.: Advanced user authentication for mobile devices, Computer and Security, vol. 26, pp.109-119 (2007)
3. Brunelli, R. and Falavigna, D.: Personal Identification using Multiple Cues. In: IEEE Transaction on Pattern Analysis and Machine Intelligence, pp. 955-966(1995)
4. Kittler, J., Matas, J., Jonsson, K. and Ramos Sanchez, M. U.: Combining Evidence in Personal Identity Verification Systems.Pattern Recognition Letters,vol.18,pp.845-852 (1997)
5. Poh, N. and Korczak, J.: Hybrid Biometric Authentication System Using Face and Voice Features. Lecture Notes in Computer Science, vol.2091/2001, pp. 348-353 (2001)
6. Ross, A., Jain A. and Qian, J-Z.: Information Fusion in Biometrics. In: Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition, Springer-Verlag, pp. 354-359, UK(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. and 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. and Dowland, P.: Behaviour profiling on mobile devices. In: International Conference on Emerging Security Technologies, pp.77-82, UK (2010)
14. Li, F., Clarke, N., Papadaki, M. and Dowland, P.: Behaviour profiling for Transparent Authentication for Mobile Devices. In: 10th European Conference on Information Warfare and Security, pp307-314, Estonia (2011)
15. Obaidat, M.S. and Sadom, B.: Verification of Computer Users Using Keystroke Dynamics. In: IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol.27, pp.261-269 (1997)
16. Clarke, N. and Furnell, S.: Authenticating Mobile Phone Users Using Keystroke Analysis. Information Security, pp.1-4 (2006)

17. Karatzouni, S. and Clarke, N.: Keystroke Analysis for thumb-based Keyboards on Mobile Devices. In:22nd IFIP International Information Security Conference, Springer, pp. 253-263, Sandton (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, pp.199-206,NJ(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, pp229 (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, p224-229 (2011)
26. Ashbourne, J.: Biometric, Advanced identity verification. The complete guide. Springer, (2000)

SMS Linguistic Profiling Authentication on Mobile Devices

H. Saevanee[1], N.L Clarke[1, 2] and S.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
email: info@cscan.org

Abstract— **It is commonly acknowledged that mobile devices now form an integral part of an individual's everyday life. As the amount of valuable and sensitive information stored on a mobile device increases, so does the need for effective security. In order to protect unauthorised access, an authentication system is required. Biometric authentication has proven to be a more reliable solution than knowledge based and token based techniques. Indeed, biometric techniques are uniquely individual, impossible to forget or lose, difficult to reproduce or falsify and difficult to change or hide. Despite the well known advantages of biometric authentication approaches, the majority of current state-of-the-art mobile devices embrace point of entry authentication systems including PIN/passwords and one time fingerprint verification. This paper introduces a feasibility study into a novel biometric technique for mobile devices, linguistic profiling. This investigation sought to authenticate users based on their writing vocabulary and style of SMS messages. The findings, based upon 30 participants, revealed the feasibility of the approach. While an overall average Equal Error Rate (EER) of 24% is unacceptably high, several users experienced an EER of 0%, suggesting significant potential to apply the technique for a subset of the population.**

Keywords-component; linguistic pofiling; user authentication; biometric; mobility

I. INTRODUCTION

The popularity of the Short Messaging Service (SMS) is one of the recognised successes of mobile communications, with over 6.1 trillion messages sent in 2010; close to 200,000 messages sent every second [1]. Meanwhile, the plethora of functionalities offered by mobile devices enables users to store increasing amounts of wide ranging types of data. Indeed, a survey of 4,000 UK and US participants revealed the following types of data were stored on end-users handsets: telephone numbers (94%) address and other contact information (65%), digital photos (83%), videos (51%), calendar information (48%) and music downloads (40%) [2]. The ever increasing network and phone capabilities, has resulted in the increase in end-users using mobile devices to store personal and sensitive data. With this in mind, previous research [2] highlights mobile users concerns of their devices being lost or stolen with 82% of respondents being worried that the personal and sensitive data stored on their device is used for fraudulent purposes. To this end, 90% of respondents were worried about the loss of crucial data in the event their mobile device is lost or stolen. The growing amount of personal and sensitive data now being stored on mobile devices juxtaposed with the increased number of fraudulent incidences there is a corresponding need for an enhanced authentication system which effectively protects valuable data from unauthorised access.

Authentication methods for mobile devices can be classified based on three fundamental mechanisms: Something you know, something you have and something you are [3]. The first mechanism is based on something you know such as using PIN (Personal Identification Number) or Password. This technique is, at present, the most common authentication method and offers a standard level of protection whilst providing cheap and instantaneous authentication. The second mechanism is based on something you have such as token or SIM. However, the problems with these two techniques arise when the passwords or tokens are lost, stolen, or misplaced, which eventually leads to fraudulent incidents. For example the selection of weak passwords that are easy to guess or carrying the SIM card around with the mobile device to present in order to access the device. A reliable solution to these authentication problems lies in the last mechanism; something you are. Biometric characteristics are uniquely individual, non-transferable to others, impossible to forget or lose, difficult to reproduce or falsify, usable with or without the knowledge/consent of the individual difficult to change or hide. Hence, it is known to be the most secure and reliable way to establish authentication more than password and token authentication mechanisms.

Current approaches are still focused upon point-of-entry authentication (e.g. PIN/passwords, fingerprint), which has a number of weaknesses. Moreover, in case of user chose to not using authentication at the first place or once the identity of the user has been verified at login, the mobile system are typically made available to user until the user exit the system. This can be lead to high risk environment which imposter targets a post authenticated session. To increase the level of authentication beyond the standard point-of-entry technique, studied [4] proposed a framework that uses a combination of secret based knowledge and behavioural biometric techniques which can provide transparent or non-intrusive and continuous authentication that authenticate user without knowledge of user

and authenticate user during usage of device rather than simply at switch-on. Among various behavioural biometric techniques, it is hypothesised that one of the possible techniques that can provide a cost-effective, non-intrusive and continuous solution for mobile devices authentication is linguistic profiling – identifying individuals based upon their writing vocabulary and style within SMS messages

The aim of this paper is to investigate the feasibility of linguistic profiling based on SMS messages, and the results of applying this novel technique as an authentication method for mobile devices. Section II provides an overview of linguistic profiling. Sections III and IV describe the methodology and results of the study. Finally section V and VI discusses the implications of the results together with conclusions and recommendations for future work.

II. RELATED WORKS

Linguistic profiling is a behavioural biometric that attempts to identify and discriminate between users based on linguistic morphology [5]. The majority of the studies conducted to date have investigated the feasibility of linguistic profiling for authorship identification in text documents and forensic analysis for criminal legal proceedings [5, 6, 7, 8, 9, 10]. 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 [5], [8, 9, 10, 11, 12]. Indeed, using combination of features revealed more successful discriminate users as previous studied [5] showed using the best combination features outperform the best individual feature. Moreover, studied [12] proved a combination of features based on shallow linguistic analysis and set of deep linguistic analysis features yields very high accuracy in a short random text.

The underlying classification algorithms utilized in linguistic profiling were analysed based on statistic and advance neural networks. A summary of literature and results of linguistic profiling in text documents was illustrated in Table1.

All of the studies have illustrated the potential of the technique, with [10] performing the best with 100% accuracy with using neural network classification algorithm. The overall of the results showed very promising results with a minimum of 80% accuracy. The best performance of systems all used machine learning approach but different choices of techniques and features used. Due to the focus on linguistic profiling based on long text message (i.e. email messages, student essays and book chapters), little work to date has been done on SMS message, which has limited space. In 2010, studied [13] utilized linguistic profiling of SMS messages to identify the authors for forensic investigations. Authorship attribution of SMS messages using N-grams approach revealed positive results, achieving the best accuracy 72%. However, a distinct lack of published literature on SMS message highlights the need for further research.

TABLE I. A SUMMARY OF LITERATURE AND RESULTS OF LINGUISTIC PROFILING IN TEXT DOCUMENTS

Document types	# of participants	Best Accuracy
Email Messages	5	85.7%
Email Messages	2	99%
Studuent Essays	8	97%
Book Chapter	2	99%
Email Messages	134	80%
Email Messages	12	100%

III. METHODOLOGY

The overall process of designing a linguistic profile authentication system of this study is illustrated in figure 1.

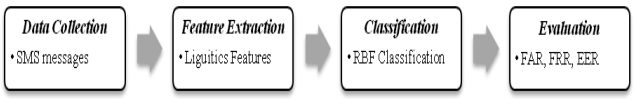


Figure 1. The Process of Linquistic Profiling authentication System

In order to analyse the effectiveness of using linguistic profiling, a real time SMS simulation program was compiled in Visual Basic and deployed onto two Smartphones (XDA II and SPV M2000). The program enabled text based communication between pairs of participants using a Bluetooth communication channel. Bluetooth was used to simulate a GSM based SMS service. Figure 2 showed Smartphone and user interface for this experiment.

(a) An XDA IIs Smartphone (b) Screenshot of user interface of experiment software

Figure 2. Mobile phone and user interface for the experiment study

In line with previous studies [4], [14], 30 participants were deemed an appropriate number. Whilst a greater number of participants would have been preferred, a trade-off exists with the availability of participants. As such a total of 30 participants were required to send at least 15 messages to each

other using a non-predictive text input method. Predictive text input was disabled in order to prevent automatic correction and adjustment of spelling and grammar and the consequential negative effect upon the discriminative information between users. Participants were encouraged to discuss any subject of their choice during the experiment in order maximise the potential level of detail and complexity within each transmitted message and to increase the likeliness of users adopting their own unique texting style and linguistic composition. The messages were saved in a text file prior to transmission.

Short Message Services (SMS) language tends to use syntactic and lexical short forms [15]. The language has developed its own unique style as email and chat-room languages [16]. Abbreviation and emotional words are commonly used in text messages. Abbreviation, which commonly uses single letters, numbers or combined to represent the whole word (for example, 'See You' can be texted as 'CU', 'Mate can be texted as 'M8'). Emotion or verbal effect represents body language such as ':)' for 'smiley face', 'hehe' or 'haha' for laughter. Therefore, the frequency distribution of abbreviations, emotional and user's favourite words were used to create user profiles, including every possible type of feature. For each message, a total of 64 discriminating characteristics were extracted as shown in Table 2. Two thirds of the dataset was used for training (316 messages), and the remaining third used for testing and evaluating the feasibility (171 messages).

TABLE II. A SUMMARY OF LITERATURE AND RESULTS OF LINGUISTIC PROFILING IN TEXT DOCUMENTS

Average Sentence Length
Average Word Length (number of characters)
Message Length
Message Length/ Total number of sentences
Total number of sentences
Total number of words in message (M)
Total number of words/ Total number of sentences
Total number of characters (C)
Total number of alphabetic characters
Total number of digit characters
Total number of capital characters
Total number of symbols in message
Total number of space after punctuation
Total number of punctuation after space
Total number of no space after Punctuation
Total number of space characters
Total number of symbols (20 features)
Total number of word length distribution (11 features)
Total number of word length distribution/M (11 features)
Total number of long words (word length > 8)
Total number of alphabetic characters/C
Total number of digit characters/C
Total number of capital characters/C
Total number of spaces/C
Total number of symbols/C

To study the feasibility of using linguistic profiling to profile and discriminate between users, descriptive statistics were utilised to illustrate the nature of the data and to provide the designer with sufficient information over its complexity and the resulting classifier required. From that analysis, an advanced neural network known as Radial Basis Function (RBF) was selected to classify users. In this study, the Equal Error Rate (EER) was calculated to evaluate the system. The EER is the value where False Acceptance Rate (FAR) is equal to False Rejection Rate (FRR), and typically used as a comparative measure within the biometric industry [17].

IV. RESULTS

In this experiment, user word profiling and linguistic features were purposed for use in authenticating users. To create word profiling, abbreviation and emotional keywords were selected. An initial analysis, frequency of keywords occurrence was calculated. Figure 3 presents each users frequency of some keywords, providing an insight into the inter and intra-class variance and therefore providing an estimate of the likely discriminative information present between users.

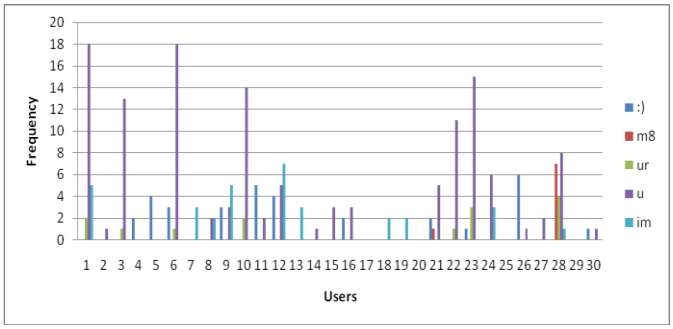


Figure 3. Frequency of words profiling of each users

As shown in Figure 3, the results from initial analysis indicate that common keywords have many users used the same words. For example, users 1, 8, 9, 12 and 24 used the same abbreviation 'u' and 'im'. However, some keywords have only two users, for example (users 28 and 30) used the same words such as abbreviation 'm8'. To analyse individual user's performance, a number of analyses were undertaken, using Radial Basis function (RBF) neuron network algorithm. Different network configurations were tested, looking for optimum performance. The best results of each user are shown in Figure 4. Table III shows this was achieved by using abbreviation and emotional based profiling.

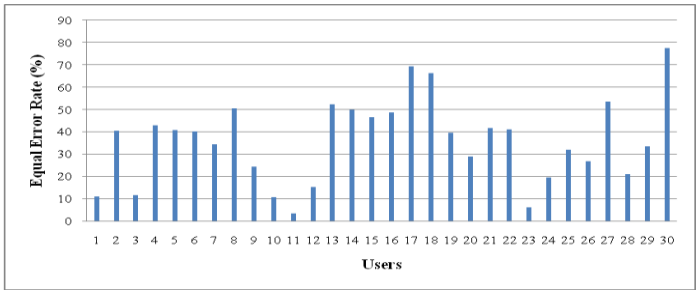


Figure 4. the Best Results of RBF Classification for each User

TABLE III. EXAMPLE OF USER WORD PROFILING

User #	User's word Profile
11	'x','u','ill', 'yeah','lol','hmm','☺','ah'

23	'r','nop','ab','c','ar','T T','bc','lots', 'u','ur','im', ‘haha’, ‘oh’, ‘wow’, ‘☺’
17	'x','q','mmm','huzzah','vessel', ‘lol’, ‘yeah’, ‘argh’, ‘ah’
18	'im','theyre', ‘oh’ , ‘hurray’, ‘yeah’

As shown in Figure 4. The finding of using words profiling to discriminate between users showed positive results. The best individual result is user 11 with the lowest EER 3.3%. However, some users generated a high EER. To overcome this issue, additional linguistic features were added to reduce error rate, in line with previous studies [5], [12]. In Figure 5, a comparison between the use of word profiling features and a combination of linguistic features is shown. The combination of linguistic profiling features results illustrated in Figure 5 is based on the features detailed in Table IV.

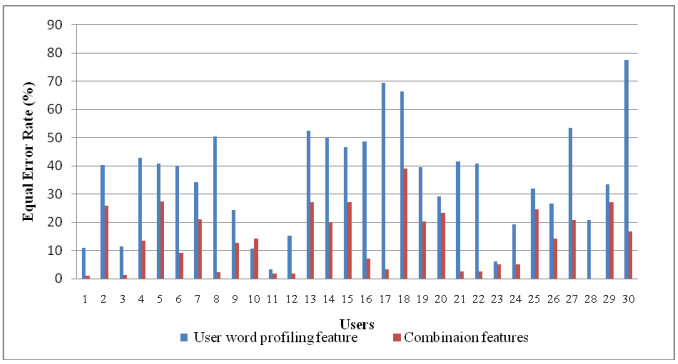


Figure 5. The Best Results of RBF Classification for each User

TABLE IV. INDIVIDUAL COMBINATION SET OF LINGUISTIC PROFILING

User #	User's word Profile
1	AbbProfile,EmoProfile,NoS,NoSpaceAfterP
2	AbbProfile,EmoProfile,NoS,NoWS
3	AbbProfile,EmoProfile,CapitalS,NoS
4	AbbProfile,EmoProfile,Comma,FullStop,SpaceAfterP,CapitalS
5	AbbProfile,EmoProfile,NoS,NoWS,CapitalS
6	AbbProfile,EmoProfile,Apostrophy,QuestionMark
7	AbbProfile,EmoProfile,NoWS,Ellipsis
8	AbbProfile,EmoProfile,NoWS,Space
9	AbbProfile,EmoProfile,NoWS
10	AbbProfile,EmoProfile,NoWS,Comma,QuestionMark
11	AbbProfile,EmoProfile,Ex,Star,Symbol,Wmorethan9
12	AbbProfile,EmoProfile,FullStop,Comma,QuestionMark
13	AbbProfile,EmoProfile,Wmorethan9,NoS,Symbol,FullStop
14	AbbProfile,EmoProfile,Wmorethan9,Symbol,FullStop
15	AbbProfile,EmoProfile,NoS,SpaceAfterP,Symbol,NoWS
16	AbbProfile,EmoProfile,CapitalS,FullStop,NoS
17	AbbProfile,EmoProfile,CapitalS,Apostrophy,Ellipsis,Comma
18	AbbProfile,EmoProfile,NoS,NoWS
19	AbbProfile,EmoProfile,NoS,Symbol,Capital
20	AbbProfile,EmoProfile,NoS,Apostrophy
21	AbbProfile,EmoProfile,Ex,SpaceAfterP,NoWS
22	AbbProfile,EmoProfile,PAfterSpace
23	AbbProfile,EmoProfile,DoubleEx
24	AbbProfile,EmoProfile,NoSpaceAfterP,FullStop,QuestionMark,Wother
25	AbbProfile,EmoProfile,DoubleEx,Comma,Dash,SpaceAfterP,Symbol
26	AbbProfile,EmoProfile,Comma,Dash,Apostrophy,NoS
27	AbbProfile,EmoProfile,NoS,Comma,SpaceAfterP
28	AbbProfile,EmoProfile,Ellipsis,NoSpaceAfterP,NoS
29	AbbProfile,EmoProfile,NoS,NoWS,SpaceAfterP,Symbol
30	AbbProfile,EmoProfile,FullStop,SpaceAfterP,NoWS

V. DISCUSSION

The feasibility study of user word profiling revealed that participants used differing sets of abbreviations and emotional words. Participants (e.g. user 11, user 23) who used a greater number of unique abbreviated or emotional words produced significantly lower error rates. Participants who used the same common abbreviation and emotional words (e.g. ‘im’, ‘u’) tended to produce higher error rates. As shown in Figure 4, using only the user word's profiling feature can be used to authentication some users. However, there are some users that generated an EER of more than 50%. Future analysis showed that adding linguistic features can improve the ability of discrimination as shown in Figure 5. The best results of each user were found by using a combination of linguistic features rather than user word profiling feature. The results show that using a greater number of linguistic features does not necessarily lower the EER. The number of linguistic features required for discrimination significantly differs between users (e.g. User 1 requires 4 different linguistic features whereas User 4 requires 6 different linguistic features to achieve a low EER) (see Table IV). Furthermore, it should be noted that certain participants (e.g. User2 and User18) use the same linguistic combinations, but produce different EER results (see Figure 5). Although the linguistic features are the same, the Number of Sentences (NoS) and Number of Words per Sentence (NoWS) values differ, thereby producing different Equal Error Rates. Therefore, selection of keywords and effective features is a critical process that can result in increased accuracy.

Figure 5 illustrates, that in some instances, the EER is lowered by the addition of linguistic features. All users achieved an improved EER of less than 40% and the best individual result is user 28 with the lowest Equal Error Rate (ERR) 0%. The overall EER of using the combination set of

linguistic features also reduced from 36% to 24%, thus increasing overall accuracy. Additionally, using a combination of structural and stylometric features to authenticate a user achieves lower error rates and higher accuracy rates than using single N-grams features approach [13].

Of the various behavioural biometric techniques, the use of linguistic profiling showed lower error rates than using the keystroke dynamics proposed by [4]. As such, the results illustrate that linguistic profiling has significant potential to authenticate users on mobile device providing transparent, non-intrusive and continuous authentication. However, it is possible to combine these techniques together to provide greater reliability than using just one single biometric technique.

The limitation of this study was manually selecting key abbreviation words, emotional words, user favourite words and sets of linguistic features to discriminate between users for each participant. Further analysis is required in order to automatically select the best discriminating features; however, it is envisaged a rule-based expert system could be devised to achieve this.

VI. CONCLUSIONS

This research investigated the behavioural biometric technique, linguistic profiling based on SMS messages, looking to apply this technique as an authentication method for mobile devices. The results showed that linguistic profiling can be successfully used to authenticate user with low error rates based on large number of participants.

The potential advantages of behavioural biometrics is that they can be utilised non-intrusively and continuously during the user session. In this way, it becomes possible to extend the authentication process beyond initial point-of-entry and verify the identity in a transparent manner, without the explicit involvement of the user. Moreover, the collection of behavioural data often does not require any special hardware and is so very cost effective.

However, individual biometric techniques such as linguistic profiling are not suited to all users and scenarios and also cannot provide adequate reliability. The next step in this research is to further analyse the behavioural biometric technique in order to provide adequate security and greater reliability for an enhanced authentication system. Once complete, the findings will be used to develop a multi-tier behavioural biometric authentication system to provide reliability and security to a wide range of mobile devices within any given mobile environment. The proposed framework 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] ITU, " ", http://www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf.

[2] Edison, "Mformation-sponsored survey reveals mobile users worried about loss and mobile fraud", http://www.mformation.com/mformation-news/press-releases/mformation-sponsored-survey-reveals-mobile-users-worried-about-loss-and-mobile-fraud.

[3] C. Kolodgy, "Biometrics: You Are Your Own Key", http://citm.utdallas.edu/research/Publications/white_papers_source/Biometrics.pdf

[4] N. Clarke, and S.M. Furnell, "Advanced user authentication for mobile devices", Computer and Security, vol. 26, pp.109-119, March 2007.

[5] H. Halteren, "Linguistic Profiling for Author Recognition and Verification", Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics (ACL 04), Association for Computational Linguistics, Morristown, NJ, USA,

[6] O. Vel, "Mining E-mail Authorship", Proceedings of sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 20-23, Boston, MA, USA

[7] O. Vel, A. Anderson, M. Corney, and G. Mohay, "Mining E-mail Content for Author Identification Forensics", ACM Sigmod Record, vol. 30, pp 55–64, ACM New York, NY, USA.

[8] S. Argamon, M. Saric, and S. Stein, "Style Mining of Electronic Messages for Multiple Authorship Discrimination: First Results", Proceeding of the ninth ACM SIGDD international conference on Knowledge discovery and data mining, Washington, DC, USA

[9] R. Goodman, M. Hahn, M. marella, C. Ojar and S. Westcott, "The use of stylometry for email author identification: a feasibility study", Proceedings of Student/Faculty Research Day, CSIS, Pace University, May 2007

[10] K. Calix, M. Connors,D. Levy, H. Manzar, G. McCabe and S. Westcott, "Stylometry for E-mail Author Identification and Authentication", Proceeding of CSIS Research Day, Pace University, May 2008.

[11] M. Koppek, J. Schler, and D. Mughaz, "Text Categorization for Authorship Verification ", Proceedings of the 21 International Conference on Machine Learning, Banff, Canada

[12] M. Gamon, "Linguistic correlates of style:authorship classification with deep linguistic analysis features", Proceeding of the 20th international conference on Computational Linguistics (COLING'04), Association fror Computational Linguistics, Stroudsburg, PA, USA

[13] A. Mohan, M. Baggili, and K. M. Rogers, "Authorship attribution of SMS messages using an N-grams approach", https://www.cerias.purdue.edu/assets/pdf/bibtex_archive/2010-11.pdf

[14] F. Li, N. Clarke,M. Papadaki, and P. Dowload "Behaviour profiling on mobile devices", International Conference on Emerging Security Technologies, pp.77-82,Canterbury, UK

[15] N. Doring, "Abbreviation and acronyms in SMS communication", http://www.nicola-doering.de/.

[16] M.S. Rafi, "SMS Text Analysis: Language, Gender and Current Practices", On Line Journal of TESOL France. http://www.tesol-france.org/.

[17] J. Ashbourne, "Biometric", Advanced identity verification, The complete guide , Springer, 2000

Behavioural Biometric Authentication for Mobile Devices

H. Saevanee, N.L Clarke and S.M Furnell

Centre for Security, Communications and Network Research, University of Plymouth, Plymouth, United Kingdom

e-mail: hataichanok.saevanee@plymouth.ac.uk

Abstract

It is commonly acknowledged that mobile devices now form an integral part of an individual's everyday life. As the amount of valuable and sensitive information stored on a mobile device increases, so does the need for effective security. In order to protect unauthorised access, an authentication system is required. Biometric authentication has proven to be a more reliable solution than knowledge based and token based techniques. Indeed, biometric techniques are uniquely individual, impossible to forget or lose, difficult to reproduce or falsify and difficult to change or hide. Despite the well known advantages of biometric authentication approaches, the majority of current state-of-the-art mobile devices embrace point of entry authentication systems including PIN/passwords and one time fingerprint verification. This paper begins by discussing the limitations of current state-of-the-art authentication approaches together with preliminary findings of an experiment using finger pressure authentication. The potential of linguistic profiling is explored and results of a secondary experiment highlighted the potential for a continuous, transparent and non-intrusive behaviour biometric authentication system for mobile devices.

Keywords

Authentication, Transparency, Behaviour, Biometrics, Mobile Devices.

1. Introduction

Presently, Mobile cellular networks are currently available for 90% of the world population including people living in rural areas. Mobile devices are widespread in excess over 5 billion users (ITU, 2010). Moreover, people are moving rapidly from 2G to 3G platforms in both developed and developing countries with 48 countries offering 3G service in last 3 years. With the increasing functionality of mobile devices the number of services, applications and information accessible to the user is significantly expanding. Moreover, Mformation (2009) highlights that mobile phone are becoming more sophisticated as significant amount of information is now stored on such devices. Mformation (2009) conducted a survey amongst 4000 people who were living in the UK and US. 90% of users store telephone numbers while 65% also store address and other contact information and 48% have calendar information on their phones including digital photos, videos, music downloads. With ever-increasing phone and network capabilities, this trend of using the phone to store

valuable and sensitive data from every aspect of life is set to continue. However, one of users' biggest concerns is the risk of their device being lost or stolen. According to the same report (Mformation, 2009), 82% of users fear that if their phones were lost or stolen, someone would use the information stored on them for fraudulent purposes; 90% were worried about the loss of their personal data if a mobile device were to go missing and 40% of respondents stated that losing a mobile would be worse than losing their wallet. In order to protect valuable and sensitive information from unauthorized users, there is an ever increasing need for an effective continuous authentication system.

The authentication on mobile devices can be classified based on three fundamental mechanisms: Something you know, something you have and something you are. The first mechanism is based on something you know such as using PIN (Personal Identification Number) or Password. This technique is currently mainly implemented and offers a standard level of protection and provides cheap and quick authentication. The second mechanism is based on something you have such as token or SIM. However, the problems with these two techniques arise when the passwords or tokens are lost, stolen, or misplaced, which eventually leads to fraudulent incidents. For example the selection of weak passwords that are easy to guess or carrying the SIM card around with the mobile device to present in order to access the device.

A reliable solution to these authentication problems lies in the last mechanism, which is something you are. Biometric characteristics are uniquely individual, non-transferable to others, impossible to forget or lose, difficult to reproduce or falsify, usable with or without the knowledge/consent of the individual, difficult to change or hide. Hence, it is known to be the most secure and reliable way to establish authentication more than password and token authentication mechanisms. However, biometrics introduce their own set of problems, in terms of performance and acceptability. Moreover, current approaches are still focused upon point-of-entry authentication, which is arguably still considerably inconvenient to the end-user.

The aim of this paper is to investigate the behavioural biometric techniques that can be applied to authentication on mobile device in a transparent, continuous and non intrusive fashion as to minimise user inconvenience and enable authentication to be performed beyond point-of-entry (e.g. PIN/passwords). Section 2 provides an overview of biometric authentication. Sections 3 and 4 then explore the feasibility of some specific behavioural biometrics, examining finger pressure and linguistic profiling. Section 5 discusses the potential biometric techniques for use in mobile authentication. Finally, conclusions and future work are presented in Section 6.

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" (IBG, 2006). 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. Some existing biometric characteristics are show in Table 1.

	Biometric Characteristic	Description of the features
Physiological	Fingerprint	Fork and ridge patterns, pore structure
	DNA	DNA code as the carrier of human hereditary
	Facial geometry	The overall structure, shape and proportions of the face: distance between the eyes, the location of the nose and eyes, the area surrounding the cheekbones
	Iris	Iris pattern
	Retina	Pattern of vein structure at the back of the eye
	Hand geometry	Features related to a human hand, finger length, width, thickness and curvatures
Behavioural	Signature (dynamics)	Pressure, direction, acceleration and the length of the strokes, dynamics number of strokes and their duration
	Voice	The Voice tract and the accent
	Keystroke dynamics	Rhythm of keyboard strokes
	Odour	Chemical composition of the one's odour

Table 1: Summary of the most well known biometric characteristics

Physical features are likely to stay more constant over time and under different conditions, and tend to be more unified within a large population (Woodward et al, 2003) so physiological biometrics tend to be used for identification-based systems as they are considered more reliable approaches. While some behavioural biometrics have very good accuracy for verification, the identification accuracy is considerably lower as the number of users in the database becomes larger (Yamploskiy and Govindaraju, 2010). This is because users act differently depending on mood, illness, stress, previous events, environment, to name a few. For this reason, behavioural biometrics tend to be used for verification-based systems.

However, one of the potential advantages of behavioural biometrics is that they can be utilised non-intrusively and continuously during the user session. In this way, it becomes possible to extend the authentication process beyond initial point-of-entry and verify the identity in a transparent manner, without the explicit involvement of the user. Moreover, the collection of behavioural data often does not require any special hardware and is so very cost effective. Clarke and Furnell (2007) proposed a framework that uses a combination of secret based knowledge and biometric techniques to meet a number of objectives for an effective authentication approach for mobile devices:

- to increase the authentication security beyond the PINs;
- to provide user convenient by operating transparent or non-intrusive authentication that authenticate user without knowledge of user;
- to provide continuous authentication that can authenticate user during usage of device rather than simply at switch-on;
- to provide flexible architecture that able to be applied with any differing hardware configurations, processing capabilities, and varying levels of networks connectivity;

In the best case, the proposed mechanism enhances PIN/password-based authentication with keystroke analysis and can provide transparent and continuous authentication of the authorised user to increase the level of authentication beyond the standard point-of-entry PIN/password technique. However, individual biometric techniques such as keystroke analysis are not suited to all users and scenarios and also cannot provide adequate reliability. For example, although keystroke analysis can provide continuous and non-intrusive application in keyboard-intensive contexts; if keystroke analysis is not working it will adversely affect the authentication system that can be providing only point-of-entry PIN/password technique. Indeed, it should be apparent from Table 1 that even if all of the characteristics exhibited the same discriminative abilities, the resulting biometrics would not be of equal value and applicability in practical contexts. The recognition that no single biometric is ideally suited to all scenarios is an important one, and several studies show that multi-modal biometric approach is superior to any single biometric approach (Brunelli and Falavigna, 1995; Kittler et al., 1997; Poh and Korczak, 2001; Ross et al., 2001) Hence, a combination of biometric techniques should provide adequate security and greater reliability for an enhanced authentication system.

In light of the foregoing exploration, a preliminary study was conducted using a novel combination of finger pressure and linguistic profiling was undertaken.

3. Studying the feasibility of Finger Pressure

According to the research of Grabham (2008), the study of a biometric based on pressure and key press duration of a user entering a PIN on an ATM type interface coupled with a component-wise verification scheme was determined. The results of this research showed that using pressure and key press duration can identify users with high accuracy and low error rate. Therefore, the feasibility of finger pressure applied to a notebook touchpad was used to simulate a mobile touch screen. Details of the experiment are explained in the subsections that follow.

3.1. Experimental procedure

In this case study, the biometric information, keystroke dynamics (i.e. typing rhythms) and finger pressure (i.e. the force applied over the finger position) were captured when a user interacts with a notebook touchpad. Keystroke dynamics can be extracted into two features: inter-key which is the duration of interval between two successive keys, and hold-time which is the duration of interval between the pressing and releasing of a single key. A total of 10 participants (n=10; female=6,

male=4) were asked to enter their cell phone numbers on 12 keys simulated on a notebook touch pad that will be recorded every 20 milliseconds, with 10 digits long, 30 times continuously and repeatedly.

This research was studied into two scenarios, namely using individual characteristics and combination of characteristics. The pattern classification process was developed using the *K*-means classification algorithm. The input to neural network depended upon the scenario with which it was evaluating.

3.2. Results

The results of the study on individual characteristics and combination biometric characteristics are showed in Table 2. Hold-time has the lowest rate of false acceptance of unauthorised users (FAR) (22.59%) and false rejection rates of authorised users (FRR) (0%). These findings illustrate that using hold-time as the key to identify users is the most effective individual measurement. In the case of combination biometric characteristics, the results show that 'The Best User' is obtained from the concatenation of hold-time and finger pressure input scenario, FAR 11.48% and FRR 0%. However, the combination of three biometric characteristics together with the combination of hold-time and inter-key time values could also be applied to identify users with the results of FAR 21.48% and FRR 0% respectively.

	The Best User		The Worst User		Average	
	%FAR	%FRR	%FAR	%FRR	%FAR	%FRR
Hold-time (H)	22.6	0	55.9	6.7	52.2	1.0
Interkey-time (I)	28.1	0	33.7	50.0	49	21.7
Finger Pressure (F)	41.8	0	47.4	3.3	44.8	3.0
HI	21.5	0	68.5	10.0	54.0	3.3
HF	11.5	0	77.8	3.3	64.2	0.3
HIP	21.5	0	67.8	3.3	53.8	2.7

Table 2: Results of individual and combination biometric characteristics

4. Studying the feasibility of Linguistic Profiling

Linguistic profiling has certainly shown its worth for authorship verification and recognition (Halteren, 2004). Hence, this experiment investigated the feasibility of using linguistic profiling to profile and discriminate between users. Previous studies (Yoon et al., 2010) have revealed that the majority of mobile subscribers use abbreviated and phonetically adapted words when interacting with text based communication facilities on mobile devices (i.e. SMS, Mobile Instant Messengers, and Email), and so this presents a potential opportunity for profiling based on patterns of text based communication.

4.1. Experimental procedure

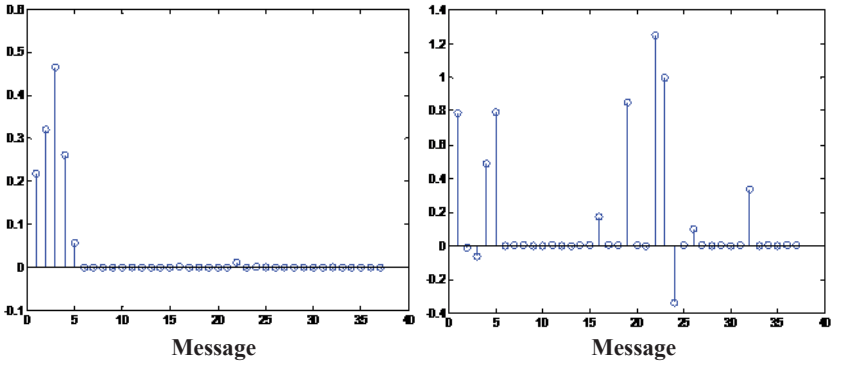
In order to analyse the effectiveness of using text based authentication, a real time SMS simulation program was compiled in Visual Basic and deployed onto two Smartphone's. The program enabled text based communication between pairs of participants using a Bluetooth communication channel. A total of 30 participants were required to send at least 15 messages each to the other participant on a subject of their choice using a non-predictive text input method. The messages were saved in a text file prior to transmission.

In linguistic profiling, many types of linguistic features can be profiled such as vocabulary, information content or item distribution through a text. According to (Yoon et al., 2010), abbreviation and emotional words (e.g. wow, hehehe, oh) are used more frequently in SMS messages so that abbreviation and emotional profiles will be used to create user profiles including every possible type of feature. Discriminating characteristics include message length, word length, average of number of words per message, and number of symbols, digits and capital letters.

The pattern used for the classification process is dependent upon the context to which it was evaluating. For example to authentication user1, the total message will count number of words found in linguistic profiling of user1.

4.2. Results

The preliminary findings from the experiment were analysed by statistic and classification techniques to study the feasibility of classification using linguistic profiling. To study the complexities of successfully discriminating between the users, mean and standard deviation were analysed to calculate the spread value. The spread of values, the range of a users input vectors that reside between lower (mean minus a standard deviation) and upper (mean plus a standard deviation), showed that users input vectors are more likely to be similar, or within similar boundaries as other users, which means it will make classification more difficult. However, the findings were positive in that linguistic profiling can indeed be used to discriminate for some users. For example, using a combination of number of words, number of sentences, number of full stops and question marks features analysed using Radial Basis Function (RBF) with 76 messages that were used for training and 37 messages were used for testing. Figure 1 clearly illustrates that the aforementioned characteristics could be used successfully to discriminate User1 from a total of 5 users (a), however this was not the case for User 4 (b) which clearly shows similarities between participants.



(a) User1 **(b) User4**

Figure 1: Result of classification using Radial Basis Function

Table 3 shows that using *K*-means classification technique elicited the best result with FAR = 17% and FRR = 13%. Table 4 shows this was achieved by using abbreviation and emotional based profiling.

User #	Good User 1	Good User 12	Bad User 7	Bad User 20	Average
Valid Attempts	15	15	14	15	16.7
Invalid Attempts	486	486	487	486	484.3
%FAR	17.1	10.5	93.4	92.2	23.4
%FRR	13.3	20.0	28.6	46.7	43.9

Table 3 Result of classification using *K*-means classification technique

User #	Abbreviation Profile	Emotional Profile
1	‘s’, ‘u’, ‘ur’, ‘fr’, ‘bout’, ‘hvnt’, ‘b4’, ‘hv’, ‘fav’, ‘i've’, ‘hvn't’, ‘yr’	‘haha’, ‘owh’, ‘yeah’, ‘erm’, ‘…’
12	‘u’, ‘cwk’, ‘q’, ‘c’, ‘txt’, ‘r’, ‘abt’, ‘x’, ‘ok’	‘lol’, ‘xx’, ‘hey’, ‘aww’, ‘um’, ‘yeah’, ‘☺’, ‘...’
7	‘u’	‘wow’, ‘hehe’, ‘ok’, ‘erm’, ‘…’
20	‘i've’, ‘i'd’, ‘id’, ‘x’	‘oh’, ‘haha’, ‘yeah’, ‘awesome’, ‘…’

Table 4 Example of Linguistic Profiles

5. Discussion

The preliminary studies undertaken using behavioural biometrics, keystroke analysis and finger pressure, show that using a combination of hold-time and finger pressure gave the lowest error rates and improving upon the findings from Grabham (2008).

The touch screen interface used in the finger pressure experiment used an interval range of 0 – 255. This range was considered to be very limiting and would prove ineffective in discriminating between a large number of users or participants. To this end, it is envisaged that the greater number of participants using this technique would result in a greater number of matches and false positives. In addition, this technique is likely to produce many false negatives, as the pressure applied by the end-user would be influenced by individual mood and energy levels.

The results in the feasibility study of linguistic profiling showed low error rates using abbreviation or emotional based profiling. Participants (e.g. 'Good User') who used a great number of abbreviated or emotional words produced significantly lower error rates. However, participants who used the same common abbreviation and emotional words (e.g. 'wow', 'u') tended to produce higher error rates. Therefore, further analysis is required in order to determine the best discriminating features. The potential advantage of behavioural biometrics is that they can be utilised non-intrusively and continuously during the user session, thereby creating an authentication process beyond initial point-of-entry.

6. Conclusions and future work

To extend the point of entry authentication mechanism currently used on mobile devices, a combination of biometric techniques is considered to maximise the potential for reliability and security. Having explored finger pressure and linguistic profiling for use in mobile devices in section 3 and 4, it is proposed that behavioural biometrics can provide transparent (non-intrusive) and continuous authentication. Moreover, behavioural biometrics, keystroke dynamics and linguistic profiling do not require any addition hardware, therefore reducing overhead costs.

Subsequent to the preliminary experiment using finger pressure, limitations concerning participant mood, tiredness and the need for specific hardware were considered. In addition, current state of the art pen based TouchPads are unable to measure the pressure of pen contact as they report all pen strokes with a constant value (Synaptics, 2001). In order to overcome these limitations linguistic profiling was explored.

The next step in this research is to further analyse the linguistic profile dataset in order to identify the most effective text based discriminators. Once complete the findings will be used to develop a multi-tier behavioural biometric authentication system to provide reliability and security to a wide range of mobile devices within any given mobile environment. The proposed framework 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.

7. References

Brunelli,R. and Falavigna, D. (1995), "Personal Identification using Multiple Cues", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Volume 17, Number 10, October 1995, pp. 955-966, doi:10.1109/34.464560

Clarke, N. and Furnell, S.M. (2007), "Advanced user authentication for mobile devices", *Computer and Security*, Volume 26, Number 2, March 2007, pp.109-119, doi:10.1016/j.cose.2006.08.008

Grabham, N. and White, N. (2008), "Use of a Novel Keypad Biometric for Enhanced User Identity Verification", 2008 *IEEE International Instrumentation and Measurement Technology Conference*, Victoria, British Columbia, Canada, 2008, pp.12-15

Halteren, H. (2004), "Linguistic Profiling for Author Recognition and Verification", *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics (ACL 04),* Association for Computational Linguistics, Morristown, NJ, USA, doi: 10.3115/1218955.1218981 http://dx.doi.org/10.3115/1218955.1218981

IBG (2006), "Which is the best biometric technology", International Biometric Group, http://www.biometricgroup.com/reports/public/reports/best_biometric.html (Accessed 4 January 2009)

ITU (2010), "The World in 2010: ICT Facts and Figures", http://www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf. (Accessed 4 January 2009)

Kittler, J.,Matas, J., Jonsson, K. and Ramos Sanchez, M. U. (1997), "Combining Evidence in Personal Identity Verification Systems", *Pattern Recognition Letters*, Volume 18, Number 9, pp.845-852, doi: 10.1016/S0167-8655(97)00062-7

Mformation (2009), "Mformation-sponsored survey reveals mobile users worried about loss and mobile fraud", http://www.mformation.com/mformation-news/press-releases/mformation-sponsored-survey-reveals-mobile-users-worried-about-loss-and-mobile-fraud. (Accessed 10 November 2010)

Poh, N. and Korczak, J. (2001), "Hybrid Biometric Authentication System Using Face and Voice Features", *Lecture Notes in Computer Science, Volume 2091/2001*, pp. 348-353, doi: 10.1007/3-540-45344-x_51

Ross, A., Jain A. and Qian, J-Z. (2001), "Information Fusion in Biometrics". *Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition*, Springer-Verlag, London, UK, pp. 354-359 ISBN: 3-540-42216-1

Synaptics (2001), "Synaptics TouchPad Interfacing Guide", http://www.synaptics.com/developers/manuals (Accessed 29 May 2008)

Woodward, J.D, Orlans, N., Higgins, P. (2003), "Identity Assurance in the Information Age" *McGraw-Hill/Osborne*, Berkeley, California, ISBN 0-07-222227-1

Yamploskiy,R.V. and Govindaraju V. (2010),"Chapter 1 Taxonomy of Behavioural Biometrics". *Behavioral Biometrics for Human Identification: Intelligent Applications,* 2010, doi: 10.4018/978-1-60566-725-6

Yoon, J. W., Kim, H. and Huh, J.H. (2010), "Hybrid spam filtering for mobile communication", *Computers and Security*, Volume 29, Number 4, pp.446-459, doi:10.1016/j.cose.2009.11.003