2017

Behavioural Monitoring via Network Communications

Alotibi, Gaseb

INFOSECURITY WITH PLYMOUTH UNIVERSITY

Behavioural Monitoring via Network Communications

by

Gaseb N. Alotibi

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

DOCTOR OF PHILOSOPHY

School of Computing, Electronics and Mathematics
Faculty of Science and Engineering

March 2017

COPYRIGHT STATMENT

Abstract

Behavioural Monitoring via Network Communication

Gaseb N. Alotibi

It is commonly acknowledged that using Internet applications is an integral part of an individual's everyday life, with more than three billion users now using Internet services across the world; and this number is growing every year. Unfortunately, with this rise in Internet use comes an increasing rise in cyber-related crime. Whilst significant effort has been expended on protecting systems from outside attack, only more recently have researchers sought to develop countermeasures against insider attack. However, for an organisation, the detection of an attack is merely the start of a process that requires them to investigate and attribute the attack to an individual (or group of individuals).

The investigation of an attack typically revolves around the analysis of network traffic, in order to better understand the nature of the traffic flows and importantly resolves this to an IP address of the insider. However, with mobile computing and Dynamic Host Control Protocol (DHCP), which results in Internet Protocol (IP) addresses changing frequently, it is particularly challenging to resolve the traffic back to a specific individual.

The thesis explores the feasibility of profiling network traffic in a biometric-manner in order to be able to identify users independently of the IP address. In order to maintain privacy and the issue of encryption (which exists on an increasing volume of network traffic), the proposed approach utilises data derived only from the metadata of packets, not the payload. The research proposed a novel feature extraction approach focussed upon extracting user-oriented application-level features from the wider network traffic. An investigation across nine of the most common web applications (Facebook, Twitter, YouTube, Dropbox, Google, Outlook, Skype, BBC and Wikipedia) was undertaken to determine whether such high-level features could be derived from the low-level network signals. The results showed that whilst some user interactions were not possible to extract due to the complexities of the resulting web application, a majority of them were.

Having developed a feature extraction process that focussed more upon the user, rather than machine-to-machine traffic, the research sought to use this information to determine whether

a behavioural profile could be developed to enable identification of the users. Network traffic of 27 users over 2 months was collected and processed using the aforementioned feature extraction process. Over 140 million packets were collected and processed into 45 user-level interactions across the nine applications. The results from behavioural profiling showed that the system is capable of identifying users, with an average True Positive Identification Rate (TPIR) in the top three applications of 87.4%, 75% and 61.9% respectively.

Whilst the initial study provided some encouraging results, the research continued to develop further refinements which could improve the performance. Two techniques were applied, fusion and timeline analysis techniques. The former approach sought to fuse the output of the classification stage to better incorporate and manage the variability of the classification and resulting decision phases of the biometric system. The latter approach sought to capitalise on the fact that whilst the IP address is not reliable over a period of time due to reallocation, over shorter timeframes (e.g. a few minutes) it is likely to reliable and map to the same user. The results for fusion across the top three applications were 93.3%, 82.5% and 68.9%. The overall performance adding in the timeline analysis (with a 240 second time window) on average across all applications was 72.1%.

Whilst in terms of biometric identification in the normal sense, 72.1% is not outstanding, its use within this problem of attributing misuse to an individual provides the investigator with an enormous advantage over existing approaches. At best, it will provide him with a user's specific traffic and at worst allow them to significantly reduce the volume of traffic to be analysed.

Contents

List of Figures

List of Tables

Acknowledgments

I would like to begin by thanking ALLAH, without whose guidance this research would not have been possible. Then I would like to express my deepest and sincere appreciation to my advisor, Prof. Nathan Clarke, for his continuous assistance, valuable guidance, understanding and patience throughout my academic growth. His breadth of knowledge and research experience continues to inspire me. I am indebted to him for his assistance and support.

I would also like to thank the other members of my PhD committee, Prof. Steven Furnell and Dr Fudong Li, for agreeing to oversee my research, dissertation studies and for taking the time to review this document.

My sincere appreciation goes to the soul of my father (ALLAH have mercy upon him), my mother and brothers; Prof. Turkey, Mohammed and Gaith for supporting and encouraging me throughout my life and all my endeavours.

Special thanks also go to General Saeed Al-Qahtany, Major General Saad Al-Kliwee and Dr Mohsen Al-Sharif, who always encouraged me to undertake my PhD and supported me from an early stage.

I would like to express my thanks to the government of the Custodian of the Two Holy Mosques and the Minister of the Interior for giving me this chance to complete my studies and be awarded this qualification for the sake of serving my country.

Finally, I would like to express my deepest gratitude to my wife, my sons Aali and Nayef and my daughter Amerah for their love, endurance, patience and support throughout my academic studies. I know that without their help and encouragement this endeavour would have been very difficult. Thanks to them, frustration and stress never stood a chance.

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 without prior agreement of the Graduate Committee.

This study was financed by the government of Saudi Arabia – the Saudi Cultural Bureau in London.

Relevant seminars and conferences were regularly attended at which work was often presented and several papers were published and prepared for publication. Other research skills development courses were also attended.

Word count of the main body of the thesis: 50,165

1 Introduction and Overview

1.1 Introduction

In recent years, using internet services and applications has become a part of our daily life, whereby a major development process is undertaken to fulfil the majority of essential tasks via such applications, such as utilising phones, managing the booking of air flights, purchasing goods from supermarkets, acquiring utility services (e.g., water and electricity) and online banking transactions. All these activities and others can now be done easily and quickly through internet services and applications. Therefore, the number of internet services being used across the world has been impacted accordingly and increased from 361 million in 2000 to 3.5 billion in 2016, this proportion representing about half the world's population (Internet Lives State, 2016). In addition, the use of e-commerce, whereby products and services are traded via the internet, has escalated dramatically in recent years. According to Ecommerce news, in 2016, this European e-commerce umbrella organisation announced that e-commerce revenue had increased among European countries, with, for instance, real revenue from e-commerce trade for the UK in 2015 was €157.1 billion, expected to reach €173 billion in 2016 (E-commerce news, 2016).

As this dependence on computers and Internet applications has grown, however, have the threats against them. Consequently, information security has become increasingly important as technology consolidates its position into our everyday lives and keeping information safe from external or internal threats has become a significant issue. Indeed, approximately 82% of large organisations have considered information security to be a high priority concern that must be addressed (PricewaterhouseCoopers, 2015). McAfee reported in 2014 that the estimated annual cost of cyber-crime to the global economy was more than $400 billion, across companies and individuals (MacAfee, 2015).

Whilst external threats remain an issue, much focus by both researchers and industry has seen the development of security controls to protect outsiders from getting unauthorised access into systems. However, these tools often do not help organisations against insider threats. Insider threats have grown to be an increasingly significant problem; this is due to the privilege and knowledge in which an insider has compared with an external attacker (Silowash et al. ,2012). While, research has grown in the area and many techniques have been proposed to detect insider threats such as Data Loss Prevention (DLP) tools, they have often focussed upon the detection rather than the response to the attack. Those studies that have focussed upon the response, tend to take a network-centric perspective where it is assumed the IP address is directly attributable to a user (Rouse, 2012). However, with the widespread use of the Dynamic Host Control Protocol (DHCP) and (more importantly) mobile computing, the Internet Protocol (IP) address is subject to change on an increasingly frequent basis. As such, whilst the detection tool will provide an IP address for the specific detected attack, for an analyst to take the next step of the investigation and understand the traffic profile of the insider – in order to provide complete coverage of the attack, identify other perpetrators, impact upon the organisation, is an incredibly challenging task. Therefore, a need exists for an approach that is able to effectively identify the network traffic associated with an individual (rather than the IP address).

1.2 Aim and Objectives

This thesis is specifically on the subject of behavioural biometrics. The research aims to investigate and determine whether a behavioural-based user profiling system would be capable of identifying users from network traffic. The objectives of this research are as follows:

- To review and identify the current state of the art regarding insider misuse detection and response.

- To create a novel feature extraction process focussed upon extracting user-oriented actions from the wider network traffic data.
- To examine the feasibility of developing a behavioural profile of users based on these user interactions.
- To develop and evaluate practical-based algorithms to determine real-world performance.

1.3 Thesis Summary

Chapter 2 provides a background of insider threats. It presents a definition of insider threats and reveals the scale of the problem based on statistical research and cases studies. This provides an impression of how insider threats can influence business and become a serious problem on an international scale. It also explains the mechanism of how existing tools are being utilised to tackle this issue by discussing pros and cons of each of them. The chapter concludes by highlighting the need for a security mechanism that can protect organisations and the need to be able to attribute this to the right person who committed the misuse.

Chapter 3 presents and discusses the feasibility of utilising biometric techniques as a solution to deal with the identified issue. The chapter begins with discussing the requirements of biometric systems, their performance and characteristics. A comprehensive summary of each category is then provided. Also, a detailed analysis and critique of behavioural profiling are presented to understand the current state of the art and what can be achieved to date. Then, an overview of the two main approaches that can be utilised to detect the insider misuse through network traffic analysis. It presents the concept of how these approaches work and highlight the serious limitations that made them incapable to identify users.

Chapter 4 proposes and explains in detail the concept of user interactions which provides the basis for identifying the discriminative information amongst users. An investigation is undertaken to analyse the extent to which it is possible to derive application-level user interactions from low-level network packet data. The chapter concludes by providing a list of user actions signatures that were determined at an application level from different Internet applications.

The main focus of Chapter 5 is to design and evaluate a behavioural-based profiling system that can utilise the aforementioned user interactions to identify users. It also explains the process flow of the classification and evaluation tasks, in addition to each component of the system and the analysis of the outcomes. The experiment utilises a data collect from 27 participants over 2 months to evaluate the performance and uniqueness of the feature extraction and classification processes.

Chapter 6 builds on the success of Chapter 5 by proposing two further techniques that can be applied to the behavioural profiling system, with a view to improving the performance. Biometric fusion is utilised at the output of the multi-classifier (more than one classifier) to provide a more refined classification decision. The second technique develops an algorithm that utilises the timeline and the IP address (over very short time periods). The chapter provides an evaluation of both.

Finally, Chapter 7 presents the main conclusions of the research, highlighting the main achievements and limitations. The future research opportunities and directions of this project are also discussed.

2 Insider Misuse and Incident Response

2.1 Introduction

The insider misuse issue becomes a major concern for many organisations; hence this chapter presents a number of research studies that were carried out regarding insider misuse to provide greater clarity and granularity to the problem. The chapter begins with various definitions of insider misuse and then recognition of the scale of the insider threat situation, supported by a number of surveys, followed by well-known examples that have emerged in recent years. The final section will address the pros and cons of the existing tools that are being utilised to respond to any insider misuse incident.

2.2 Definition and Type of Insider Misuse

There are two kinds of attacks which may occur to any organisation; internal and external. Both of them can damage the system, but the intentional insider threat is arguably more dangerous than other attacks because the insider knows more details about the sensitive information and is familiar with the kind of system valued within the company or organisation more than others (Phyo and Furnell, 2004; Hunker and Probst, 2011). Thus, Insider misuse defined as "*a current or former employee, contractor, or other business partner who has or had authorized access to an organization's network, system, or data and intentionally misused that access to negatively affect the confidentiality, integrity, or availability of the organization's information or information systems*" (NCCIC, 2014)

Researchers have categorised insider threats into different types, some of them are based on their level of access whilst others according to their role. One of the first researchers who suggested to categorise insider abuse was Anderson (1980). He describes three types of internal

threat, namely; masqueraders, clandestine and misfeatures. Masqueraders are internal users who have the authority to access the system but use their privilege to exploit the weaknesses of the authentication system, thus gaining the identity of another legitimate user account. A clandestine user is connected to an authorised user, where such a user can have capabilities to evade audit, control and access resource mechanisms in a specific computer system. Finally, misfeatures are insiders who abuse the system by misusing their privileges. Magklaras and Furnell (2002) suggested that insider threats can be classified into three classes as follows, system masters (includes all the authorised users who have full administrator privileges), advanced users (includes all the authorised users who have wide knowledge of the organisation's internal system and processes) and application users (includes all remaining authorised users within the organisation). This category clearly clarifies that user is the key element of the organisation in different levels and shows the level of privilege that may each level has, thus the consequences of insider threat mainly relies on in which level the user located in and this is due to the privilege that system masters have. Furthermore, Cole and Ring (2005) have divided insiders into four categories based on the level of access as follows: pure insider (users with full privileges and access), insider associate (users with limited authorised access), insider affiliate (a friend or partner who compromises an employee`s identification to access the system) and outside affiliate (an outsider person who tries to exploit open access to gain access to an organisation`s resources). One of the important point that organisations try to mitigate and control it is limited the access to their resources from outside. However, there are some issues may occur from users who inside the originations, thus dividing the insiders based on their access might control the kind of information and boundary that need to be available for them.

Even though the aforementioned researchers have discussed the insider misuse from a theoretical point of view, a number of surveys have categorised insider misuse based upon incidents that were committed by employee misuse in which various scenarios were designed based on the collected data as outlined below:

1. Espionage: is a method which includes human sources (agents) or technical means to gain information which is not normally available to the public (Security Service, 2014).
2. IT Sabotage: is an insider incident in which the insider utilises information technology to target an organisation or individual (Capplli et al. 2012). This kind of attack could go beyond targeting individuals to target governmental agencies (Chee, 2007).
3. Fraud: is a set of activities against an organisation that can be committed by an employee, such as falsely claiming for expenses, colludes with an outside accomplice(s), and stealing a database containing personal details of customers, which is used by a criminal gang to obtain money, credits, goods or services fraudulently from other organisations (The Security Company, 2013).
4. Intellectual Property Theft: an insider who uses the IT systems to steal intangible assets that were generated and owned by an organisation which is important to attaining its mission (Capplli et al., 2012).
5. Lack of compliance with policy: all employees must comply with the information security policy and guidelines. This can be carried by monitoring the daily activities of the users where any lack of compliance needs to be addressed by management (Hostland et al., 2010).
6. Unauthorised Access is to gain access to other computers or using someone else's account to gain access without any permission.
7. Misuse of Confidential Information: is an unlawful use of sensitive information.

8. Loss or Leakage of Confidential Information: describes the action of intentionally or accidentally leaking confidential information to the public.

It is evident that user is the key player for the majority of the incident. However, the intention or the purpose to do that is altered. In addition, the way of employee misuse is also subject to change based on various elements such as privilege, skills and so forth.

2.3 The Level of Impact of Insider Threats

According to the Information Security Breaches surveys that were conducted by PricewaterhouseCoopers (PwC) in 2013, 2014 and 2015, insider attacks have become serious threats to organisations. Figure 2.1 reveals how insider threat incidents have gradually increased during the three years period. Indeed, unauthorised access to systems in organisations represents the highest threat in the last three years, whilst, information leakage has achieved a significant increase in a number of incidents with a 15% increase to 66% in 2015. Also, there is a steadily rising in the misuse of information incidents between 2013 and 2015 as can be seen in figure 2.1 below.

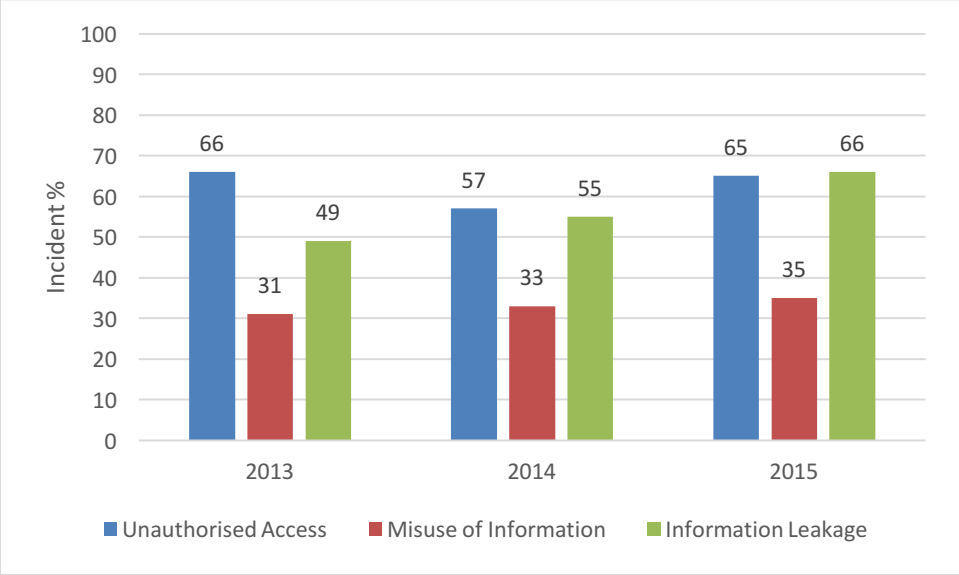

Figure 2.1: Level of Insider Threats incident in organisations in 3 years (PwC,2015)

These threats have originally originated from inside the originations by one of the employees as mentioned above, thus the consequences may be quite costly due to the knowledge and privilege the insider has. PwC (2015) stated that individual security breaches can cost organisations more than £1 million in damages. It can be quite difficult to specify the exact impact of insider threats that would occur within an organisation because many may prefer to solve this kind of problems internally rather than prosecuting the insider for a malicious attack. According to the PwC (2011), about 88% of the respondents in a survey they conducted stated that insider threats had been solved internally without an intervention from any third party regardless of whether it was private or public. This was done to avoid further risks or consequences and to maintain the organisation's reputation.

As shown in figure 2.2, both small businesses and large organisations have suffered from the insider misuse threat, with the latter being impacted more seriously (i.e., large organisations have achieved high numbers of incidents when 81% of them have experienced various security breaches in 2015). This is due to the number of employees that are working within this type of

organisations and mobile computing which increased the possibility to be exposed to the insider misuse (Saed.F, 2016). In addition,

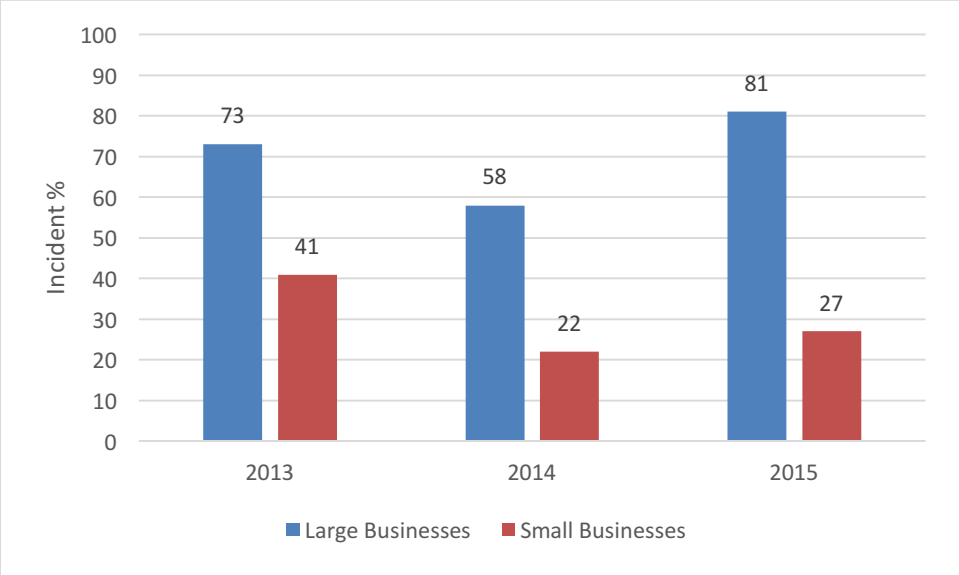

Figure 2.2: Insider Misuse Incident staff between 2013 and 2015 (PwC,2015)

Although organisations have tried to hide insider misuse incidents from the public and solve it internally, there are various cases that are well known across the world either in the private or public sector. The cases that are going to be discussed refer to common incidents occurred in the private and public sectors. Indeed, they describe to what extent insider misuse could be harmful even for the advanced organisation that supposed to have a strong security infrastructure such as Central Intelligence Agency (CIA) in the USA that can be explained and discussed as outlined below.

Edward Joseph Snowden was an American systems analyst and a former employee of the CIA and National Security Agency (NSA). While he was working for these two agencies, he uncovered some illegal actions that, in his opinion, had breached the Internet usage privacy. The agencies were using Internet surveillance programs, such as PRISM, XKeyscore and Tempora, as well as intercepting the US and European telephone metadata (CNN, 2013). He

felt obliged to bring this story into light by speaking to the press in order to shame the US government and put an end to this kind of surveillance programs. As a result, he had leaked secret NSA documents to different media outlets.

Since Snowden was a specialist in information security, and he was responsible for the computer network in one of the US embassies in Europe; this gave him an opportunity to access a lot of confidential information up until leaving in 2009. This means he followed the policies and instructions of these agencies and did not arouse any suspicion at that time. However, he had copied all the documents that he was interested in. As an IT expert, he could collect information from the system in one of the most secret agencies in the world without arousing any suspicion that would indicate that there was a serious problem with the internal security policy of the organisations.

Another example happened in the middle of August 2012, there was an attack on a Saudi Arabian Oil Company (known as Saudi Aramco), one of the 10 biggest oil companies in the world (Borken and Ringas, 2013). The attack utilised a Shamoon virus, which affected around 30,000 windows-based personal computers operating on the company`s network. The main aim of the attack was to stop all Aramco's activities, especially the flow of oil and gas to local or international markets (Infosecurity, 2012). Shamoon spread through the company's network and wiped the computers' hard drives clean. Saudi Aramco stated, "damage was limited to office computers and did not affect systems software that might hurt technical operations" (Finkle, 2012). However, it took Aramco around two weeks to completely restore the network and recover from the disruption to its daily business processes caused by data loss and disabled workstations resulting from the incident (Borken and Ringas, 2013). Although the results of the final investigation into this attack have not yet been released, a source familiar with the

investigation and forensic examination said, "It was someone who had inside knowledge and inside privileges within the company" (Finkle, 2012). This idea is supported by different parameters, such as the attack was launched during a national holiday and the external detection tools did not raise any threat could have occurred from outside the organisation.

A further example of a dangerous insider threat case comes from the military in the US. Bradly Manning leaked a diplomatic cable and other classified documents to WikiLeaks website, which included, “an unprecedented look at back-room bargaining by embassies around the world, brutally candid views of foreign leaders and frank assessments of nuclear and terrorist threats” (Shane and Lehren, 2010). Bradly is an authorised user who had accessed top secret information and saved it for more than 3 years before releasing it on the Internet. In doing so, he utilised the privileges given to him to cause damage and abused the trust of the organisation. Nobody can imagine the scale of damage to the US government in terms of their reputation and relationship with other countries and the time taken for the US government to regain trust (Shane and Lehren, 2010).

Furthermore, Abdulkader Smires was a chief software engineer at Internet Trading Technologies. Smires was not satisfied with what he had earned from his employer, so he complained about his low salary and asked for a pay rise with additional benefits, but his request was denied. Thus, he decided to take revenge against the organisation by using his previous account at another organisation (Queen College) to launch a Denial of Service (DoS) attack. Internet Trading Technologies lost 3 days of revenue due to the consequence of this attack (Radcliff, 2000). Smiles's case shows us how an expert individual may retain sensitive information in his mind until he decides to utilise it. Smire knew the infrastructure of Internet Trading Technologies very well, and he had the ability to exploit the vulnerabilities in the

system to launch a DoS attack. Another point is that he used his previous legitimate access to another organisation (Queen College) when this account should have been removed from the system since the employee had left the organisation.

In 2001, an infamous insider misuse case occurred in the US. According to (CSC/FBI, 2001), former FBI agent Robert Hanssen abused the trust of the FBI by using his authorisation to access the FBI system than handing over some important information about ongoing investigations that he retrieved from the system to Russian agencies. His motivation for doing this was financial gain. Using his IT expertise, Hanssen utilised an unusual technique to hide the information so that only advanced digital forensics could find it. His technique relied upon using a special process to hide the information in 40 track diskettes, where it appeared as a blank space. By using this method, he handed sensitive information that might have been used to damage U.S. National Security over to the Russian Intelligence Services for 22 years between 1979 and 2001.

Based on the previous cases it is quite evident that when someone intends to misuse information, different attack vectors will be used to achieve the aim. The above cases were intentional because all the perpetrators planned the misuse; however, their motivations for doing so vary, for example, revenge, money or personal opinions about certain issues. Indeed, what they all have in common though is that they utilised their legitimate user status to commit this kind of crime and abused the trust of the organisation they attacked to commit a crime.

Subsequently, identifying the individual insider misuse is the challenging and major concern of organisations in different sectors and countries. According to a survey that was conducted by CERT in the US, among 557 respondents, 75% of insider intrusions are handled without

legal action while just 10% of insider intrusions handled with legal action (CERT, 2014). This is due inability to determine the individual responsible user as well as lack of evidence. In the same survey, 37% of respondents think that insider threat cannot identify the individual responsible for committing Crime. Consequently, a further research needs to be carried out regarding identifying individual insider threat in order to gain people's trust.

2.4 Incident Response

In a normal response to the increased number of insider threats, both public and private sectors have been seeking a technique that can reduce the scale of insider threats, thus, minimise the damage they may cause. The following sections will discuss two types of technologies that are commonly utilised for incident response process of insider misuse threat; Data Loss Prevention (DLP) and Security Information Event Management (SIEM), which is considered to be a core solution of these types of threats and being used in incident response.

2.4.1 Data Loss Prevention

DLP is a tool that incorporates predefined policies and requirements for monitoring and utilising the sensitive information held in different locations in the organisation's system (Sans, 2008). DLP first emerged on the market in 2006 and became popular in early 2007 (Sans, 2008). It is described as a plan that encourages end users to not send sensitive, confidential or critical information outside of the responsible official network. It is also utilised to describe software products that contribute to a network administrator being able to control what data end users can transfer (Rouse, 2012). Other definitions from the security companies of DLP are more comprehensive. For instance, Symantec has recently defined DLP as a comprehensive data security solution that discovers, monitors, protects and manages information wherever it is stored or used (Symantec, 2015). Data in the system can be found in different situations,

across the network, storage and endpoint systems, and DLP can be used to prevent losing data from those situations. Figure 2.3 below describes the concept of how DLP tools work to keep the data safe. The tool classifies data location into three parts: data at rest, data in motion and data at the endpoint, and in each part, there are a number of actions such as Databased which is representing data at rest, email represents data at motion and USB driver that represents data at Endpoint. Moreover, the predefine policy is the core task that controls the status of the data in each part.

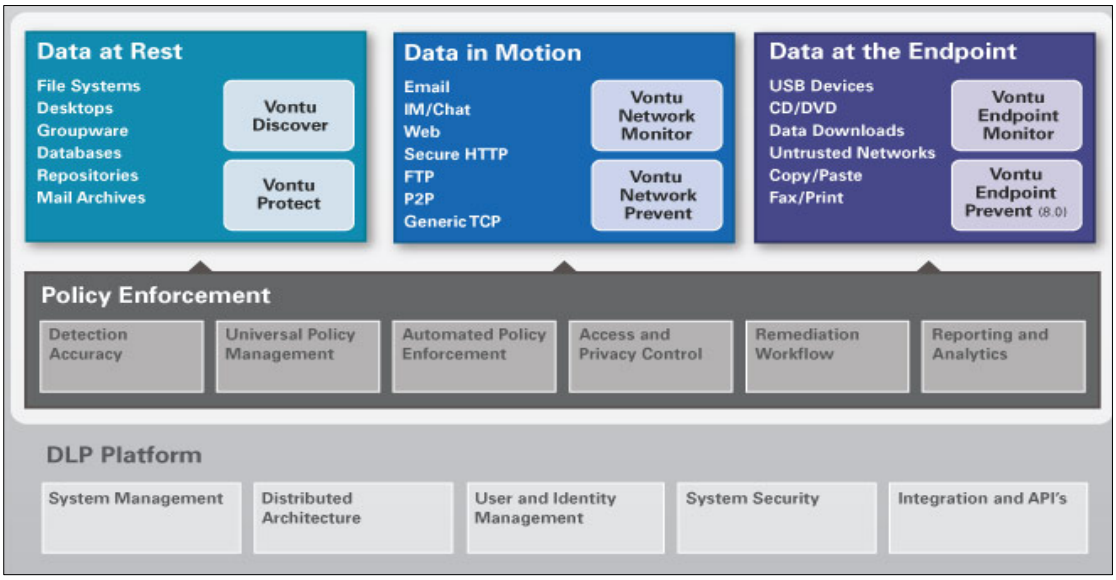

Figure 2.3: DLP Design (Imagined Security, 2014)

Preventing information leakage is one of the most important aims of companies and governmental organisations. Any leakage might harm a company's competitiveness and reputation and lead to lawsuits or regulatory consequences of poor security (Liu and Kuhn, 2010). Data protection has become an important issue, and therefore, sensitive data needs to be protected against leakage and improper handling. The majority of security systems try to protect data from external attacks and prevent the data leaks from outside the organisation. However, the insider threat is considered as a critical issue for organisations in terms of the disclosure of sensitive data (Wuchner and Preschner, 2012). In addition, most insider threats

that culminate in the disclosure of sensitive information can be attributed to the accidental or deliberate activities of company employees (McCormick, 2008). According to the Information Systems Audit and Control Association (ISACA) (2010), three key objectives are included in most DLP solutions:

1. Identify and list confidential information stored throughout the enterprise. This means there is a predefined policy of where the sensitive data is located and who has an authorisation to access these data.
2. Monitor and control any movement of sensitive information inside the organisation network. So that any department that has an access to this type of data will be covered by DLP tool.
3. Monitor and control the movement of sensitive information on the end user system which means there is a restriction on the user usage of the data

Since, the sensitive information needs to be defined based on a cooperation between IT and business departments in order to rectify what the sensitive information is and who should be granted access to it, to avoid any leakage or loss (ISACA, 2010). Liu and Kuhn (2010) indicate that data loss can be divided into two categories; leakage (when the data is not under the organisation's control) and disappearance or damage (when an accurate data copy is no longer available to the organisation or company). Therefore, keeping sensitive information under monitoring during its movement in the network is quite important to protect it from any loss.

The capabilities of DLP can be described in four functionalities as outlined below.

1. Managing; which includes the different issue that needs to be handled properly, such as defining important data, usage policies, establishing reports of data loss incidents and define who has authority to access what and how.
2. Discover; is a second capability which includes identifying where the sensitive data will be stored and managing deleted requests for these data completely or partly.
3. Monitoring which means how can data be monitored at all times.
4. Enforce some policies to keep the data inside the system safe from any change from a unauthorised person by restricting printing, saving, coping and downloading (Liu and Kuhn, 2010).

Although DLP provides some kind of support to help the organisation to keep their information safe, it has some key limitations. These fundamental limitations are: DLP does not support any encryption, which means it cannot decrypt existing encrypted files, if a user who has an ability to encrypt a file, DLP will not be able to decrypt it unless it puts this in the tools itself, in addition, it controls certain data while other parts of the system could be exposed to attacks. Moreover, it cannot identify the individual who's responsible for the misuse; because, it focuses on data and provides monitoring and reporting about it. Any attempt to change from one vender to another or integration with an acquired organisation's solution might require significant work to imitate, such assert the same rules in different devices and the ability to monitor one application in one system may not be done in synchronisation with other application (ISACA, 2010).

DLP has made a significant contribution to reducing the leakage of sensitive data from inside local networks. However, it relies on the predefined polices and instructions that manage and control the mobility of the data inside local networks. Hence, it cannot be considered to be a

perfect solution. This is due to some of these policies might change and new methods may emerge to counter the instructions and these changes might happen intentionally or accidentally from legitimate users, thus sensitive data may be exposed to insider threats. In addition, these instructions are unable to prevent loss of data from users who have authority to see and utilise it, this is a vulnerability if the leakage comes from them. Moreover, identifying individuals who commit this type of crime is out of scope and as explained before, one of the main reasons that organisations do not tend to reveal or report insider threats is the inability to identify who did the crime.

2.4.2 Security Information and Event Management (SIEM)

SIEM system is another category of a security tool that is utilised to protect organisations from insider threats. It emerged on the market in early 2000 due to lack of existing technologies in their abilities to deal with increasing amount and accuracy of security data and inability to centralised visibility. The abbreviation SIEM is derived from a combination of Security Information Management (SIMs) and Security Event Management (SEM). SIM tools are able to aggregate and store log data for different systems, applications, server and router to analyse the traffic in order to detect any threat. While SEM tools apply some sort of analysis to this log data in order to assist in identifying potential threats by concentrating on different activities of perimeter devices within a network, such as intrusion detection system and intrusion prevention system, in order to improve the incident response abilities of those devices. The combination of these two linked capabilities produced the modern SIEM (Schultz, 2009). Security Information and Event Management (SIEM) is a system that collects logs from different devices and applications of a network to provide centralised log handling. If the system detects an attack, it can react through its incident management channels that include alerting personnel and even initiating counter measures. A SIEM may also help an organisation to meet the terms

of regulations pertaining to data retention, and it can be helpful in cases of e-discovery and forensics (Karlzen, 2009).

Gartner (2012) defines SIEM as "*Security information and event management (SIEM) technology supports threat detection and security incident response through the real-time collection and historical analysis of security events from a wide variety of event and contextual data sources. It also supports compliance reporting and incident investigation through analysis of historical data from these sources. The core capabilities of SIEM technology are a broad scope of event collection and the ability to correlate and analyse events across disparate sources*"

There are trends to utilising security management in an organisation's system. According to the SANS survey in 2012, 37% of 600 respondents had used SIEM tools in their organisations. The motivations behind the use of SIEM vary; compliance, insider threats and a large number of devices being the most important reasons for using SIEM (Karlzen, 2009).The different SIEMs tools, sold by, for example, Cisco and Symantec, have some general features in common. In order to understand the capabilities of SIEM tools, it is quite important to know about their general features (Callahan, 2013). These are described below:

- Collection (Node logging): different devices in the network can provide information through the log to the SIEM collection. Then the logs provide records of activities that can potentially be collected and analysed using a SIEM tool (Schultz, 2009). The most common logs are windows servers, security devices, network devices (switch, router) and network system (IDS/IPS/antivirus) (SANS, 2012).
- Event Normalisation makes utilising the event easier in the report step.

- Correlation: predefines the relationship between the event and SIEM provides this service.
- Filters: among a different number of logs, it is quite important to filter out some log events and focus on the important logs.
- Rules: predefined rules are used to evaluate events received from the normalisation (Callahan, 2013).
- Dashboards: SIEMs have utilised dashboards to make it possible for the security officers to monitor an organisation`s network activities and keep an eye on the level of the security on all devices.
- Alert: the alert will be raised automatically when the event threshold occurs.

  Log storage: in the initial set-up of the SIEM, an organization can label important records for log storage in terms of size limits and time limits in order to comply with relevant regulations (Schultz, 2009).

Figure 2.4 : SIEM Architecture (EventLog,2017)

As shown in Figure 2.4 above, the tool relies on information that is collected from different resources such as servers, routers, switches and so forth. This data provides vital information

by analysing these different logs in order to aid a security officer to concentrate upon security incidents that might happen to the system rather than just matching different signals from different devices (RSA security brief, 2012). Figure 2.4 shows how SIEMs tools work and from where the data comes to the tools. It is quite evident that there are many events connected to the SIEMs tool. Hence, analysing all these logs file from SIEM provide vital support for the security officer. Table 2.1 below discusses SIEM capabilities and limitation.

SIEM Capabilities	Limitation
Performing various functions automatically such as collecting, archiving and reporting of log and event data from several devices such as firewalls, IPS, IDS and antivirus.	The design of the SIEM was not built to deal with a huge number and the volume of information security required today, which causes a problem in term of data handling.
Centralising the security data by collecting them in the same repository and giving the Security Operation Center (SOC) authorization to access to this repository.	It is clear that a large amount of data is collected from different logs. However, these data represented a small amount of the available resources and this gives security analyst small amount of potentially relevant activities might belong to the certain threat.
SIEM is offering a centralized repository by the combine of logs and data event from different sources in order to produce a complete security data warehouse.	Although the amount of data that exist within SIEM system is huge, possibility to being utilized for handling incident investigation is poor.
Providing a various control reports that reflect which level the regulations that release from the government and industry were applied.	Although making a compliance report is important, does not mean the security situation of the organisation will be improving or potential risk will reduce.
Using a correlation rules and signature of known attack in order to produce alert mechanism.	This mechanism is not accurate because it limited to the signatures attacks, while the advanced attacks do not have a predict signatures which means the tool is not able to detect advanced attacks

Table 2.1: Capabilities and limitation of SIEM system (Siemplify.com,2016)

Although SIEM may produce a sufficient information to the security officer (analyst) which would aid towards insider threat measurement, it presents three challenges to the user and organisation. The first of these is the regulatory requirements; internal policies should be

considered at the time of applying the tools. The second is an organisation has to be able to deal with a number of devices in addition to inside and outside threats, and determine which threats warrant investigation in order to save staff time and the organisation money. Lastly, gathering logs from the applications and devices is not an easy job and managing a large amount of real time data and store the data is a hard and complex task.

2.5 Conclusion

There is an increasingly wide range of technologies developed to help detect and manage this process but when it comes to response part, there is still a gap (Sperotto et al., 201.; Crawford and Peterson, 2013; Young et al.,2013; Senator et al,2013). This gap is what comes after the detection step which is attributed to the misuse incident by finding out the individual who perpetrated the crime. As discussed above, insider misuse detection tools such as SIEM have done an excellent task in terms of monitoring users and detect a misuse. However, their work to fulfil this part is by depending solely upon IPs address; however, IP is not a reliable measure. With mobile computing and DHCP, the IP address is constantly changing making it challenging for an analyst to appreciate and understand the traffic associated with the attack versus the wider network traffic. Subsequently, there is an urgent need to go further into the investigation phase to find out a new approach that is able to link the traffic to the right person without dependency on IP address.

3 Biometrics and Behavioral Profiling

3.1 Introduction

Biometrics is a technique that can identify verify a person based upon some unique information about them. Well known biometric techniques include fingerprint and face recognition. Behavioural profiling is a biometric technique that is based upon utilising information gathered about what a person is doing (e.g. in a mobile context, who they call, when they call and for how long). It, therefore, has the potential to capitalise upon patterns in user behaviour as a means of identifying who they are. This chapter provides a background into an understanding of what biometrics are, their characteristics and performance in order to better understand whether this technology is suitable for this research. The chapter concludes with a critical evaluation of the prior work completed within behavioural profiling.

3.2 Biometric System

Biometrics as a characteristic can be defined based on ISO/IEC 27001 as "*A measurable biological (anatomic and physiological) and behavioural characteristic that can be used for automated recognition"*. There are two types of biometrics; physiological and behavioural. The former one means the physical characteristics of a person such as fingerprints and hand geometry. The second one is biometric characteristics that are learnt or an outcome of the environment in which the person has lived, such as keystroke and behavioural profiling (Shrivastava, 2013). These different types of biometrics mainly utilized in verification and identification mode (Lewies, 2002; Clarke, 2011). As a result, biometric applications can be useful in many different environments. Alphonse Bertillon, chief of the criminal identification division of the police department in Paris, established and used a kind of measurement of criminals in the middle of the 19th century (Jain, 2004). Later in the 19th century, the distinctiveness of the human fingerprint was discovered and law enforcement departments

started storing criminal's fingerprints. This view is further supported by Jain et al. (2000), who suggested that biometrics can be utilised in a variety of areas, as illustrated in Table 3.1.

Forensic	Civilian	Commercial
Criminal investigation	National ID	ATM
Corpse identification	Driver`s license	Credit card
Parenthood determination	Welfare disbursement	Smart Phone
	Border crossing	Access control

Table 3.1 : Biometrics applications

In verification mode, the system validates a persons' identified by making a comparison between the recently captured sample and his/her previously stored template(s) resulting in a one-to-one comparison. To enable this comparison, the individual will claim an identity (e.g. in an IT context through the use of a username). In identification mode, the system uses a one-to-many method, which means the system recognises an individual's biometric and searches through the database to find out who matches the individual from all users (Shrivastava, 2013). For instance, when it comes to a scenario such as when the police have a picture of a suspicious person and want to know who is this person, identification is required to search for this a person by using face recognition in the citizen database.

A biometric system consists of a number of components, which rely on an individual set of biometric characteristics. Jain (2004) lists four main modules that make up the biometric system. The first is a module which undertakes the first step of the process by capturing the biometric data of an individual, such as a fingerprint sensor, which records the user's fingerprint. Secondly, the feature extraction module extracts or selects the discriminating features from the sample resulting in a feature vector. Thirdly, the matcher module makes a comparison between the recently acquired feature vector and the stored reference templates. Finally, the decision module makes a final determination of the user through the application of a threshold to the matching subsystems output. However, Clarke (2011) has added a new

component named by storage; utilised to store the features for next step; before match component. Figures 3.1 and 3.2 present an overview of the four main modules of the verification and identification biometric system, considering the differentiation between the verification and identification mode, in which, the first mode the user claimed identity and the system deals with one matcher of the sample and returns true or false whilst the other one there is no identity claimed and deals with number (N) of matcher in which it would result with user's identity or user not being identified.

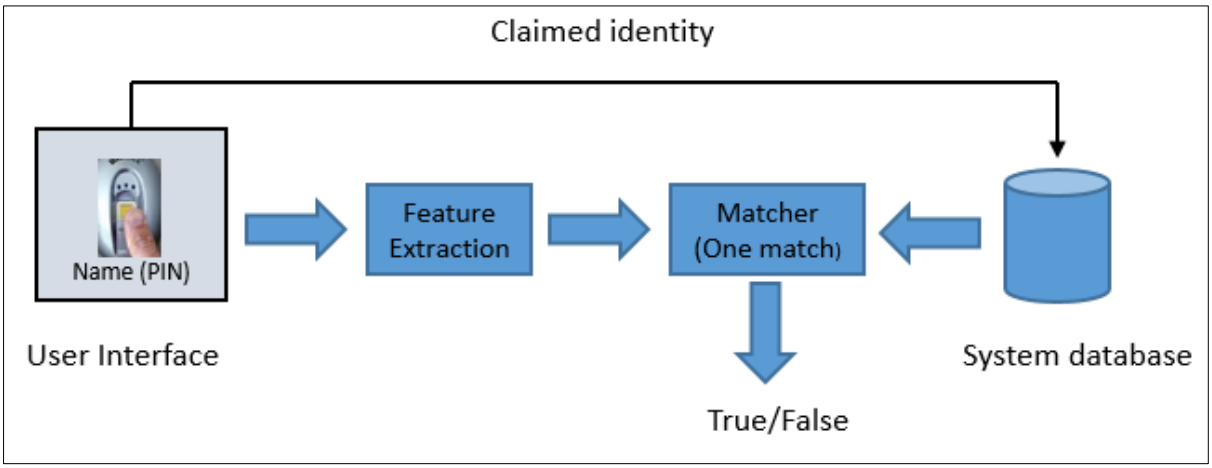

Figure 3.1 : A biometric verification process

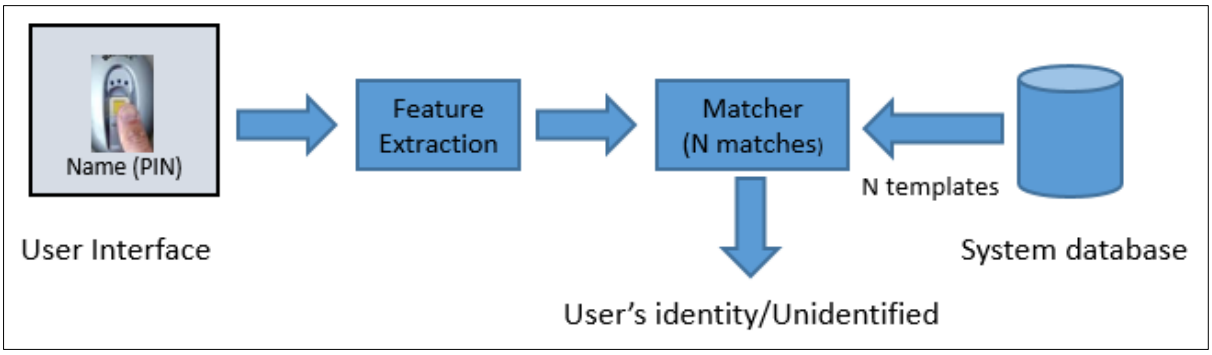

Figure 3.2: A biometric identification process

The key aspect of the process involves enrolment and authentication. The enrolment process creates a reference template of the user for use in any subsequent authentication. At this stage, the sample is based on the enrolment template in the initiation step, and it is pertinent to ensure that it is of good enough quality to be used. The authentication process covers the remaining

two processes in the biometric system by comparing an input sample of biometric against one reference sample in the verification system, and with all reference samples with identification system. This process starts with capturing a biometric sample and extracting the features using appropriate sensors. The sample must then be compared with the reference template. Finally, the threshold will define whether the sample is valid if there is a match between the sample and the reference. If it does not, the decision will be to reject it.

3.3 Biometric System Characteristics

Biometrics have number of characteristics that describe various aspects of the system. Thus, setting certain criteria for the biometric characteristic or trait is very important as it can often play a significant factor in whether the biometric will work in a particular application. The characteristics are Jain et al. (2004):

- **Universality**: refers to the degree in which an individual has the necessary biometric sample. For instance, whilst fingerprints are a popular biometric, there is a subset of the population without the use of their hands and for whom such an approach would not be suitable.
- **Uniqueness** refers to the degree in which the technique is unique. There are differences between biometrics characteristics, for example, physiological techniques tend to be more unique than their behavioural counterparts.
- **Permanence** indicates how long the biometric characteristic is constant (without any change) over time. For example, fingerprints of a person are very resistant to change over time; however, the way a person walks varies based on age, fitness level and even the weather (Bohannon, 1997).

- **Collectability** refers to the ability to collect the biometric sample from the user. For example, an iris image needs a special camera with infrared, and it takes a little bit longer compared with a facial recognition, which only needs a couple of seconds using a normal camera.
- **Performance** is a measure of the recognition performance that can be achieved. The performance is based upon a range of factors such as uniqueness and permanence but is also directly related to the underlying matching sub-system.
- **Acceptability** refers to how acceptable a certain biometric is to the population of users.
- **Circumvention** is the means or the ability of the system to be circumvented or tricked. For instance, fingerprints biometric can be fooled by using a fake finger.

Whilst all biometrics contain these characteristics no biometric technique exists that meets all of them perfectly. As such, the selection of a biometric technique is often dependent upon what the developer wants to achieve – with higher and more secure biometrics typically being more inconvenient and thus less acceptable. The nature of where the biometric is to be deployed also tends to self-select or reduce the available techniques.

3.4 Biometric Performance

A function of the matching system, which effectively provides a measure of similarity, is to introduce error rates, which subsequently are used to determine performance characteristics. According to Kindt (2013), a variety of metrics are used to rate the performance of a biometric system. However, the most common performance metrics in verification system are False Acceptance Rate (FAR) and False Rejection Rate (FRR).

- FAR is the possibility that the system incorrectly gives permission to non-authorized persons due to incorrectly matching the biometric input with a template.

- FRR represents the possibility when the system does not grant access to an authorised user and considers it as an unauthorised user.

(Jain et al,2011)

As shown in Figure 3.2 below, the relationship between the two rates is mutually exclusive, which means when one rate decreases, the other one increases. Clarke (2011) indicates that the relationship between the two rates is quite special. He emphasises that there is a strong correlation between them and that this will be clear in the trade-off for the system designer between low user convenience and high security (tight-threshold setting) or high user convenience and low security (slack- threshold setting).

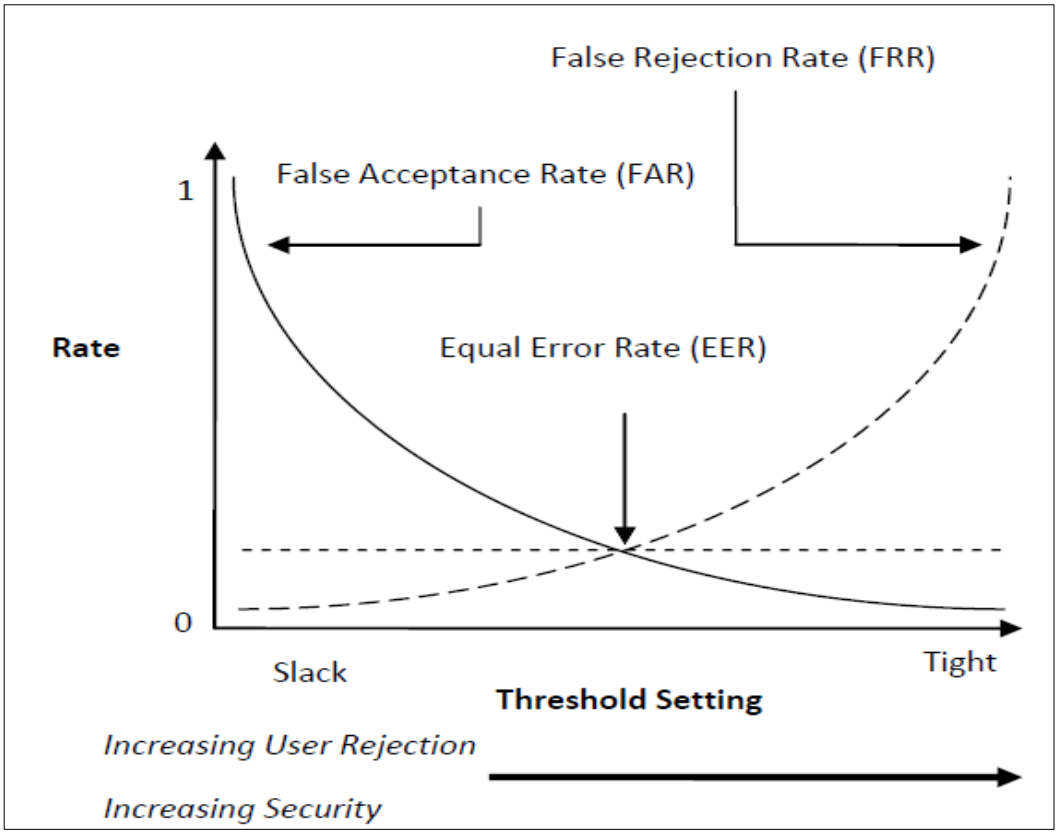

Figure 3.3 : Biometric Performance Characteristics (Jain et al.,2011)

In order to find an appropriate way to compare the performance of different biometric techniques, the Equal Error Rate (EER) can be utilised as a measurement parameter (Kindt,2013). The EER is a value which portrays the point where the FAR and FRR are equal.

Li (2012) states that the recorded performance of a biometric system is seriously dependent on the number of participants, the uniqueness of each participant and the sophistication of the employed classification approach. Thu, biometric technique with a smaller EER tends to be more accurate.

In contrast, other performance parameters may be used when designing a biometric authentication system, such as Failure To Acquire (FTA) and Failure To Enrol (FTE).

- FTA means the rate of biometric samples not being successfully extracted at the capture stage to initiate a valid template.
- FTE gives the rate of the user who is not able to complete enrolment onto the system.

In identification mode, the performance characteristics are calculated differently. Indeed, with a number of (N) identities enrolled, output will be a set of identities corresponding to the top *t* matches ($1 \leq t << N$) ,where, the identification rank is defined as the rank of user's that has the correct identity in the top *t* match return by the biometric system.

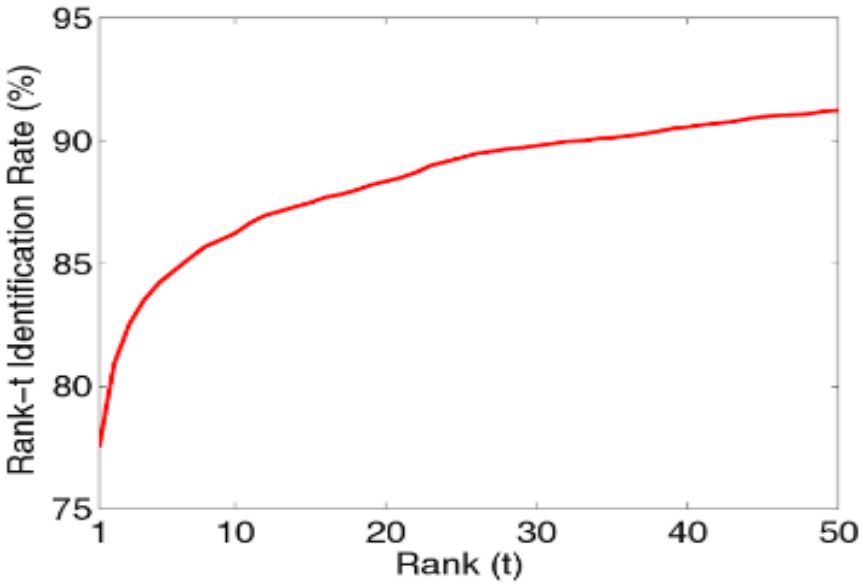

Figure 3.4 : CMC curve example (Jain et al.,2011)

The identification system error rates can be measured by calculating two common performance metrics as outlined below (Jain et al, 2011)

- True Positive Identification Rate (TPIR): occurs when the input to the biometrics system truly matches one or more templates in the database.
- False Positive Identification Rate (FPIR): occurs when the input to the biometrics system falsely matches one or more templates in the database.

To summarise, different values of *t*, the Cumulative Match Characteristic (CMC) curve is used to produce rank-t identification as demonstrated in Figure 3.3 below. As Figure 3.3 illustrated, the overall performance of Identification rate increases when the value of t increases.

3.5 Literature of Behavioral Profiling

The research in this area started in the 1990s and since then, there have been a variety of studies that have examined the behavioural profiling from different perspectives, such as fraud detection, intrusion detection and authentication. Behaviour profiling is utilised to verify a user based on the previous history, it then creates a user template which can be used to decide whether this kind of activity belongs to a legitimate user or not. The sections below discuss the different systems that utilised behavioural profiling.

3.5.1 Fraud Detection

Fraud detection is an issue that many organisations in many countries suffer from. The measurement of the efficiency of each technique mainly used detection rate (DR) where the system correctly finds out the suspicious activities. One of the earliest studies on fraud detection was part of the European commission-funded ACTS ASPeCT (Advanced security for Personal Communications Technologies) project (Burge and Taylor, 1997). This study was

conducted on the Vodafone network in the UK. The approach used behavioural profiling to detect any abnormal activity on the number. Toll ticket tools used to extract all the necessary information about calls activities. In addition, two methods for utilised behavioural profiling in this study Current Behaviour Profile (CBP), which means the short sequence of activity and Behaviour Profile History (BPH) to save the long sequence of activity and both of them are managed based on the queue system. Over a two-month period, they used a neural network technique with an unsupervised approach where no fraudulent examples are required for training to classify the call activity, based on three parameters: national calls, intentional calls and use for supplementary reasons. Ultimately, the approach detected 75% of fraudsters and only 4% of misclassifications of valid subscriptions were included within the fraudsters list.

In 1997, Moreau et al. published a paper in which they presented the first prototype of a tool based on a supervised neural network, an unsupervised neural network and knowledge-based systems. This research provided a framework for the detection of fraud in mobile communication as a part of the European commission-funded ACTS ASPeCT project. They gathered the information for the study from toll ticket because it includes all the necessary information that could help in the detection process. The features were International Mobile Subscriber Identity (IMSI), the start date of calls, the start time of calls, duration of calls, dialled telephone numbers and national/international calls. It is quite unacceptable to detect just 64% of fraudsters by using the unsupervised approach, which has a 5% false alarm rate, due to the fact that a huge number of fraudsters will not be detected, while the main reason for this kind of technique is to determine what the problem with the communication is and to identify the potential solution to mitigate it. The other approaches are significantly better, but they still have a high number of false alarms that may disrupt communications for the valid subscribers (Moreau et al.,1997).

Samfat and Molva (1997) investigated the ability to perform intrusion detection in the visited location and within the duration of a typical call. The experiment was carried out in the simulation environment and the dataset has a similarity of GSM network, then applying a statistical classifier in order to have a user behavioural based on location and telephony features, such as the start and end of call and duration. The results of this experiment were 82% detection rate and about 40% FAR, which means a huge number of the illegitimate user is considered as an authorised user. Although the detection rate is quite satisfying, FAR is quite high which may need further research to fix it.

In 2002, Boukerche and Notare provided a fraud detection model on mobile phones by utilising a Radial Basis Function (RBF) neural network model. In order to detect call service fraud, the model uses user-calling features such as time and duration of the call. The model sends an alert message to the user and network administrator immediately when a suspicious call is identified. The dataset was collected by an unknown telecommunication service provider and contains 4,255,973 telephone calls. The result achieved by applying the RBF neural network and using 110 neurons in the hidden layer of the network, was 97% detection rate and 4.2% system error rate.

Hilas and Sahalos (2005) published a paper titled “User profiling for fraud detection in telecommunication networks”. The authors mention that the amount of telecommunication fraud has dramatically increased. Hence, searching for effective detection is important to avoid further fraudulent activities. The paper suggests obtaining user profiling to detect fraud activities by using a statistical machine learning method, which constructs user profiles of each user and compares them with the future activities of the users to detect any abnormal activities. The methodology of this paper is to create a profile for each user from their previous history,

which were stored on the organisation's database. Consequently, the authors mention the similarity measure, which consists of different parameters, such as calls made to local destinations, the duration of local calls, the number of calls to mobile destinations, the duration of mobile calls, the number of calls to national and international destinations and their corresponding duration. There are two levels of similarity, the first being to examine the equality of the number of calls in the same category, and the second is to compare the total calls durations across categories. The dataset contained more than 5,000 users, and it was collected over a one-year period at a university (Hilas and Sahalos, 2005). The technique used was the statistical machine learning to present that constructs user profiles for the detection of fraudulent activities in telecommunications networks. The experiment relied upon a different number of parameters and extracted these parameters from the call inquiries in the large dataset. In addition, the greatest similarity achieved was 80%, which is a promising result. Furthermore, this paper does not mention anything about malicious activities from a particular user. It merely compares known parameters for known durations for known users, and that might not be enough to detect fraudulent activities in the telecommunication sphere. The same authors in 2007, tried to perform another classifier to improve the detection rate, however, the result almost remains about 80%

Ogwueleka, (2009) published a paper titled “Fraud Detection in Mobile Communications Networks Using User Profiling and Classification Techniques”. This study utilised call data for describing behavioural patterns of users by using unsupervised Self Organisation Map neural network (clustering). The evaluation of this experiment was done using a dataset that includes 180 participants during a period of 75 days, in which, the results of this research concludes that FAR is about 3% which means the fraudulent transactions that accepted as

legitimate and the mobile communication detection system detect most of the fraudulent transactions.

Similarly, Oayyum and Mansoor in 2010 have stated that in order to improve the detection rate of the fraud issue, two techniques could be utilised; let the user think what is the most feature that may represent the fraud and create a weight that is association priorities based on user input. The values that were given from the users to the most features that might represent the fraud by giving the highest features high value to presenting the level of dangerous that has. The study used a neural network as a classifier and evaluate the technique using a dataset that includes 300 participants. The final result from this approach was an ability to detect up 70% of the fraudulent calls.

In 2015, Subudi and Panigrahi have published a paper titled "Quarter-Sphere Support Vector Machine for Fraud Detection in Mobile Telecommunication Networks". This paper has utilised a set of features such as call duration, call type, and call frequency along with location and time. The experiment was carried out by using the Reality Mining dataset which includes 94 users' information. The authors utilised support vector machine classifier in order to distinguish the malicious behaviour from the normal behaviour. The result of this approach was promising compared to the previous studies in this fields where the detection rate attained is 97%. The aforesaid studies found that using behavioural profiling in fraud detection technique provide promising results once the behavioural profiling template creates with sufficient information. With respect to performance, many studies have achieved positive DR and FAR and over the time where more information became available and being able to include when the user profile created such as user location.

3.5.2 Intrusion Detection System

This section explains using behavioural profiling in IDS which is analogous to anomaly detection. Buschkes et al.,1998 has introduced a new approach that using anomaly detection; they produced their procedure to test their hypothesis by using Bayes decision rule; In probability theory and statistics, Bayes' theorem (alternatively Bayes' law or Bayes' rule) describes the probability of an event, based on prior knowledge of conditions that might be related to the event to detect any abnormal usage (Zou,2006). The study mainly focused on different features, such as call activities and location of users. The result shows that the system was able to score 87.5% detection rate.

Similarly, Sun et al. (2004) have published a paper in titled "Mobility-Based Anomaly Detection in Cellular Mobile Networks". This paper examined cell IDs transferred by a user as the features value. In addition, various services were included in this study for a user as an option in order to create an accurate user profile. Consequently, a higher order of Markov model for the detection part was used. The paper has applied the concept to a simulation environment for evaluation purposes and the final finding form this research was 87.5% Detection Rate (DR) and 15% False Alarm Rate (FAR). The second experiment that was conducted by Sun et al in 2006 has achieved a better DR (89%) and low FAR (13%). This is due to more care which was being considered in relation to the location of the user. Nevertheless, location information is quite sensitive for users, in which this approach gives the user an ability to turn it on/off as required.

Yazji et al.(2011) published a paper in the title "Protecting private data on mobile systems based on spatiotemporal analysis". The research sought to produce a model used to detect theft in smart phone mobile by creating an anomaly spatiotemporal patterns. They used network

access patterns and file system activities to build a behavioural model that permitted attack detection with a latency of 5 minutes and an accuracy of 90%. The study evaluated on Reality Mining dataset consisted of 100 participants. The authors used two techniques spatiotemporal and Trajectory. The former used when the users do not have many locations in their dataset whereas the later utilised with sufficient locations, in order to improve the performance. The result shows that the system was capable of detecting an attack within 15 minutes with 81% accuracy.

Finally, Yazji et al. (2014) produced an approach that addressed mobile device issues by making an effective correlation between the user's location and time data. They developed two statistical profiling approaches for modelling the normal captured behaviour of the user: the former based on an experimental collective probability measure and the other one is based on Markov which provides a way to find dependencies between the current information and compare it with the previous one that owned (Clemson University,2017). Reality Mining and Geolife dataset were utilised to evaluate this approach. The result shows that the system is capable of detecting a potential intrusion with 94% accuracy. This high proportion of accuracy achieved due to the set of features utilised in this research and the duration that was taken into account in order to create a user behaviour profile template, which was 9 months.

3.5.3 Authentication

The majority of authentication-based systems rely on a behavioural-based mobile security system. These systems use different techniques to assess user's identity based on one or more behavioural characteristics, such as keystroke and gait. The following section provides a short description of these studies.

Aupy and Clarke (2005) conducted a preliminary study to determine at which level people present a unique behavioural profile while utilising their computer desktop. The paper aimed to provide a transparent and continuous authentication of users relying on their interaction with computers, what applications are used and when, and recently visited websites. The authors used the 'logger' application to capture user interactions. The experiment sample was 21 participants, and the data was collected over 60 days. The study utilised different features extracted by the logger application, as shown in Table 3.2 below.

Code	Continuously action update
KEY	When the word has been typed in, the title of the window where it has been entered and the word is recorded.
OPN	This records the name and class of the opened window.
CLO	This records the name and class of the closed window.

Table 3.2: Logger Features Extraction

The study used a neural network classification known as the Feed Forward Multi-Layered Perceptron (FF-MLP). FF-MLP has been commonly used in implementing the K-classification module for the character recognition (Oh and Suen, 2000). The result of the study was promising, with an overall EER average of 7%. Finally, the study faced some problems, such as the number of participant's actions being very small. Therefore, the authors decided to repeat a number of actions from the dataset many times to raise the input features.

Yazji et al. (2009) proposed an implicit user re-authentication system for the portable computer that has limited in application change and hardware replacement. The main aim of this research is to reduce the probability of unauthorised access or theft by creating a user behaviour based on a combination of filesystem activities and network access that is able to distinguish between

normal user and anomalous. There are a set of features that can be obtained from file access system where the log file has a lot of useful information such as timestamps, names of the process responsible for access, locations of accessed files, and operations on the file. Different parameters of network activities were included in this study, such as the timestamp of each access, source IP address, destination IP address, protocol identifier. The study utilised K-means clustering, and in order to evaluate the approach, a period of two weeks was allocated to gather a dataset of 8 participants. The result was promising where the system was capable to score detection rate of 90% with FAR 14% (it means 14% of nonauthorized access considered as authorised) and FRR 11% (it means 11% of authorised access considered as unauthorised access or imposter).

Li et al. (2010) published a paper titled "Behaviour Profiling on Mobile Devices". The study proposed a novel behaviour-based profiling technique that is capable of building up on the weaknesses of current mobile device authentication systems by developing a comprehensive multilevel approach to profiling. One of their main features collected form device usage includes the day, time, duration and weekday. The approach was evaluated using 30 users from MIT Reality dataset and Neural network with different neuron size was applied. The findings of this are; 35.1% average EER, where top user achieved 1%EER. With this high number of EER, the authors proposed another approach that is concentrating on mobile user's application usage. They used a behavioural-based host mobile security mechanism, using some of the user's behaviour characteristics to assess the legitimacy of current users. The experiment was conducted using a dataset that is available from the MIT Reality Mining project. The dataset was collected over 10 months as the duration of time starting from September 2004 to Jun 2005 for 106 participants. There were different types of features collected from the mobile including application level (default application, such as a phonebook), application specific (additional

discriminating information, such as Internet browser) and multi-instance application (a mixture of application level and application specific). The user behaviour was created based on two types of profile techniques: static and dynamic. After applying the data on the mathematical equation, the results were 13.5%, 5.4%, 2.2% and 10% for the general application, telephone, text messaging, and multi-instance application usage respectively. This result seems to be encouraging and points to the use of this method in order to identify misuse within the mobile phone sphere. However, some limitations may affect that decision, such as the dataset was from 2004/2005 when smart phones first emerged, hence the data is not up to date. The entity data is not sufficient for various reasons, because, for instance, applications have changed since 2004 (Li et al, 2011). Therefore, an additional experiment has conducted on a dataset consisted of 76 users in order to build a novel behavioural profiling framework that is able to collect user behaviour based on applications usage. So, the system would reject the user's access based on a number of consecutive abnormal applications usages rather than single application usage. By applying this technique, the system has achieved better performance achieved 9.8%EER, FAR 4.1% and FRR 4.4%.

Similarly, Shi et al. (2011) published a paper titled "Implicit Authentication through Learning User Behaviour". The aim of this paper is to create a new model for performing implicit authentication. The model utilised different features, such as SMS and telephony along with Browser History: they examined browser history and recorded the obfuscated domain name of each URL along with the number of visits to this URL done previously. The technique was

evaluated based on two weeks of collected data from 50 users and the result was quite promising where the system achieved 95% detection rate.

Salem and Stolfo in 2011 produced modelling user search behaviour for masquerade detection technique. Their hypothesis relied on the level of knowledge that a user has about his system, in which encourages him to do a search for information in limited, targeted and unique fashion. Subsequently, this kind of behaviour might aid to distinguish legitimate users form masquerade. There were 18 participants involved in this experiment and a support vector machine as a classifier to detect the masquerade. The result showed that by using this kind of user behaviour, the system was able to detect all masquerades with a very low false positive rate of 1.1%.

In 2013, Abramson and Aha created a technique for continuous authentication purpose based on user web browsing behaviour. They claimed that using web browser has become a common activity in different places for different purposes, thus there is a chance to uniquely identify someone based on their Web browser behaviour. The main features of this technique come from the type of web browsing, number of sessions, length of session, and timestamp. The experiment was conducted using a dataset that included 10 users by using ensemble classifier. This classifier is often more accurate due to a set of classifiers whose individual decisions are joined in some way to classify a new example (Dietterich.T,2001). The results show that 24% best EER achieved.

Similarly, Fridman et al. (2015) presented an active authentication model on mobile devices based on user behaviour that is collected form four biometrics modalities: Text enters via soft keyboard, Application usage, Web site visited and Physical location via GPS. The evaluation

of this system was conducted based on a dataset from 200 participants collected for at least 30 days. Decision Fusion classifier has been utilised in order to improve the overall accuracy of the system. The result showed that high accuracy can be achieved with EER 3%.

In 2016, Sbeyti (2016) suggested a novel authentication model be used complementary to the existing models. Indeed, the research sought to produce an implicit authentication through the capture of user's discriminative behavioural pattern. The system aims to reduce the number of explicit authentication. Its purpose is not to replace the common authentication approaches, but rather to accompaniment them. That is, the user can still use his chosen authentication method, but once the phone is unlocked, the implicit authentication takes responsibility to determine if the user is indeed the owner or an attacker. And to accomplish this, the author relied on two features applications usage and duration of each usage for each particular user. Then a mathematical algorithm has implemented and being work at the run time to determine if the user is an owner of the smart phone or an attacker.

The experiment evaluated on Android dataset consisted of 30 users and the result shows that the system was able to achieve True positive attained 76%.

3.5.4 Identification

Much of the prior art has focused on the application of behavioural profiling in the simpler verification mode. In identification mode, there are a few papers discussed using behavioural profiling model. Yang (2010) published a paper about web user behavioural profiling for user identification. The author tries to build a user profile based on web usage patterns. He creates a user profile for different known users based on their web activities and anonymous web sessions.

The framework using Lift-based profiling; performs best when the number of users is small; consists of different stages, starting from initiating a user behavioural pattern, including the number of web sessions, overall user behavioural patterns, retrieve patterns for all the users, and measure the distance between two profiles. The dataset consisted of 50,000 users over the one-year period, and it was provided from a commercial data vendor. It included the web browsing history of a panel of users who were included in this experiment. However, the author selected a few users (with id # 2, 5, 10, 20, 50, 75, 100) and chose 10 different sliding window sizes,1, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 in order to investigate how sliding window size affects the predictive accuracy. The accuracy of the classification methods showed that the level of accuracy attained 90%. Regardless of the fact that accuracy rate was promising, the study has few serious issues which may affect the reliability especially when the author has evaluated the approach on just 10 users from the large dataset that contained 50,000 users.

Similarily, Herrmann et al. (2010) proposed a re-identification model based on behavioural profiling that is collected form web users' sessions in order to overcome the changing of IP addresses issue. They mentioned that the websites that are retrieved by an individual reflect – at least to some degree – his or her interests, habits and social network. The URL of some pages may even disclose the user's identity. So, they tried to ink the web sessions of a given user only based on a record of his past activities on the web. To overcome some privacy issue the features of the study were collected form HTTP requests, in which they used destination host name only. The evaluation of the approach was conducted in a real-world dataset from 28 users by using Naïve Bayes classifier and assuming a limited knowledge of an attacker who can only observe the host names visited. The result showed that the system was able to re-identify up to 50 % of the users, about 80 % of the time.

This technique relied on the observed access frequencies of hosts based on the name of destination hosts in order to create user profiles. There is a scalability issue in this study where just 28 users utilised to gauge whether the provider can correctly identify their users or not based on web site visited history. In addition, perhaps using sophisticated attacks, it can be argued that this study is not going to be efficient.

3.6 Discussion

The table below summarises the different studies of behavioural profiling. Although fraud detection and IDS research have contributed in a view of understanding the technique, further research is needed with reference understanding the applicability of behavioural profiling within a networking-based application. A variety of studies have examined user behavioural profiling from different perspectives such as fraud detection, intrusion detection, authentication and identification. The techniques mainly utilised to verify a user by storing the pervious user activities that used it to create the user profile and comparing it with current user activities to be able to decide whether this user is legitimate or illegitimate. The nature of the behavioural profiling features has played a core key into creating an accurate user profile. In fraud detection where the majority of this research has been supported by communications and mobile companies, the main features that utilised to create a user profile extracted from telephony activities such as call number, time and duration and later country of the calling included. Thus, the accuracy does vary across studies with high FAR scores in some of them. This is because of the features that were used to create a user behavioural profiling were quite limited and the samples were quite huge. In contrast, there are various researches have been conducted to explore the possibility of applying user behavioural profiling to increase the level of security of the mobile device. Indeed, the early studies in this field have utilised anomaly- based detection to determent any abnormal activities with the mobile devices such as studies (10-15)

as illustrated in table 3.4. whilst recent research tends to create a user profile based on application usage within authentication and identification model and achieved a high accuracy. This is due to that nature of features that included which were contributing towards achieving a higher performance as can be seen in table 3.3 studies (16-28).

No.	Author(s)	BP source	Classifier	# User	Performance%	Category
1	Burge and Taylor, 1997	Telephony	Neural Network	110	DR=75, FAR=4	Fraud detection
2	Moreau et al.,1997	Telephony	Neural Network	600	DR=90, FAR=10	Fraud detection
3	Samfat and Molva,1997	Telephony	Statistical	non	DR=82.5, FAR=40	Fraud detection
4	Boukerche and Notare,2002	Telephony	radial Basis function	100	DR=97, FAR=4.2	Fraud detection
5	Hailas and Sahalos,2005	Telephony	statistical machine learning	5000	DR=80	Fraud detection
6	Hailas and Sahalos,2007	Telephony	decision tree	5000	DR=80	Fraud detection
7	Qgwueleka,2009	Telephony	Neural Network and probabilistic	180	FAR=3	Fraud detection
8	Qayyum and Mansoor,2010	Telephony	Neural network	300	DR=70	Fraud detection
9	Subufhi and Panigrahi,2015	Telephony	Support Vector Machine	94	DR=97	Fraud detection
10	Buschkes et al,1998	Mobility	Bayes decision	Non	DR=87.5	IDS
11	sun et al,2004	Mobility	Markov	Non	DR=87.5, FAR=15	IDS
12	Hall et al, 2005	Mobility	Decision rule	50	DR=50, FAR=50	IDS
13	sun et al,2006	Mobility	Markov	Non	DR=89, FAR=13	IDS
14	Yazji et al.,2011	Mobility	Spatio-temporal model and trajectory-based	100	DR=81	IDS
15	Yazji et al,2014	Mobility	non	100	DR=97	IDS
16	Aupy and Clarke,2005	PC usage	Neural Networks	21	ERR=7	Authentication
17	Yazji et al., 2009	File access activity and network event	K-Means	8	DR=90, FAR=14, FRR=11	Authentication
18	Li et al, 2010	Telephony, device usage, Bluetooth and network scanning	Neural Networks	30	EER=13.5,35.1, 35.7	Authentication
19	Shi et al,2011	Telephony, SMS, Browsing and Mobility	Gaussian Mixture Model	50	DR=95	Authentication
20	Salem and Stolfo,2011	file access activity	SVM	18	FAR=1.1	Authentication
21	Li et al, 2011	Applications Telephony, SMS	Neural Networks	76	EER=13.5,2.2, 5.4	Authentication
22	Dimitrios et al,2012	Telephony, SMS, Browsing	machine learning algorithms	35	DR=98.5, EER=1.6	Authentication
23	Abramson and Aha, 2013	web browsing	ensemble	10	EER=24	Authentication
24	Li et al 2014	Application usage	Neural Networks	76	EER=9.8	Authentication
25	Fridman et al,2015	text, application, web and location	probabilities	200	EER=3	Authentication
26	Sbeyti,2016	Application usage	Mathematic algorithm	30	TP=76%	Authentication
27	Yang,2010	web browsing	decision trees	7	DR=91	Identification
28	Hermann et al.,2010	web browsing	Naïve Bayes	28	DR= 50	Identification

Table 3.3: Behavioural profiling overview

There are different classifiers utilised from the studies above such as SVM, Neural Network, decision tree and so forth. However, it is clearly obvious Neural Network machine learning is a classifier that is mainly utilised with behavioural profile most of the studies in different categories due to its ability to solve the non-linear problem (Chtious et al.,1997).

According to the numbers of studies in Table 3.3, there were various sources to create a user behavioural profiling such as applications usage, text and web history, thus the performance does vary based on the features that extracted to create a user profile. The performance of the studies that relied upon one user behavioural profiling may not quite accurate compared with other researches that building their user profile from different behavioural profiling sources as demonstrated in Table 3.3. However, there is a chance to improve the performance by using different multi-biometrics system categories such as multi-sensor and multi-algorithms. Indeed, Combining the information from these differing sources is called fusion (Jain et al., 2008).

Based on the discussion of the aforementioned studies, it can be argued that using behavioural profiling could help in distinguishing users for various purposes in different performance based on the number of activities that can be provided in order to aid the system to create an accurate user profile. Subsequently, using behavioural profiling is an appropriate solution that may contribute in associating the insider misuse to the right person.

3.7 Conclusion

Although behavioural profiling has been used in different studies particularly in a verifications mode and scoring a positive performance, there are a few studies that have discussed and examined the identification mode by using behavioural profiling and shown that it is viable, even if performance could be improved to overcome some of the scalability issues. However, it has been shown that using behavioural profiling can contribute towards creating an accurate template of the user profile.

4 User Interactions: A Novel Feature Extraction Process using Network Traffic Metadata

4.1 Introduction

As highlighted in the previous chapter feature identification is a key element of the biometrics design process. Identifying discriminative features will enable better classification and better performance. Therefore, it is important when looking to use network-based traffic that more investigation needs to be undertaken to better understand what features might be more discriminative than others in order to create the most optimal user profile. This chapter begins with a discussion of the traditional network analysis techniques, packet based and flow based detection technologies, in order to understand how network data is typically modelled and why this would not be appropriate in this context. This leads to a novel proposition in terms of the creation of application level interactions for network traffic followed by a detailed examination and set of investigations in order to be able to identify and understand whether user level interactions can be derived from low-level network data.

4.2 Network Analysis Approaches

There are two different kinds of network traffic analysis approach that utilised in the existing tools for preventing /detecting insider threats which are packet and flow based.

4.2.1 Packet based Network Analysis Method

A network packet contains various data fields, including headers, trailers and payload; and it can be obtained directly from network traffic. The packet based network analysis method (also known as Deep Packet Inspection (DPI)) inspects not only the header information of a packet but also (more importantly) its payload data in a bit by bit manner to detect misuse. From a network observation point, the raw information of the packet is constantly analysed and

compared with various signatures of known threats; any matching could indicate the presence of an attack. With the aim of enhancing the performance and effectiveness of the packet based network analysis, a wide range of research has been conducted, including (Verizon,2014; Cisco,2014; Merkle.L,2008; Baehr et al.,1995). Also, a number of tools (both open source and proprietary) have been developed, including Cain and Abel, TCP dump, Wireshark, Xplico and Microsoft network monitor, assisting network security analysts with an easy capture of packet information and a better understanding of how the attack was formed.

Mahoney and Chan (2001) emphasize that the main aim of their approach is Packet Header Anomaly Detection (PHAD) for identifying hostile network traffic. Their algorithm relied on learning the normal values for each packet field at the different protocols, IP, TCP, UDP, and ICMP. They determined the normal and unusual behaviour during the training stage and conducted all experiments on the 1999 DARPA dataset. The outcomes from these experiments were different, however examining packets and fields in isolation, and by using simple non-stationary models that estimate probabilities based on the time since the last event rather than the average rate of events gave the best value with a detection rate of 65%.

Wang and Stolfo (2004) presented anomaly network intrusion system based on packet payload which called PAYL. PAYL was designed along with some objectives such as no human intervention "hands-free", high accuracy, and ability to operate within high bandwidth situation. The model segregates to the different protocol relying on port number and length for instance FTP has port 20 and 21, HTTP utilises port 80 and so forth. Then, they utilised n gram analysis, Mahalanobis and Z-string in order to analyse, compare and process the data. The evaluation of this model was conducted using 1999 DARPA dataset, in which the result was nearly 100%, accuracy was achieved with 0.1% false positive rate on TCP port 80.

To faster the process of packet analysis, Bolzoni et.al proposed new anomaly network intrusion detection system that consists of two parts; The first part is using self-organisation map (SOM) for classifying payload data, and the other part is a slight modification of the well-known PAYL system which was discussed above. The modification upon PAYL is by using SOM after the PAYL uses the class value given by SOM instead of the payload length. This approach made a combination between SOM with a modified PALY algorithm. The experiment was carried out based on the 1999 DARPA dataset. There were 97 attack instances (payload-based attacks) in the dataset. The duration of the experiment was 4 weeks randomly (weeks 1 and 3 for training with 244,591 attacks free packets and weeks 4 and 5 for testing). The overall detection rate result was 73.2%. Different studies have utilized classification approach to detect any abnormal traffic. In 2012, Al-Bataineh and White produced a study that used entropy and byte frequency distribution of HTTP POST request content as features to identify data stealing. The study relied on studying the behaviour of a data stealing bot, then provides insights into how it could be distinguished from another typical web usage of benign applications. After applying a classifier, the outcomes were 99.97% detecting rate on HTTP traffic.

He et.al (2014) proposed a novel network-based solution that relied on entropy technology and user's network behaviour to detect data exfiltration in the cloud environment. Because of the DPI cannot detect encrypted data leakage efficiently, this study tried to build a network behaviour profile based on features, such as network and application protocol type, destination IP, port and time of occurrence, and then applying Naïve Bayes algorithm to classify expected and unexpected results successfully. The final result showed that the rate of detection was 90% and the false positive rate was 1%.

Table 4.1 lists a number of studies that have been done by using packet based to detect several types of threats where the insider threats tools have been executed based on the same concept. Since information is encapsulated within the payload of the packet, the packet based network analysis can be effective against malware attacks and information leakage; also the method is utilised to detect other network related attacks, such as buffer overflow attacks, DoS attacks and network intrusions. Nonetheless, the packet inspection approach is a time-consuming process due to the nature of the bit by bit comparison.

No.	Studies	Applications	Methods	Performance
1	Mahoney and Chan (2001)	Anomaly based network IDS	Learns the normal ranges of values for each packet header field	65% DR
2	Wang and Stolfo (2004)	Anomaly based network IDS	Mahalanobis distance for the detection phase.	100% DR with 0.1% FP for port 80 traffic
3	Zanero (2005)	Anomaly based network IDS	Uses Self-Organising Map to analyse the payload of TCP packets	66.7% DR with 0.03% FP
4	Wang et al 2005	Zero-day worm detection	Correlates ingress/egress payload alerts to identify the worm's initial propagation and also automatically generates signatures.	Over 95% of DR with less than 0.5% of FP
5	Bolzoni et al (2006)	Anomaly based network IDS	Exams the packet payload via the combination of Self-Organising Map and a modified PAYL system	73.2% DR with less than 1% of FP
6	Wang et al (2006)	Buffer overflow attack blocker	Extracts instruction sequences from HTTP request and determines if the instruction contains malicious code	100% DR on HTTP traffic
7	Ahmed and Lhee (2011)	Malware detection	Classifies the packet payload based upon three categories: multimedia (e.g. jpg), text (e.g. asp) or executables.	FN (4.69%) and FP (2.53%)
8	Al-Bataineh and White (2012)	Detection of data exfiltration	Uses entropy and byte frequency distribution of HTTP POST request contents to detect the presence of data stealing botnets	99.97% DR on HTTP traffic

Table 4.1: Packet-based detection studies

4.2.2 Flow based Network Analysis Method

Flow-based approaches seek group IP packets passing through an observation point in the computer network within a certain time interval based on a (typically TCP/UDP) connection profile. All packets that belong to a specific flow have a set of shared properties. These properties may exist in the header or in other different parts in the packet itself or both (Wang

et.al, 2010). The advantage of using flows is the vast reduction in data that needs to be analysed in comparison with packet-based approaches. The flow record normally consists of various fields, such as the time and date stamps, the IP addresses of the communication source and destination, their port numbers, the length of the total payload, and the type of protocols. The flow is normally generated from the raw traffic by using third party applications, such as NetFlow, SFlow, JFlow and IPFIX (Wang et al.,2010; Ahmed and Kyung,2011; Al-Bataineh, 2012; Gaofeng,2014). These applications perform different tasks based on flow based analysis, such as traffic monitoring, identifying unauthorised network activity and tracing the source of DoS attacks. Typically, this is performed by analysing the current traffic flow and identifying any abnormality based upon the historical traffic profile. Subsequently, this concept of how the network traffic being analysed, many methods and tools have been proposed and devised within the flow based network analysis domain (Dharmapurikar et al.,2003). With increasing network bandwidth and encryption, the flow-based approach has become the main approach in the market in terms of investigating network traffic issues. In addition, the analysis in large capacity networks consumes more time considerably. Subsequently, this approach is more efficient in detecting network scans and intrusions, the widespread of malware attacks, and monitoring general network usage. However, although the approach solves issues of data volume and encryption, the approach is based on including all communication traffic.

As shown in Table 4.2, there have been number of studies conducted about using IP flow to detecting threats. Pao and Wang (2004) published a paper titled "NetFlow based Intrusion Detection System". They emphasised that the security approach needs to be improved to be able to detect insider and outsider attacks. Therefore, they produced a system that can detect several types of network attacks from inside and outside organisations. They created an access

control list using IP addresses and port numbers. The result of this system was being able to detect the Ping sweep, DoS and port scan attacks.

In 2006, Crotti et al presented a classification technique based on the statistical analysis of network traffic implemented at IP-level in order to find out a protocol fingerprinting. The study included three protocols HTTP, SMTP and POP3 for network traffic on both directions between client and server. The finding of the experiments shows that HTTP protocol was scoring the promising result where the detection rate attained 99.41% from the server and 97.46% from the client side.

Similarly, Song et.al (2006) produced a novel Flow-based Statistical Aggregation schemes (FSAS) for network anomaly detection. The technique relied upon an aggregation of traffic within a certain time to reduce the amount of data and improve the ability to handle a high amount of statistical and packet data. Five parameters were mainly focused on as listed below, source and destination IP addresses, source and destination port numbers and Layer 3 protocol type. Neural network Classifier (Back-Propagation Networks) was utilised, in which, the approach was evaluated using two datasets DARPA 98 dataset and CONEX testbed dataset and the result was quite positive with 94% detection rate and 0.2% false positive rate. Muraleedharan et. al (2010) applied chi-square detection mechanism to detect anomaly activates based on IP flow characteristics. They provided a solution to identify some of the anomalies attack like scan a flood attack by studying the behaviour of the network traffic, in which, the system concentrated on three protocols, TCP, UDP and ICMP that they selected.

They used different parameters in traffic analysis, such as the number of packets, average packet size in bytes, average flow duration in second, average packet per flow, number of single

packet flows. The total number of attacks was 15 and they belonged to two categories, which are Scan and Flood. The output of this system was compared with SNORT detection system to gauge the detection rate. The result was quite excellent where the detection rate was 100%, but there was a long delay compared to Snort. Another limitation that can be discussed is that this system was not able to detect other attacks, such as the worm. Moreover, various studies have been conducted to improve the detection rate of IDS verifies the IPs flow. For instance, Braga et.al utilized Self-Organization Map (SOM) to detect DoS attack. However, other researchers have used different classifiers to analysis IP flow in order to detect anomaly or malware attacks.

No.	Studies	Applications	Methods	Performance
1	Pao and Wang, 2004	Signature based network IDS	Creates an access control list by using IP addresses and port numbers	The method was able to detect the Ping sweep, DoS and port scan attacks.
2	Crotti et al, 2006	Traffic classification	Uses a statistical method to analyze network traffic at the IP flow level; traffic from both directions between the client and the server were considered.	99.4% and 97.5% DR for server's and client's HTTP traffic respectively
3	Song et al 2006	Anomaly based network IDS	Employs the back-propagation neural network to exam the statistically aggregated flow data	94% DR with 0.2% FP
4	Muraleedharan et al, 2010	Anomaly based network IDS	Utilises a chi-square detection mechanism to analyse IP flow characteristics	100% DR on 15 attacks
5	Braga et al, 2010	Anomaly based network IDS	Verifies IP flows by using Self-Organising Map to detect distributed DoS	99.11% DR and 0.46% FP
6	Winter et al, 2011	Anomaly based network IDS	Uses One-Class Support Vector Machines for the analysis of IP flow data	98%DR
7	Tegeler et al, 2012	Malware detection system	Utilises a statistical method to create training and identifying models for detecting botnet traffics	90% DR with 0.1% FP
8	Jadidi et al, 2013	Anomaly based network IDS	Utilises the Multi-Layer Perceptron neural network with a Gravitational Search Algorithm to analyse the flow data	99.43% DR
9	Hofstede et al, 2013	Anomaly based network IDS	Mainly employs Exponentially Weighted Moving Average (EWMA) for mean calculation algorithm	95% DR with 1% FP

Table 4.2: Flow-based studies

In conclusion, the aforementioned studies discussed a set of behavioural characteristics of the packet or flow based. Hence, by using anomaly approach, in depth knowledge of their features

has been acquired. However, looking to investigate to what extent using this technique can contribute towards identifying individuals activates is required. Subsequently, this work on networks is based upon anomaly detection and is NOT what we are looking to achieve. Indeed, this research is focused on the identification of people which is a more challenging classification problem than anomaly detection has.

4.3 Application-Level User Interaction Features

The aforementioned studies have relied on monitoring all network traffic; however, this method is arguably not effective due to noise that is introduced (e.g. any traffic that is not directly related to the user activity/problem at hand) . Whilst the packet-based inspection approach offers the opportunity to use the payload (which itself is likely to contain a very rich source of discriminative information) the increasingly and widespread use of encryption makes this more challenging. Proxies or gateways are a mechanism for allowing organisations to decrypt and encrypt such traffic but the increase in the computational overhead and increasing concerns related to privacy make this undesirable. Conversely, the nature of the Transport Layer Security (TLS) protocol, in that the session remains open for a period of time (in comparison to the stateless HTTP protocol) results in the flow-based approach not having the appropriate level of resolution or granularity required to understand user interactions. For example, several user interactions could exist within a single encrypted session).

Subsequently, it is necessary to devise an appropriate feature extraction approach that is a better measure of the user activity. Therefore, the focus of this research was on the derivation of user interactions from network traffic metadata. This approach is quite complex because in order to capture user interaction signature, both traffic directions (from user to application server and reverse) need to be investigated. Whilst intuitively the user action at the application level is

encapsulated in one form or another in the network layer when it comes identifying and extracting this behaviour it is not a simple task. In fact, packets responsible for the user action get transmitted with other packets which have different tasks such as, establishing connection, connection control, downloading web-page JavaScript and so forth. On top of that the traffic may encompass packets associated with other application layer and machine to machine related protocols which makes capturing a certain packet in both directions is a challenging task. The figure below (Figure 4.1) simulates a reality by capturing user interactions among different network traffic. Even the most simply web-applications would result in multiple data flows from a single mouse click.

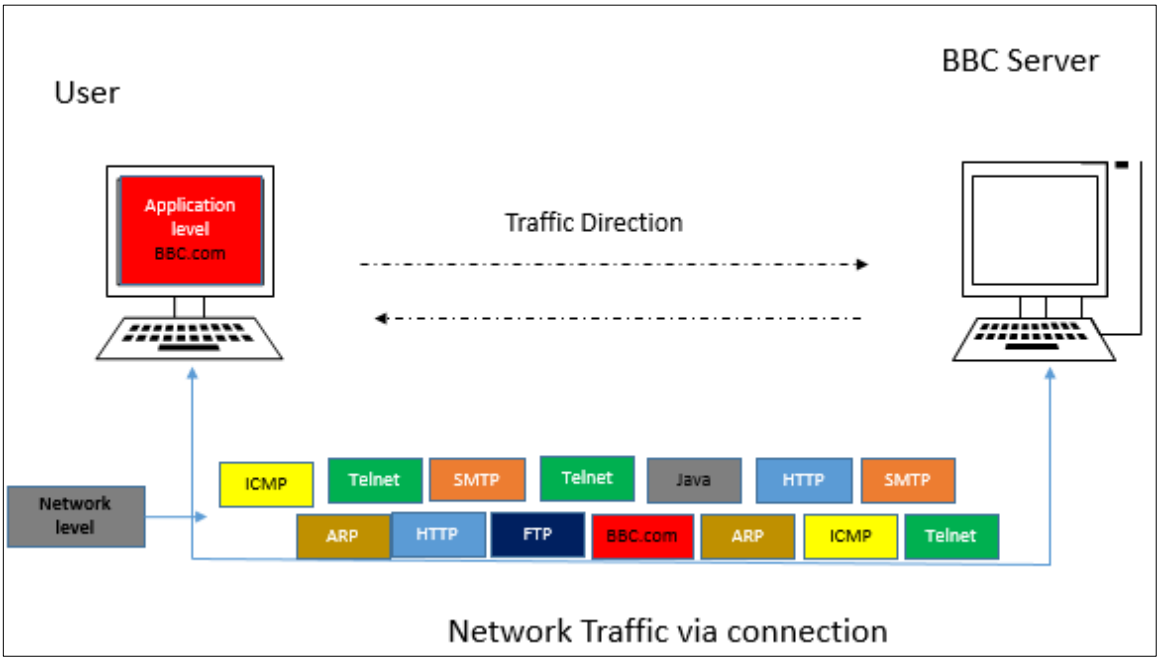

Figure 4.1 : Various protocols traffic in low network level

As shown in Figure 4.1, identifying user action, such as BBC.com at the network level is challenging due to the very large number of packets that may pass at the same time. Although the previous studies relied on capturing and analyzing the whole traffic between the two nodes, the user interaction approach has reduced the size of packets that were captured and analyzed by focusing and extracting just the packets that represent user actions from network level while omitting the rest.

Since using TLS has become popular; TLS is the standard security technology for establishing an encrypted link between a web server and a client's browser; efficiency of flow based approach of anomaly detection significantly decreased. This is due to a conceptual work of the SSL technique that establishes a secure channel between a user and an application server, in which, such a channel stays open until the connection has finished. Thus, the previous studies have suffered from a level of noise at their network traffic, because they had considered the connection between a user and an application server as a one interaction. However, scientifically this is might not be right because each user action on the application level has a specific IP flow that carries out the request and the respond in both directions, yield there might be more than one IP flow within the TLS connection. Thus, this research focuses on the IP flow of the user actions in both directions and omits other non-relevant flow. However, this technique increases the level of complexity in this approach. This is because a user may establish different TLS connections, thus, the number of packets from user to other applications and reverse that belongs to these TLS connections that will increase the level of noise and make capturing the associated user action IP flow more difficult. However, these features that can be obtained from the application level, will reflect the reality of user activities within those applications, which may contribute towards identifying individuals.

Accordingly, to determine whether user interactions could be derived from network traffic different web-applications across a range of categories needed to be investigated. For example, a search engine, online shopping, social network and so forth (web developer, 2015). Based on this concept, it was decided to choose certain services in different categories in order to make sure a variety of Internet services were included for analysis. For instance, webmail, social networking, news, online storage, file transfer, search engines, VoIP application, video/TV

streaming, online documents, and e-shopping, as shown in Table 4.3 below, have been selected to be part of this investigation.

No	Service	Example Applications
1	Information search	Google
3	Blogging/microblogging	Twitter
4	Video sharing/streaming	YouTube
5	Online documents	Google Docs
6	Online shopping (E-Commerce)	E-bay
7	Online encyclopedia	Wikipedia
8	Email web mail	Outlook
9	Social/Professional Networking	Facebook
10	News	BBC
11	Online storage	Dropbox
12	VoIP	Skype

Table 4.3 : Commonly used Internet services and example applications

The table below (Table 4.4) illustrates a sample of the user actions that are embedded within Skype application.

Service	Application	User Actions
VoIP	Skype	Text message
		Voice call
		Video call
		File attach
		Idle

Table 4.4 : Skype user actions

The section below discusses the viability of user interactions whether it is even possible to identify it or not.

4.4 Methodology

The approach is based on the theory that states how users interact with Internet-based applications on their computer produces a (relatively) unique network packet signature which can subsequently be used to identify the activity.

In order to test this hypothesis, an investigation was undertaken to determine to what extent these signatures could be developed. In the first instance, ten of the most popular Internet-based applications were selected for analysis in order to obtain comprehensive user actions date in various applications (eBizMBA, 2016). These applications are Google, YouTube, Skype, Facebook, Dropbox, Outlook, Twitter, Wikipedia, eBay and BBC. To ensure the resulting analysis is reliable, three researchers were tasked with the collection and analysis of network traces against a predefined set of network captures against user activities by following patterns validation model as illustrated in Figure 4.2, which themselves were repeated 20 times in order to allow for any variance in the resulting network traffic.

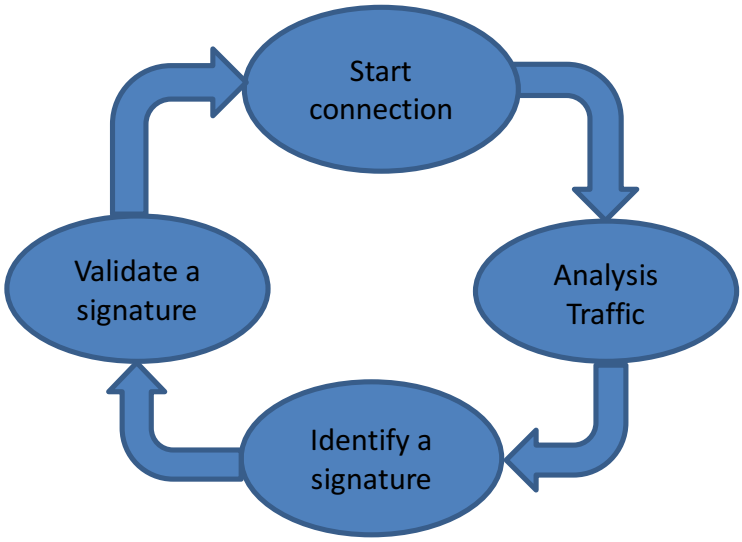

Figure 4.2 : Pattern validation

Consequently, each researcher swapped his list with another researcher to verify the same user actions and then comparing the final result that were obtained by the three researchers. Finally, the most constant and stable signature of user actions are going to be utilized in this research. The main advantage of this process is to make sure the three researchers identified the same user actions signatures. This is because some user signatures are not stable in their form or/and value. For instance, the text message user action signature in the Skype application achieved a different packet length and this is because when the user increase message length the packet

value of signature changes as well. Whilst the typing action on the Facebook application has a constant value but different number of packets that represents this signature. On top of that there are some unrelated packets that come alone between the signature packet as can be seen in Figure 4.3, which makes the feature identification process of user actions more difficult.

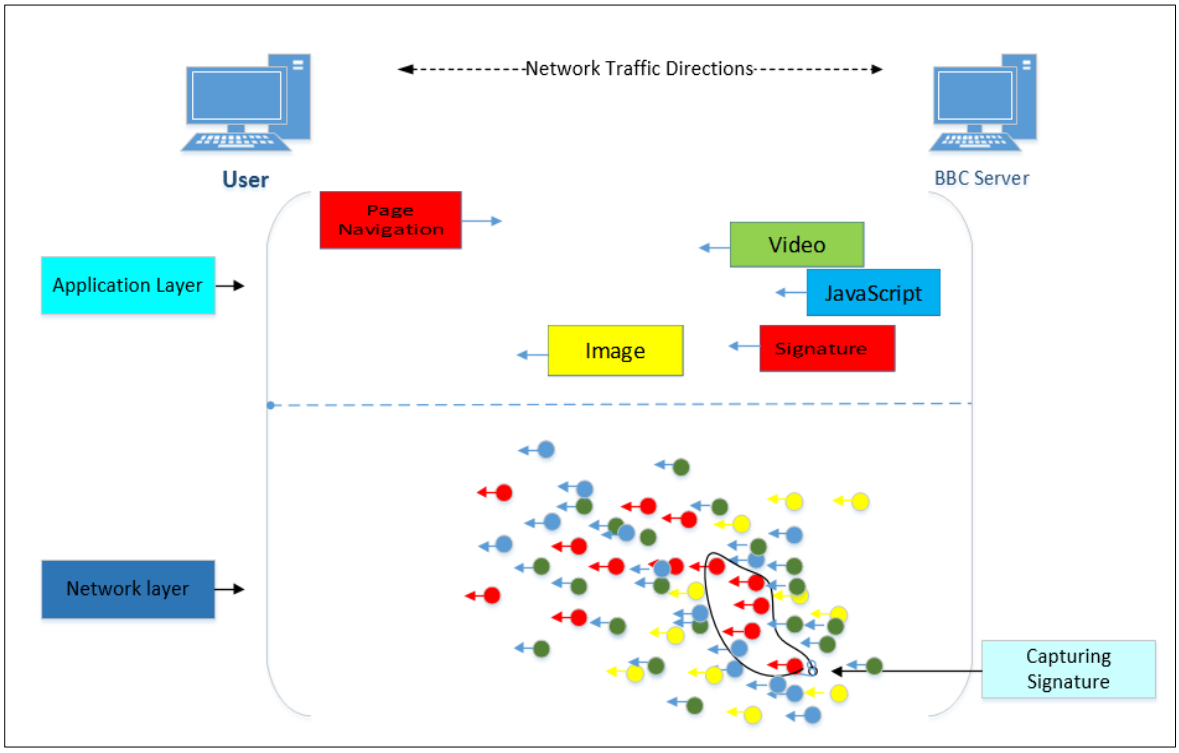

Figure 4.3 : Identifying user interaction from network traffic

Figure 4.3 demonstrates the level of complexity that exists in one of the investigated applications (which is actually the simpler BBC application). The monitoring process of the signature starts when the user presses on one of BBC sections at application layer. This request is then directed to the BBC server. However, the response from BBC server about this request is not limited to this action, it usually returns with a very large number of packets associated with it as can be seen in Figure 4.3. These packets that were generated from BBC server may represent different things, such as video, images, JavaScript and so forth. Hence, the challenge is to extract the related signature of the page navigation from this larger set of packets

associated with it. This noisy and variable output was a key consideration in the decision to assign three researchers for validating each of the patterns.

The table below (see Table 4.5) illustrates the whole user interactions across the different applications that are included in this study. The researcher assured that the main user actions in different applications is a part of this investigation, for instance YouTube application has two main tasks, upload and download,

Application Name	User Actions	Application Name	User Actions
Skype	Text messages	Outlook	File attachments
	Audio calls		Compose an email
	Video calls		Reply/forward an email
	File transfer		Insert a recipient
	Click on contacts	Facebook	Post
	Idle		Comment
Online documents (Google Docs)	Creation		Share
	Editing		Find friends
	Share with other s		Page loading
	Delete file		Attach files
Wikipedia	Search		Chat
	Download		Typing
YouTube	Watch videos	BBC	Page navigation
	Video upload		Watching video clips
Search engines (Google)	Keyword search		Listening to audio
	Page navigation	Dropbox	Download files
eBay	Browsing products		Upload files
	money transaction		Share files/folders
	Comment		Folder navigations
Twitter	Idle/Reading		
	Tweet		
	Download		
	Upload		
	Click Contact		

Table 4.5 : Applications user actions

Despite the fact that network traffic has got a numbers of features, the methodology that are used to obtained the user interaction has focused upon features that have a vital influence of representing user action at the application level. Therefore, a list of network metadata

parameters was utilized for the analysis, there are as follows: the date and time stamp, the IP addresses of source and destination, their port numbers, protocol ID (either TCP or UDP), the length of a datagram, and several TCP flags (e.g. SYN, FIN, ACK and PUSH). Based on these parameters, a number of features were collected for better representation of user interaction as illustrated in Table 4.6 below.

No.	Feature Name	Example
1	Start Time of interaction.	2014.11.11.10:48:19.769086
2	End Time of interaction.	2014.11.11.10:48:19.817979
3	Source port number.	58823
4	Service IP.	216.58.208. %
5	Service port number.	443
6	# packets send from (client to server)	2
7	Total size of packets sends (client to server)	2000
8	# packets send from (server to client)	6
9	Total size of packets sends (server to client)	2868

Table 4.6 : User Interaction Features

4.5 Experimental Result

There are 9 applications included in this investigation in order to find out a unique signature of user actions after excluding two application File transfer and eBay due to lack of usage of the former and noise signature for the later. The following sections will explain each application's user action signature and provide a summarized table of the signatures followed by discussion on how the researcher captured user actions from network traffic.

There are some abbreviation words mentioned in the following sections that need to be clarified before starting the analysis process, such as stream (10 packets or more transmitted from one side to another), multiply (between 3 and 8 packet convey and may not in sequence order) and MTU (maximum transmission unite).

4.5.1 News (BBC)

The BBC application has been chosen because of its popularity in the news category. However, its user interactions are quite limited to the three main tasks which are listed in Table 4.7 below, associated with their signatures.

User interactions	# packets	Total length (bytes)	Directions
Page navigation	Stream	Various	Server > Client
Watching video	Stream	MTU (Almost)	Server > Client
Listening to audio	Stream	120-280	Server > Client
Searching	×	×	×
Comment	×	×	×

Table 4.7 : BBC user action signature

Extracting a pattern from BBC website was a challenging step. This is due to the number of IPs address that appeared during the experiment and the ability to utilize all the data traffic during certain time was quite difficult as demonstrated in Figure 4.4. Hence, there were several attempts to validate the user action signature by repeating the capturing and analyzing processes of user action, and this step was done across all investigated applications.

BBC application has three different types of user actions identified as being highlighted, page navigation, watching video and listening to audio. The former signature is identified through a stream number of packets that have various lengths that are transferred from BBC server to the client side as can be seen in Figure 4.4.

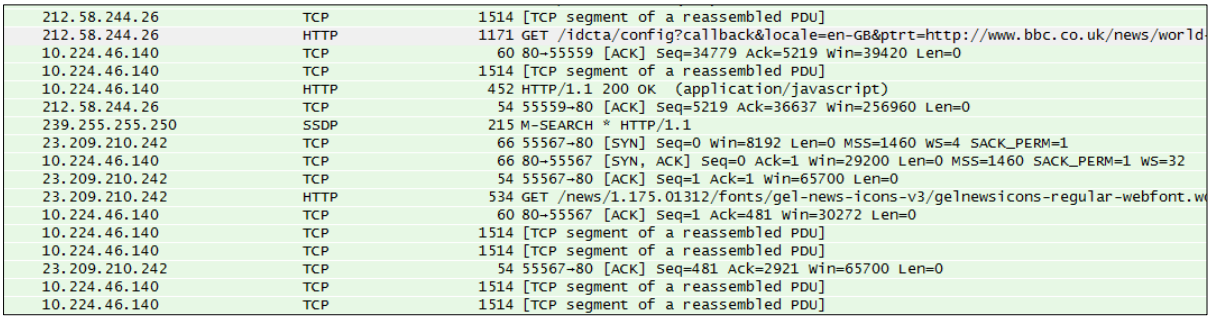

Figure 4.4 : Starting navigation on BBC

Regarding the other user actions (watching video and listening to audio), the length of packet is different as demonstrated in Table 4.7. This kind of differences can help to overcome of the

similarities that exist between the user actions signatures, such as the number of packets stream from BBC server to client with total length almost MTU and from the client to the BBC server mainly a stream of TCP ACK packets are observed as can been seen in Figure 4.5.

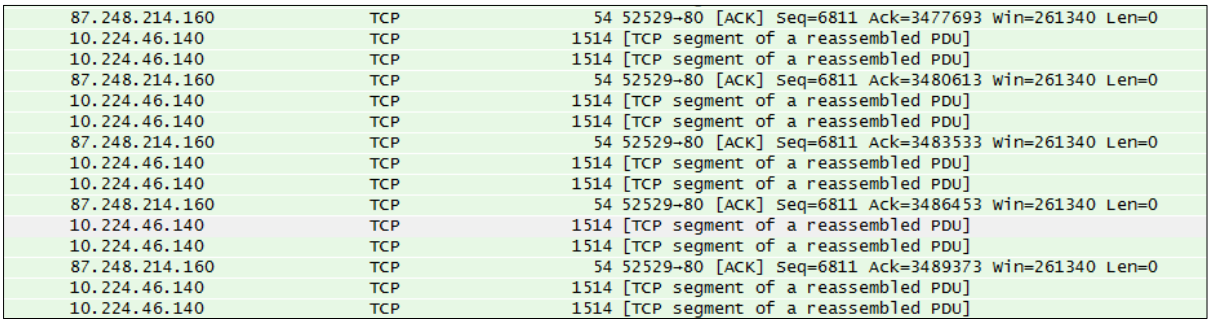

Figure 4.5 : Audio and video user action signature

4.5.2 Outlook

The process of Outlook services is quite complex. This is due to number of IPs addresses that appeared in the experiments were subject to change, which has increased the disruption of pattern extraction. The Outlook login website has number of IPs within the same range, such as 131.253.61.66 or 131.253.61.84 and after a user accesses its account, another IPs have raised up based on the interaction selected. However, there is a time of stability where the IP address of the Outlook application can be detected.

Identifying a signature that belongs to the Emails interactions is one of the main targets to this particular service. Therefore, this research has tried to cover all interactions within the Outlook application account. And, some of email interactions can be possibly determined as demonstrated in Table 4.8 below, in association with their signatures.

User interactions	# of packets	Total length (bytes)	Directions
File attachments	Stream	MTU (Almost)	Client > Server
Compose an email	One	971	Server > Client
Insert a recipient	One	971	Client > Server
Main page content	×	×	×
Send	×	×	×
Reply	×	×	×
Delete	×	×	×
Forward	×	×	×
Search for email	×	×	×

Table 4.8 : Outlook application signatures

1. File attachment

File attachment user action signature is relatively easy to identify. This is due to the form of its traffic at the network level as a stream number of packets transmit from client to Outlook server with total length almost MTU as illustrated in Figure 4.6 below.

192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TLSv1.2	1434	Application Data
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	411	[TCP segment of a reassembled PDU]
192.168.200.58	204.79.197.210	TCP	1434	[TCP segment of a reassembled PDU]

Figure 4.6 : Attach file

2. Compose email and Add recipient

In these user actions, four packets are transmitted from client to Outlook server. These packets have total size around 4450 bytes (this number calculated based on similarity of time stamp and identical source and destination). It is transmitted by PSH and ACK form. When the packets are safely delivered to the server, the server sends one packet with the size of 971 bytes in compose user action; as shown in Figure 4.7. While in add recipient the same length of packet send from client to application server.

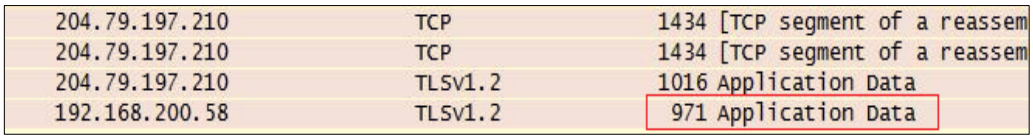

Figure 4.7 : Pres compose icons and add recipient

4.5.3 Skype

Skype application is one of the accurate applications in terms of type of signature. Indeed, it is considered as one of the applications that have many user actions from different categorizes and it is rich in signature side where the majority of its interactions are determined with a high level of accuracy as shown in Table 4.9. However, text message was the only action that was often altered, hence further investigation was carried out to explore this issue, in which, the researcher found that the length of the message was different, thus, the size of the packet was changed.

User interactions	# packets	Total length (bytes)	Directions
Text messages	One	794+	Client > Server
Audio calls	Stream	117-147	Both Clients
Video calls	Stream	1165-1365	Both Clients
File transfer	Stream	MTU (Almost)	Sending Client> receiving client
Click on contacts	One	747	Client> Server
Idle	One	587	Client > Server
Add contact	×	×	×

Table 4.9 : Skype application signatures

1. Online Text Message

The researcher was able to identify the length of the message by testing this user action many times over as demonstrated in Figure 4.8 below. The baseline of testing a message could be identified by sending a different length of text message form the client to the Skype server where the result of this process was determined in which 794 bytes is the baseline of text message user action.

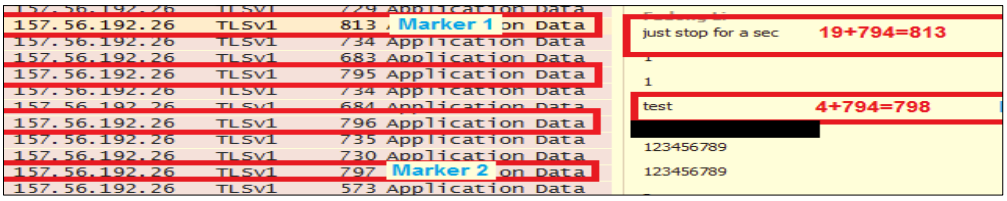

Figure 4.8 : Skype message

2. Audio user action

In audio user action, the connection is established between two nodes directly after the Skype server sets up the communication channel as can be seen in Figure 4.9. The connection is established by using UDP protocol and the data was flown from node 1 to node 2 with a range of sizes between 117 to 147 bytes. Once the audio call has finished the UDP connection also stops but the connection with Skype server remains valid.

192.168.200.58	192.168.200.63	UDP	98 Source port: 12354 Destination port: 57742
192.168.200.63	192.168.200.58	UDP	68 Source port: 57742 Destination port: 12354
192.168.200.63	192.168.200.58	UDP	119 Source port: 57742 Destination port: 12354
192.168.200.63	192.168.200.58	UDP	125 Source port: 57742 Destination port: 12354
192.168.200.58	192.168.200.63	UDP	111 Source port: 12354 Destination port: 57742
192.168.200.63	192.168.200.58	UDP	125 Source port: 57742 Destination port: 12354
192.168.200.63	192.168.200.58	UDP	120 Source port: 57742 Destination port: 12354
192.168.200.63	192.168.200.58	UDP	117 Source port: 57742 Destination port: 12354
192.168.200.58	192.168.200.63	TCP	72 60400→57742 [PSH, ACK] Seq=2839 Ack=2825 Win=
192.168.200.58	192.168.200.63	UDP	114 Source port: 12354 Destination port: 57742
192.168.200.63	192.168.200.58	UDP	119 Source port: 57742 Destination port: 12354
192.168.200.63	192.168.200.58	UDP	150 Source port: 57742 Destination port: 12354
192.168.200.63	192.168.200.58	UDP	119 Source port: 57742 Destination port: 12354
192.168.200.63	192.168.200.58	UDP	123 Source port: 57742 Destination port: 12354
192.168.200.58	192.168.200.63	UDP	122 Source port: 12354 Destination port: 57742
192.168.200.63	192.168.200.58	UDP	116 Source port: 57742 Destination port: 12354

Figure 4.9 : Audio action

3. Video interaction

This user action signature has the same signature of audio user action in terms of connection between the two clients without passing through skype server but the difference is in the size of packet, where it is between 1165 to 1380 bytes as demonstrated in Figure 4.10.

192.168.200.58	192.168.200.62	UDP	1362 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1199 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1198 Source port: 12354 Destination port: 50196
192.168.200.62	192.168.200.58	UDP	82 Source port: 50196 Destination port: 12354
192.168.200.58	192.168.200.62	UDP	1367 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1362 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1203 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1203 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1203 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1202 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1367 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1362 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1053 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1053 Source port: 12354 Destination port: 50196
192.168.200.58	192.168.200.62	UDP	1052 Source port: 12354 Destination port: 50196

Figure 4.10 : Video conference

4. Sending File

It can be considered as similar to the previous two user actions in terms of number and direction of the signature of user action packets. However, this user action has something different that can be used to distinguish it from others which is the total length of packets. Figure 4.11 shows that packet size is the main sign to differ sending file from others; the size of packet here is MTU.

10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TLSv1.2	1494 Application Data
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	832 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]
10.224.46.140	13.79.153.60	TCP	1494 [TCP segment of a reassembled PDU]

Figure 4.11 : Sending file

5. Click on Contact and Idle

These two user actions have a similarity in terms of number of packets exist in this signature. In both of them, the user actions signatures were identified as one packet pattern. However, the size of the packet is the main pattern for both of them that can be used to distinguish each user action from another. As can be seen in Figure 4.12, click on contact is represented by packets sent from client to the Skype server with size of 747 bytes by using TCP protocol and this amount is 587 bytes in idle action as shown in Figure 4.13.

141.163.187.157	TCP	60 443→59923 [ACK] Seq=1
157.56.193.43	TLSv1	747 Application Data

Figure 4.12 : press on contact

157.56.193.43	TLSv1	587 Application Data
141.163.187.157	TLSv1	155 Application Data

Figure 4.13 : Idle status

4.5.4 Facebook

Facebook application is one of the applications that has a number of user actions investigated. The majority of its user actions signature were less accurate comparing to other applications

user signatures, such as Skype signatures and it is quite complex to determine. For instance, typing user action is represented by two packets sent from client to the Facebook server, however, these packets could be transmitted with the same overall length and number of packets but the length in each packet changes every time. The table below explains the user actions and their identified signatures from network traffic followed by explanation of each one.

User interactions	# packets	Total length (bytes)	Directions
Page loading	Stream	Various	Server > Client
Attach files	Stream	MTU (Almost)	Client > Server
Typing	Multiply	1,502	Client > Server
Post	×	×	×
Like	×	×	×
Comment	×	×	×
Add or remove	×	×	×

Table 4.10 : Facebook application signature

1. Typing

When the user starts typing on Facebook, a number of packets are sent from the client to a Facebook server. The total size of two of these packets is 1,502 bytes (i.e. 1434+68 or 1169+333). These packets are sent in less than one millisecond timeframe (as demonstrated in Figure 4.14).

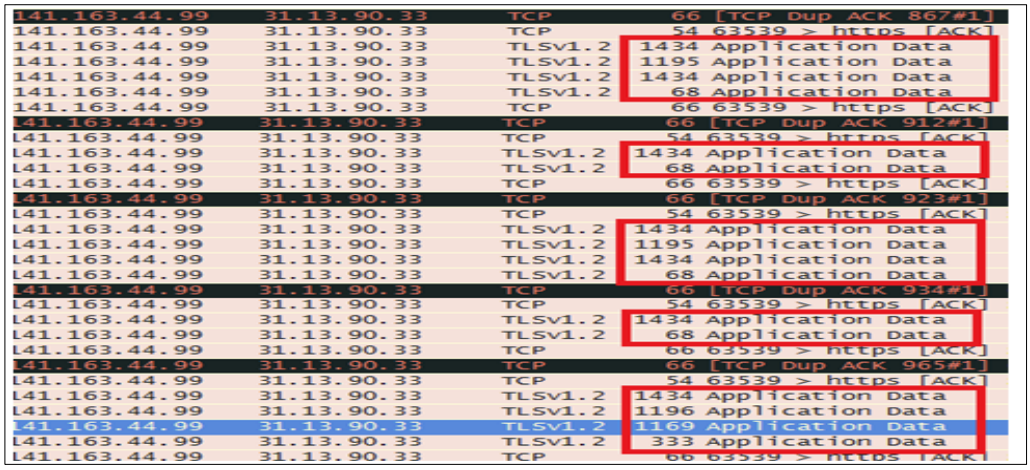

Figure 4.14 : User Typing on FB

2. Attach file

The files are transmitted from client to server in group of packet (MTU size over TCP protocol) as shown in Figure 4.15 followed by group of Acknowledgments back from server.

Filter: ip.addr==192.168.200.62 &&ip.addr==31.13.90.2 &&tcp Expression... Clear Apply Save

No.	Time	Source	Destination	Protocol	Length	Info
1098	2015-03-16 17:34:52.434681000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=162817 Win=16560 Len=0
1099	2015-03-16 17:34:52.434759000	31.13.90.2	192.168.200.62	TCP	1434	[TCP segment of a reassembled PDU]
1100	2015-03-16 17:34:52.434768000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Application Data
1101	2015-03-16 17:34:52.434775000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=165577 Win=16560 Len=0
1102	2015-03-16 17:34:52.434814000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Application Data
1103	2015-03-16 17:34:52.434824000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1104	2015-03-16 17:34:52.434828000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=168337 Win=16560 Len=0
1105	2015-03-16 17:34:52.434835000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1106	2015-03-16 17:34:52.434887000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1107	2015-03-16 17:34:52.434897000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=171097 Win=16560 Len=0
1108	2015-03-16 17:34:52.434904000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1109	2015-03-16 17:34:52.443200000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1110	2015-03-16 17:34:52.443232000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=173857 Win=16560 Len=0
1111	2015-03-16 17:34:52.443249000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1112	2015-03-16 17:34:52.443268000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1113	2015-03-16 17:34:52.443273000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=176617 Win=16560 Len=0
1114	2015-03-16 17:34:52.443279000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1115	2015-03-16 17:34:52.443285000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1116	2015-03-16 17:34:52.443289000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=179377 Win=16560 Len=0
1117	2015-03-16 17:34:52.443295000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1118	2015-03-16 17:34:52.443298000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1119	2015-03-16 17:34:52.443304000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=182137 Win=16560 Len=0
1120	2015-03-16 17:34:52.443310000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1121	2015-03-16 17:34:52.443312000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1122	2015-03-16 17:34:52.443316000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=184897 Win=16560 Len=0
1123	2015-03-16 17:34:52.443322000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1124	2015-03-16 17:34:52.444189000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1125	2015-03-16 17:34:52.444200000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=187657 Win=16560 Len=0
1126	2015-03-16 17:34:52.444208000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1127	2015-03-16 17:34:52.444212000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1128	2015-03-16 17:34:52.444223000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=190417 Win=16560 Len=0
1129	2015-03-16 17:34:52.444229000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1130	2015-03-16 17:34:52.444232000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1131	2015-03-16 17:34:52.444240000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=193177 Win=16560 Len=0
1132	2015-03-16 17:34:52.451196000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1133	2015-03-16 17:34:52.451214000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1134	2015-03-16 17:34:52.451225000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=195937 Win=16560 Len=0
1135	2015-03-16 17:34:52.451241000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1136	2015-03-16 17:34:52.451257000	31.13.90.2	192.168.200.62	TLSv1.2	1434	Continuation Data
1137	2015-03-16 17:34:52.451262000	192.168.200.62	31.13.90.2	TCP	54	57827 > https [ACK] Seq=16382 Ack=198697 Win=16560 Len=0

Figure 4.15 : File Attach

3. Page loading

Figure 4.16 demonstrates the main page load action, when a user presses main icon, stream of data transfers from Facebook server to the client and the size was MTU.

Filter: ip.addr==192.168.200.62 &&tcp Expression... Clear Apply Save

No.	Time	Source	Destination	Protocol	Length	Info
3865	2015-03-27 16:48:11.986107000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3866	2015-03-27 16:48:11.986127000	192.168.200.62	31.13.90.6	TLSv1.2	189	Application Data, Application Data, Application Data
3867	2015-03-27 16:48:11.986225000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3868	2015-03-27 16:48:11.986234000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1687848 Win=260820 Len=0
3869	2015-03-27 16:48:11.986241000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3870	2015-03-27 16:48:11.986250000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3871	2015-03-27 16:48:11.986255000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1690608 Win=260820 Len=0
3872	2015-03-27 16:48:11.986264000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3873	2015-03-27 16:48:11.986267000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3874	2015-03-27 16:48:11.986273000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1693368 Win=260820 Len=0
3875	2015-03-27 16:48:11.986279000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3876	2015-03-27 16:48:11.986282000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3877	2015-03-27 16:48:11.986286000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1696128 Win=260820 Len=0
3878	2015-03-27 16:48:11.986292000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3879	2015-03-27 16:48:11.986378000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3880	2015-03-27 16:48:11.986385000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1698888 Win=258060 Len=0
3881	2015-03-27 16:48:11.986392000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3882	2015-03-27 16:48:11.986395000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3883	2015-03-27 16:48:11.986399000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1701648 Win=255300 Len=0
3884	2015-03-27 16:48:11.986407000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3885	2015-03-27 16:48:11.986410000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3886	2015-03-27 16:48:11.986414000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1704408 Win=252540 Len=0
3887	2015-03-27 16:48:11.986420000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3888	2015-03-27 16:48:11.986424000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3889	2015-03-27 16:48:11.986427000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1707168 Win=249780 Len=0
3890	2015-03-27 16:48:11.986433000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3891	2015-03-27 16:48:11.986509000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3892	2015-03-27 16:48:11.986514000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1709928 Win=247020 Len=0
3893	2015-03-27 16:48:11.986519000	23.61.254.121	192.168.200.62	TCP	1434	[TCP segment of a reassembled PDU]
3894	2015-03-27 16:48:11.986526000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3895	2015-03-27 16:48:11.986528000	23.61.254.121	192.168.200.62	TCP	1434	[TCP segment of a reassembled PDU]
3896	2015-03-27 16:48:11.986539000	192.168.200.62	23.61.254.121	TCP	54	58051 > https [ACK] Seq=1146 Ack=57891 Win=66240 Len=0
3897	2015-03-27 16:48:11.986545000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3898	2015-03-27 16:48:11.986549000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1712688 Win=244260 Len=0
3899	2015-03-27 16:48:11.986554000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3900	2015-03-27 16:48:11.986557000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3901	2015-03-27 16:48:11.986561000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1715448 Win=241500 Len=0
3902	2015-03-27 16:48:11.986637000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3903	2015-03-27 16:48:11.986643000	23.61.254.121	192.168.200.62	TCP	1434	[TCP segment of a reassembled PDU]
3904	2015-03-27 16:48:11.986646000	31.13.90.6	192.168.200.62	TLSv1.2	1434	Continuation Data
3905	2015-03-27 16:48:11.986650000	192.168.200.62	31.13.90.6	TCP	54	58013 > https [ACK] Seq=4134 Ack=1718208 Win=238740 Len=0

Figure 4.16 : Page loading

4.5.5 Online Document (Google)

Google Doc application has been chosen to be the example of online document category in this study in order to have a different number of user actions. However, one user action was identified at network level as shown in Table 4.11. This is due to the complexity of identifying user action signatures when the signature continues to change during the experiment and this has become clear in the validation process of identifying user actions methodology. In fact, the three researchers found that user actions in Google application achieved different signatures in each researcher's list and the only user action that remains stable is edit document as show below.

User interactions	# packets	Total length (bytes)	Directions
Edit	Multiple	530+	Client > Server
Create	×	×	×
Share	×	×	×
Delete	×	×	×

Table 4.11 : Google Doc. application signature

1. Edit

There were two packets sent from client to server as illustrated in Figure 4.17. One of them in the beginning of interaction with size ranging between 530 up to 650 bytes and this is the signature of this user action on Google application that can be determined.

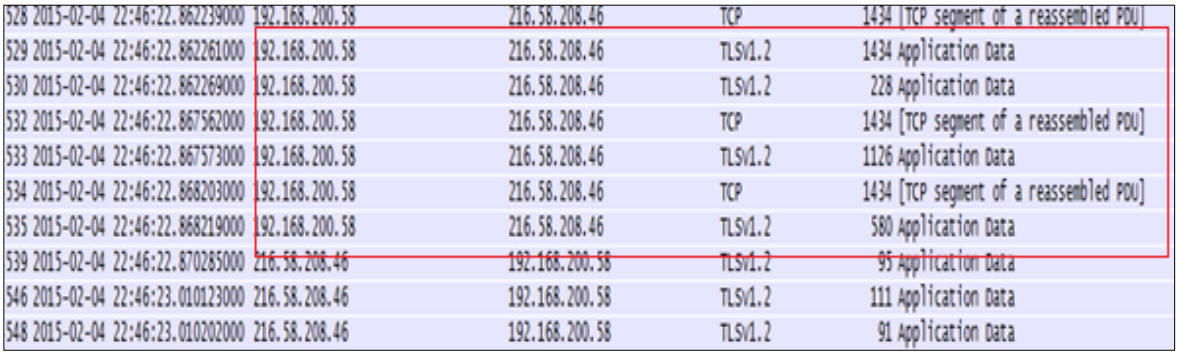

Figure 4.17 : Edit document

4.5.6 Video (YouTube)

There are two main user actions in this application which are watching video and uploading video. Thus, the investigation process has focused upon these two user actions. Table 4.12 shows that watching and uploading user actions have got a signature.

User interactions	# packets	Total length (bytes)	Directions
Watching Video	Stream	MTU (Almost)	Server > Client
Upload	Stream	MTU (Almost)	Client > Server

Table 4.12 : YouTube signatures

1. Watching Video (download)

When a user start chooses to watch a video, stream of packets is transferred t from YouTube server to the client as shown in Figure 4.18 with MTU size. These packets are intervened by Acknowledgement reply from the client side from time to time.

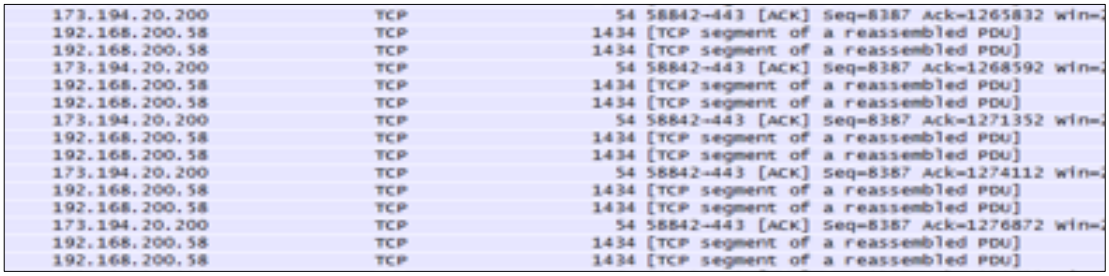

Figure 4.18 : Watching Video

2. Uploading Video

This option requires an account to be created on YouTube website in order to upload videos. After that, when the upload process starts a stream of packets are transferred from client side to the YouTube server by MTU size of each packet as demonstrated in Figure 4.19.

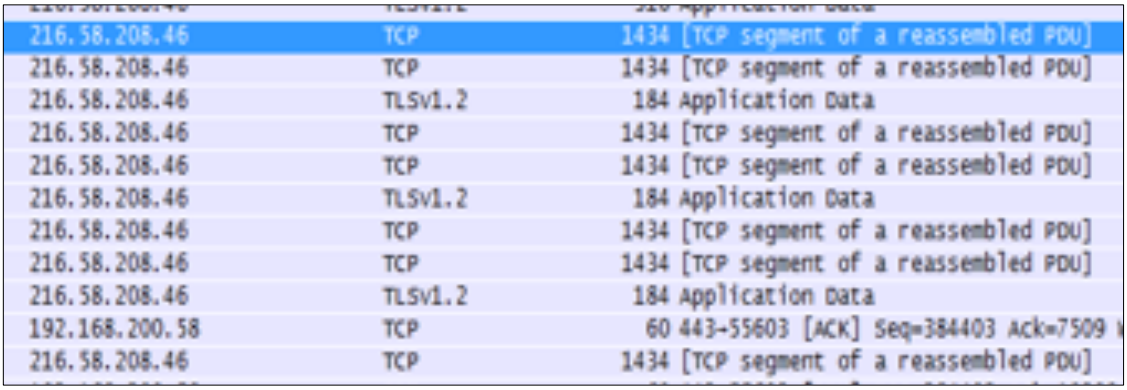

Figure 4.19 : Uploading Video

4.5.7 Online Storage (Dropbox)

The researcher has chosen Dropbox application to investigate cloud storage. This is because Dropbox services are becoming more popular. There are two main interactions investigated and their signatures explained below.

User interactions	No. of packets	Total length (bytes)	Directions
Download	Stream	MTU (Almost)	Server > Client
Upload	Stream	MTU (Almost)	Client > Server

Table 4.13 : Dropbox application signature

1. Upload

When a user uploads a document or file, the signature represents a group of packets with MTU size transferred from client to Dropbox server as shown in Figure 4.20. After each group that sent from client, there is a group of ACK packet sent from server to client. This process continues up to the document or file sent to server. The transmission is over TCP protocol and the data is sent and received through port 443 on the server side.

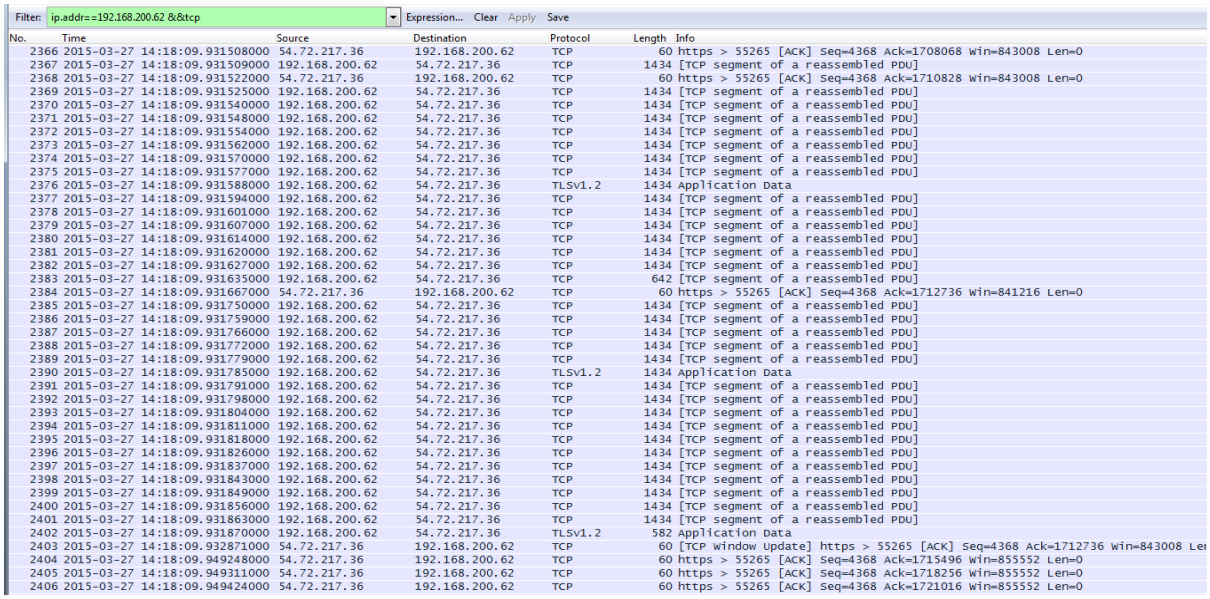

Figure 4.20 : Uploading document

2. Download

As can be seen in Figure 4.21, two to three packets are sent from server to client with MTU size followed by one ACK packet sent from client. This process continues up until all documents or files are completely delivered to the server. The transmission is over TCP protocol and the data is sent and received through port 443 on the server side.

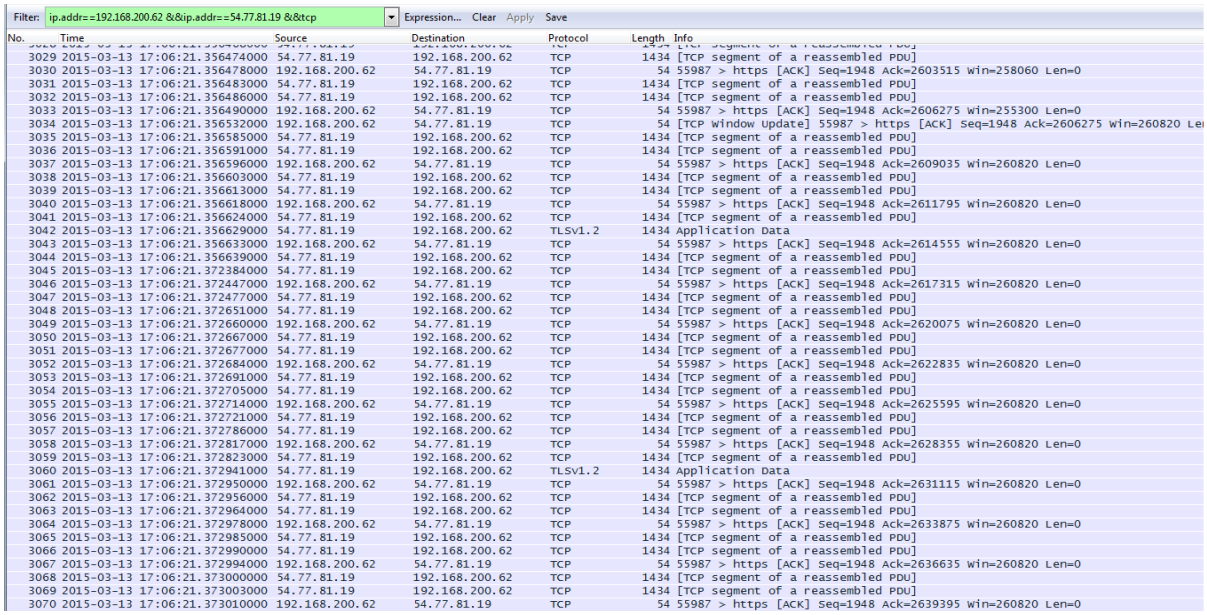

Figure 4.21 : Downloading document

4.5.8 Wikipedia

This application is the one that has been frequently visited as it is an open source for different subjects. So it is important to explore whether some of user actions can be captured here or not. The table below (see Table 4.15) shows that two user actions signature have been identified form this application; page viewing and download, in which they are going to be explained in the following paragraph.

User interactions	No. Packets	Total length (bytes)	Directions
Page viewing	One	Various	Client > Server
Download	Stream	MTU (Almost)	Server> Client

Table 4.14 : Wikipedia application signature

1. Page viewing

The data is moving from client to server through port 80. It is possible to identify each navigation action within websites by determining SYN and FIN, ACK using Wireshark as sown in Figure 4.22.

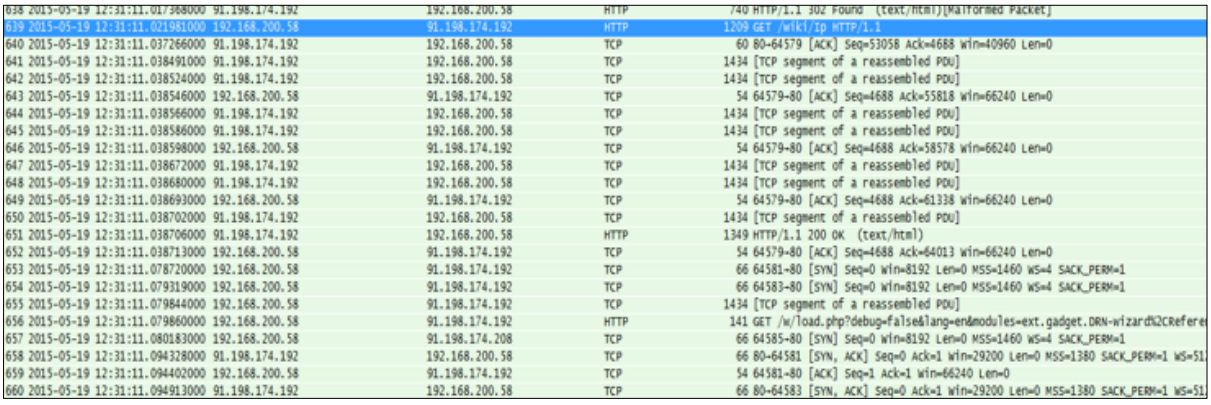

Figure 4.22 : Page viewing

2. Downloading

In the downloading user action, the data is moving from server to client in group of two to four packets with MTU size until the download is complete as can be seen in Figure 4.23.

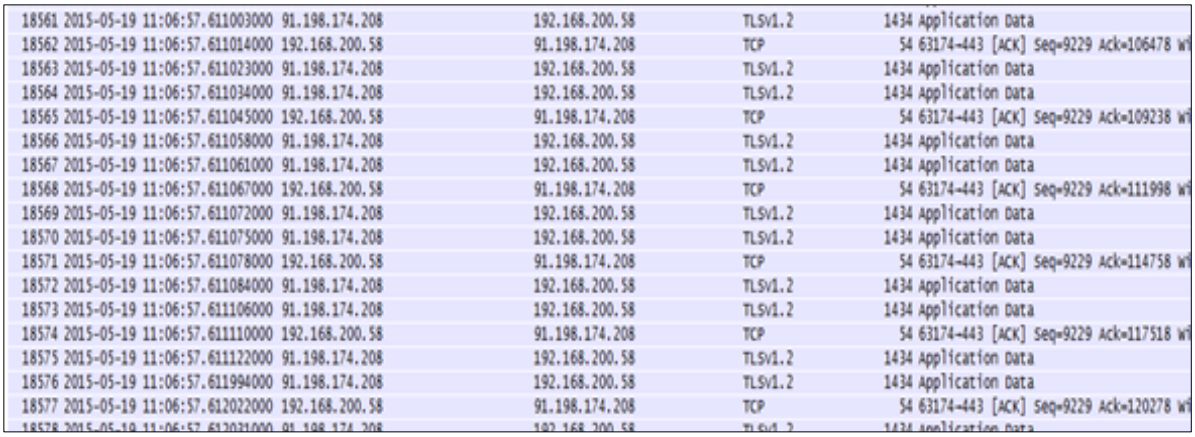

Figure 4.23 : Downloading file

4.5.9 Twitter

Twitter application has got a number of user actions that are commonly utilized by users due to the low level of complexity that has, which may encourage everyone to experience their action. However, table below (see Table 4.15) shows that three different user actions signature were successfully identified with different levels of accuracy as explained below.

User interactions	# packets	Total length (bytes)	Directions
Text message	One	192+	Client>Server
Upload	Stream	MTU	Client>Server
Click contact	×	×	×
Follow	×	×	×
Plying Video	×	×	×

Table 4.15 : Twitter application signature

1. Tweet

Doing tweet interaction several times gave a visible baseline. This baseline is valid with text tweet only. 182 bytes is an initial number. The experiment started with a tweet that is 10 char. long and found out that the packet size is 192 bytes as can be seen in Figure 4.27. Subsequently, number of characters was increased to 20 char. thus the packet size has increased to 202 bytes as illustrated in Figure 4.24. Therefore, 182 bytes considered as a bassline of tweeting user action.

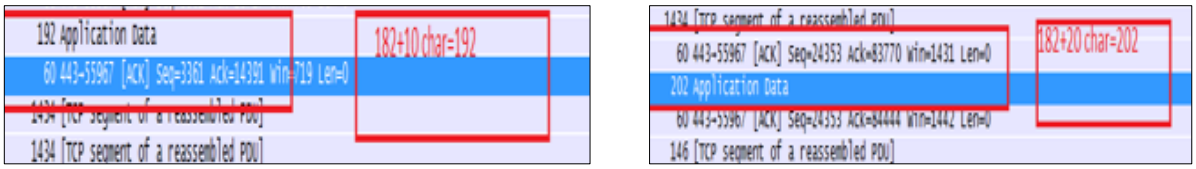

Figure 4.24 : Sending 10 and 20 characters

2. Uploading

This signature becomes obvious when a user starts uploading an image through tweet interaction. Then, a group of packets are transmitted from client to twitter server over TCP protocol. The size of these packets is dependent on the size of image, but it generally ranges from 1434 bytes. After the packets are received by the server, the server sends a group of acknowledgment packets to the client. The process continues until the image is delivered to Twitter server.

1611	2016-02-25 21:45:41.099713000	104.244.42.203	141.163.187.157	TCP	60 443→58868 [ACK] Seq=2472 Ack=962055 Win=1763 Len=0
1612	2016-02-25 21:45:41.099723000	141.163.187.157	104.244.42.203	TLSv1.2	151 Application Data
1613	2016-02-25 21:45:41.099748000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1614	2016-02-25 21:45:41.099758000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1615	2016-02-25 21:45:41.099765000	141.163.187.157	104.244.42.203	TLSv1.2	1434 Application Data
1616	2016-02-25 21:45:41.099771000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1617	2016-02-25 21:45:41.099778000	141.163.187.157	104.244.42.203	TLSv1.2	1434 Application Data
1618	2016-02-25 21:45:41.099784000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1619	2016-02-25 21:45:41.099790000	141.163.187.157	104.244.42.203	TLSv1.2	441 Application Data
1620	2016-02-25 21:45:41.107473000	104.244.42.203	141.163.187.157	TCP	60 443→58868 [ACK] Seq=2472 Ack=963532 Win=1763 Len=0
1621	2016-02-25 21:45:41.107598000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1622	2016-02-25 21:45:41.107613000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1623	2016-02-25 21:45:41.107621000	141.163.187.157	104.244.42.203	TLSv1.2	1434 Application Data
1624	2016-02-25 21:45:41.107629000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1625	2016-02-25 21:45:41.107636000	141.163.187.157	104.244.42.203	TLSv1.2	1434 Application Data
1626	2016-02-25 21:45:41.107642000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1627	2016-02-25 21:45:41.107649000	141.163.187.157	104.244.42.203	TLSv1.2	1434 Application Data
1628	2016-02-25 21:45:41.107655000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1629	2016-02-25 21:45:41.107662000	141.163.187.157	104.244.42.203	TLSv1.2	1434 Application Data
1630	2016-02-25 21:45:41.107669000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]
1631	2016-02-25 21:45:41.107676000	141.163.187.157	104.244.42.203	TLSv1.2	1434 Application Data
1632	2016-02-25 21:45:41.107683000	141.163.187.157	104.244.42.203	TCP	1434 [TCP segment of a reassembled PDU]

Figure 4.25 : Uploading image

In the same context, the examination of the 9 services is done similar to what have been explained before. Forty-two user actions across twelve services have been investigated, and the majority of them have a signature that belongs to the one of the categories mentioned above.

4.6 Discussion

At first glance, and by looking to the very large number of packets that pass through the network layer, capturing user actions to be impossible; however, applying the methodology of feature extraction process for network traffic metadata resulted in a simple concept of how user actions can be identified. Table 4.16 presents user actions signatures of all investigated

applications. The experiment has proved that identifying user action at application level is possible and unique as described in Table 4.16.

Services	User Action	Signature			Applications IP
		No. of packet	Size of packet	Direction	
Outlook	File attach.	Stream	MTU	Client>Server	131.253.61.%
	Compose email	One	971	Server>Client	204.79.197.%
	Insert recipient	One	971	Clientr>Server	204.79.197.%
Facebook	Share	Multiple	Various	Server>Client	31.13.%.%
	Page loading	Stream	MTU	Server>Client	
	Attach files	Stream	MTU	Client>Server	
	Chat	Multiple	2,625+	Client>Server	
	Typing	Multiple	1,502	Client>Server	
BBC	Page Navigation	Stream	Various	Server>Client	212.58.244.% 212.58.246.%
	Watching Video	Stream	MTU	Server>Client	
	Listening Audio	Stream	MTU	Server>Client	
Dropbox	Download files	Stream	MTU	Server>Client	108.160.172.
	Upload files	Stream	MTU	Client>Server	
	Page navigate	Multiple	Various	Server>Client	
Skype	Text messages	One	794+	Client>Server	157.56.193.% 157.56.126.%
	Audio calls	Stream	129-147	Both Clients	
	Video calls	Stream	1165-1365	Both Clients	
	File transfer	Stream	MTU	Both Clients	
	Click contact	One	747	Client>Server	
	Idle	One	587	Client>Server	
Google	Editing	Multiple	530+	Client>Server	216.58.208.% 173.194.78.% 216.58.209.%
	File download	Stream	MTU	Server>Client	
YouTube	Watch videos	Stream	MTU	Server>Client	216.58.198.%
	Video upload	Stream	MTU	Client>Server	
Twitter	Tweet	One	192+	Client>Server	185.45.5.% 199.96.57.% 104.244.42.%
	Upload	Stream	MTU	Client>Server	
Wikipedia	Download	Stream	MTU	Server>Client	91.198.174.%
	Viwing	Stream	MTU	Server>Client	

Table 4.16 : User actions signatures

Accordingly, 30 user actions across all applications out of 45 were successfully identified. Indeed, some user actions within the same application have a level of consistency in their signature, such as YouTube application where user action (upload/download) is represented by a stream of packets and MTU packet length but in different direction based on user action and others have not, for instance on twitter application each user action has different signature. In addition, user action signature uniqueness is measured by to what extent user actions signature is precise. With reference to Table 4.16, it is evident that user action signatures have fluctuated across the investigated applications. However, there are some applications that have achieved

a sufficient level of accuracy in their signature, such as Outlook, YouTube and Drobox. In fact, Skype application is considered as the most accurate application among them due to the high level of accuracy in its user action signatures in terms of number and total length of packets.

Although all applications that are included in this research have achieved an acceptable level of user action signature uniqueness such as Facebook, Twitter and Wikipedia, eBay application was an exceptional case where no single user action signature could be identified due to the noise of its traffic, in addition, the whole user actions in this application are encrypted since the website open. Ultimately, it is clear that there is a strong chance to determine the number of user actions signature in different applications, and this outcome may contribute towards an accurate user profile that would reveal what kind of actions being utilized within different applications.

4.7 Conclusion

Although the existing network traffic analysis based on anomaly detection is widely used in preventing/detecting insider threats, it suffers from different issues that are reducing their capability which subsequently causes serious concerns for organizations. Therefore, applications level user interactions proposed as a new approach required to overcome of the previous approaches limitations.

The user actions investigation shown showed that how complex identifying user actions signature at network level is. However, a set of experiments have been conducted on different applications that proved that a number of user interactions in different application can be identified.

5 User Behavioral Profiling from Network Traffic Metadata

5.1 Introduction

The previous chapter has shown that it is possible to identify user interactions with Internet-enabled applications from network traffic meta data. This chapter builds upon these to create feature vectors and develop and evaluate a network-based behavioural profile biometric. Behavioural profiling may have the potential to identify a user based on the user interactions signatures that predefined in the previous chapter. This leads to utilised a comprehensive methodology that is including some of the key functions such as classification and evaluation processes to measure the success of this approach.

5.1 Methodology

The experiment process consists of a number of steps which are demonstrated in Figure 5.1. that members the sequence for each function. The user's behavioural profiling flows from one function to another through the directional arrows. The methodology aims to measure to what extent it is feasible to use behavioural profiling based on user interactions. Data collection is the first step in focusing on the targeted applications highlighted before. Then selecting the features after applying the user action signatures to create a new dataset that includes user actions from different applications. However, this data needs to be normalized to be more beneficial to the classifier as demonstrates in the following sections.

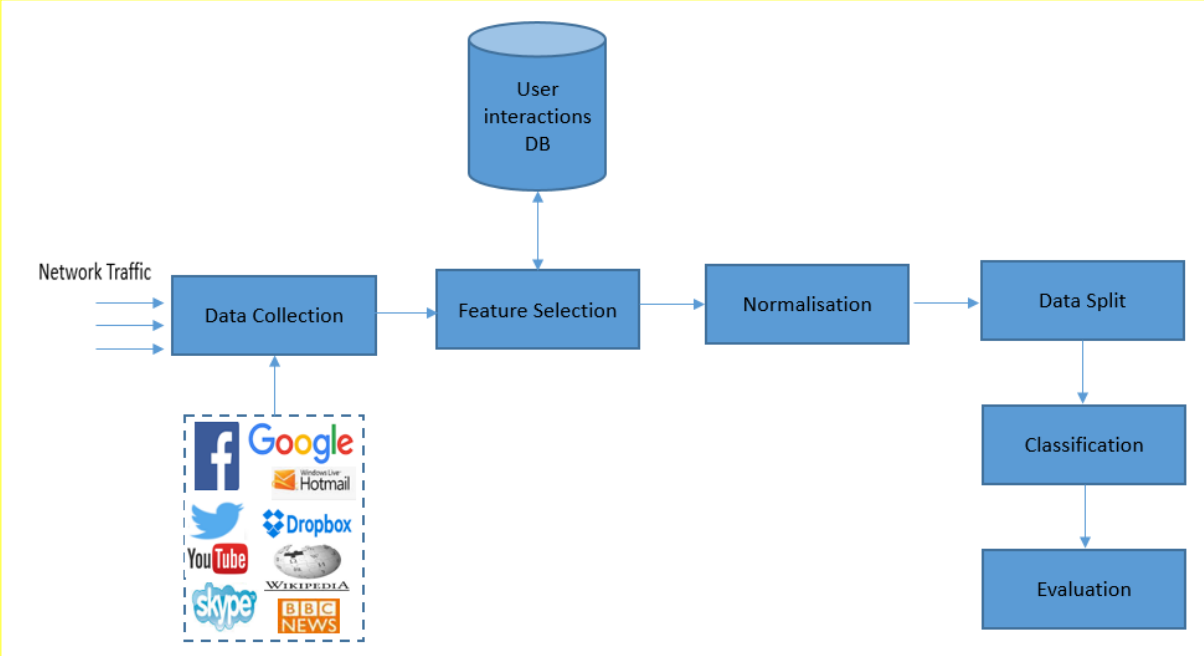

Figure 5.1: User behavioural profiling processes

5.1.1 Dataset

The effectiveness and robustness of an experiment such as this are largely dependent upon the quality and quantity of data. After a search for open-access data sets failed to identify any suitable sources, a significant data collection activity was undertaken. Network-based traffic data was collected centrally from the local research Centre (ethical approval appendix B). This enabled the researchers to set static IPs within the network in order to provide the ground truth and avoid changing IPs. Twenty-seven participants took part in the study that lasted for 2 months; from 7th November 2014 to 7th January 2015. This process focused purely on the collection of network metadata. The size of the complete data set attained at the end was 62.4 GB with over 140 million packets. The source IP address was anonymized in order to maintain privacy. Ethical approval was sought and granted (please refer to Appendix B for more information).

The raw data describes the natural manner of connection between users and all others services. This traffic includes a number of packets for establishing the connection with the server and

all applications that users have used. Furthermore, the database has been built to reveal the key information about each connection and this includes, time of the request, source and destination port, service IP addresses, size of packet and type of packet (as illustrated in Figure 5.2).

	Time	S_IP	S_port	D_IP	D_port	Length	Tags
	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1	2014.11.11.10:48:19.769086	user1	64584	224.0.0.252	hostmon	25	1
2	2014.11.11.10:48:19.817979	user1	49160	141.163.222.43	13000	0	S
3	2014.11.11.10:48:19.820678	141.163.222.43	13000	user1	49160	0	S.
4	2014.11.11.10:48:19.820866	user1	49160	141.163.222.43	13000	0	.
5	2014.11.11.10:48:19.820915	user1	49160	141.163.222.43	13000	0	F.
6	2014.11.11.10:48:19.822387	141.163.222.43	13000	user1	49160	0	.
7	2014.11.11.10:48:19.822492	141.163.222.43	13000	user1	49160	0	F.
8	2014.11.11.10:48:19.822567	user1	49160	141.163.222.43	13000	0	.
9	2014.11.11.10:48:19.823273	user1	49161	141.163.222.43	13000	0	S

Figure 5.2: Raw packet header information

There are some parameters that exist in the dataset as shown in the last column in Figure 5.2 and is explained in details in Table 5.1 so that they can be understood. These abbreviations represent the status of the packet in TCP protocol (TCP flags) and number 1 refers to the UDP traffic. As discussed in the previous chapter, every IP flow may contain separate interactions; hence it is important to realize when the connection is started. Synchronize abbreviation (S) represents the start of user action when the request is sent from the client to the server and the following packets represent a response from the server (Syn/Ack), end of action is described in the database as (F) and it is a required confirmation packet from the other side to show that the action is finished and this is represented as abbreviation (F.) that means 'Finish Acknowledgment'.

Column		Description
Time		Time when packet sending
S_IP		Source IP address
S_port		Source Port
D_IP		Destination IP address
D_port		Destination port
Length		Size of packet
Tags	1	UDP traffic
	S	Synchronize
	S.	Synchronize Acknowledge
	.	Acknowledgment
	F	Finish
	F.	Finish Acknowledgment
	PUH	Push

Table 5.1: Dataset description

The dataset has included all traffic that was captured between the participants' machines and any Internet services. Hence, some of preprocessing actions need to be done in order to orient the research to the main point that was discussed by focusing on the 10 Internet services that were chosen in the previous chapter.

5.1.1.1 Targeted Services Selection

The raw data included the metadata of the whole traffic that passed through the network gateway. This means that it is not limited to the nine Internet services that were predefined and investigated in the previous chapter. Subsequently, it had to be categorized based on the application, by doing a segregation process to enable every user to have a separate part that contains his traffic on each application and this process needs to be applied to 140 million packets that are representing the entire data set. Amongst this traffic, the experiment sought to focus upon 9 services that were previously analysed for user interaction signature in the previous chapter; after excluding eBay application due to the high level of noise during the signature identifying phase. Table 5.2 below shows that the selected applications, their network traffic and the number of participants in each application. It is clear that YouTube, Facebook and Google applications have the largest number of network traffic compared to other

applications and this is predictable due to their type of signatures as well as a large number of monthly visits across the world as mentioned in the previous chapter.

No.	Application	No. of packets	No. of Participants
1	YouTube	21,131,316	27
2	Facebook	5,727,953	27
3	Google	1,857,420	27
4	Twitter	747,584	27
5	Wikipedia	1,250,302	24
6	Outlook	703,711	24
7	Dropbox	17,480,739	18
8	BBC	201,263	17
9	Skype	575,030	12

Table 5.2: Total number of packets and participants per application

The traffic reveals that the majority of participants have utilised all applications. In fact, there are 4 Internet-based services across 9 services that have been utilized by all 27 participants, in which; two of them are into in domain social networks which are Facebook and Twitter while the other two are Google and YouTube services. Throughout the experiment, more than half of participants went through the whole 9 services. Skype has got the lowest usage in this experiment by 12 participants. In contrast, BBC and Dropbox applications have been utilized by more than 17 participants. The table above (see Table 5.2) demonstrates a number of participants in each application and it shows that Google, YouTube, Facebook, Twitter, Outlook and Wikipedia seem to be the core applications in this experiment and have enough popularity to have been used by the majority of participants. Unfortunately, as the purpose of the data collection methodology was to collect real users behaviour over a significant period of time, it was always likely that some applications would be not be used or would not be used enough to be able to be incorporated with the study. The overall figures, however, suggest there is sufficient volumes of interactions and users across a good number of applications to proceed.

Table 5.3 demonstrates the number of packets in each application at a user level. The amount of user traffic in each application relies on two factors. The first one is how frequently the user

has utilized an application; some applications have been used quite often during the day; such as Facebook and Twitter. The second factor is the nature of application traffic, for instance, Dropbox application traffic is quite massive and this is due to download and upload capabilities that it offers.

ID	BBC	Dropbox	FB	Google	Outlook	Skype	Twitter	Wiki.	YouTube	Total
1	545	0	2559	14479	65	0	1593	1658	66846	87745
2	9173	16029	123930	30390	834	91	90385	6084	496536	773452
3	937	0	69853	3301	1088	94	13913	7895	162538	259619
4	0	68303	540992	54971	36912	0	9643	481	6977320	7688622
5	0	0	5238	82272	276	0	3814	474	363268	455342
6	259	29895	6250	31108	2033	53026	1723	0	242199	366493
7	1176	62	24144	115720	236	0	20844	3253	535746	701181
8	0	0	669241	39609	90	0	1542	2241	175470	888193
9	2081	78294	591562	70626	101335	60283	95826	9544	2860391	3869942
10	0	0	45233	32867	13351	40742	12280	1277	165192	310942
11	505	0	37665	27085	19715	40694	5229	678	212851	344422
12	18734	1347	44985	295503	9935	36676	58610	5703	2572454	3043947
13	3045	45295	346115	113603	24	0	8679	1078	112193	630032
14	2440	251812	7751	439	2088	45430	14123	969	127571	452623
15	170	49672	652252	21130	339	0	46797	27236	504684	1302280
16	0	16474	2846	12706	1264	0	720	3034	10130	47174
17	51	0	803	12247	0	0	270	5444	11828	30643
18	7394	10859	21555	117756	204	0	6582	2128	13262	179740
19	18457	19219	13943	48101	65957	41497	4913	7567	195874	415528
20	0	887	255886	51059	283	0	13906	21992	1048721	1392734
21	756	20929	1538	19806	1593	41442	1823	421	95363	183671
22	0	0	2398	2839	0	0	232	0	13871	19340
23	0	0	314270	5724	115	0	1563	4753	80459	406884
24	2577	5727	7167	24045	118360	42839	3044	12661	93149	309569
25	0	1059	8976	15942	0	0	1329	1056	92418	120780
26	0	18	228082	11673	8	0	1499	1066	81918	324264
27	1157	15830	27477	332077	115908	42282	50081	3044	370566	958422

Table 5.3: Number of packet in each application per participant

There are some users who have acquired a small amount of network traffic across the services that have been utilized, such as Users 1 and 22 as illustrated in Table 5.3. These kinds of users reveal the daily usage of Internet services and type of applications that are preferred. For instance, both of them (user 1 and 22) did not utilize Dropbox and Skype applications, in contrast, Facebook, Twitter and YouTube applications were usually used based on the number of interactions they have.

Another example that is shown in Table 5.3, some participants have utilized all applications and the number of their network traffic is quite high. User 2, 9, 12 and 27 have a good number

of network traffic which roughly ranges between 700 thousand and 3.5 million in all applications. However, there are some differences among them, e.g. a number of packets in Outlook application for user2 are quite low while user 27 has got a high number of packets with over than 100 thousand. In contrast, this amount is not the same on BBC application where user 2 has a high number of packets compared to user 27. However, Twitter and YouTube applications have acquired the highest number of packets for all of them.

In terms of the applications usage, there are some participants who did not use some of the applications. With reference to the Table 5.3 above, user 22 has achieved the lowest point, in which it has utilized four applications out of nine; which are Facebook, Google, Twitter and YouTube. Nevertheless, there are also a good number of participants who did use the nine applications, such as users 9 and 27. The percentage of the number of users who used at least eight applications is quite promising, which is 60%. Hence, the applications usage boundary is between four; which represents 3%; and more than five applications; which represent 96%; for each individual and this is in accordance with what the previous chapter proposed about choosing the top applications' usage to be investigated in this experiment.

5.1.1.2 User Interaction Creation

The previous section shows the volume of network traffic for each participant on each application. However, this traffic represents the raw data the users have, regardless of whether it represents user actions or not. Subsequently, it is important to apply the user interactions signatures that were identified in Chapter 4 to investigate whether this data has a meaning or not. Hence, after applying the user interactions, the outcomes reveal that there is valuable data on the majority of applications. Table 5.4 illustrates the total number of interactions that exist after applying user interactions signatures upon each application.

No.	Application	No. Interactions	No. Participants	Data Reduction %
1	YouTube	1,322,848	27	93.8
2	Facebook	386,741	27	93.3
3	Google	194,404	27	89.6
4	Twitter	71,403	27	90.5
5	Wikipedia	5,719	20	99.5
6	Outlook	122,989	19	82.6
7	Dropbox	98,555	16	99.4
8	Skype	178,686	12	69
9	BBC	4,180	12	98

Table 5.4: Number of Interactions per application with rate of data reduction

With reference to Table 5.4 above, it is evident that there was a large reduction in data occurred after applying the user interaction signatures. Wikipedia and Dropbox have acquired the maximum amount of this by omitting about 99% of their network traffic. However, skype and Outlook are the two applications that have achieved a low data reduction with 69% and 82.6% and this is due to (as mentioned above) the nature of their interactions. Nevertheless, this scale of reduction will have a huge impact on the data processing requirements for the recognition system and over packet-based analysis systems.

According to the table above, there is sufficient interactions across a number of applications such as YouTube, Facebook, Google and Twitter that have data from the entire population of participants. In the same context, other applications have acquired a sufficient number of interactions in their traffic with at least 60% of participants, such as Wikipedia, Outlook and Dropbox.

Figure 5.3 demonstrates that the proportion of each application interaction from the total number of interactions during the experiment period. YouTube application cumulatively represented 55% of the total users' interactions due to the nature of user action signature in which it has been utilized by all participants.

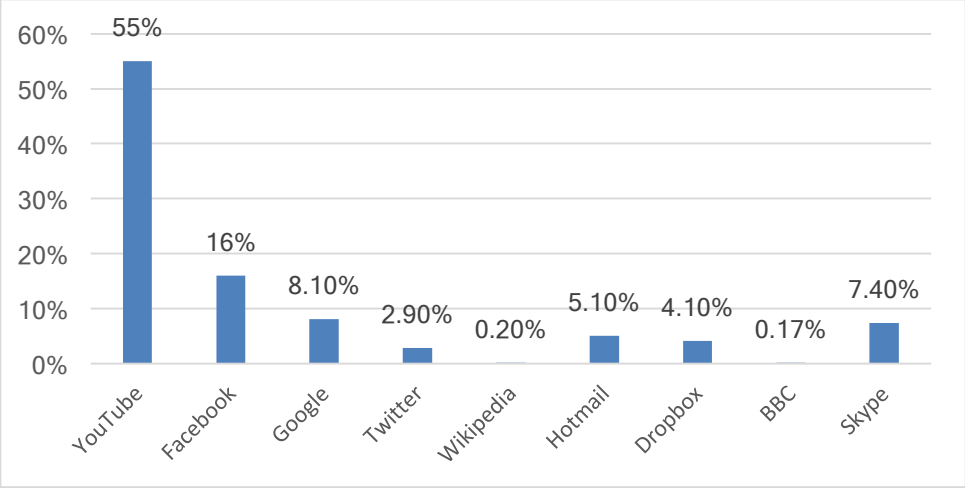

Figure 5.3: Overview of user interactions per application

In the same vein, Facebook and Google applications achieved the 2nd and 3rd applications in terms of user interactions since they have been utilized by all users. Although there are some applications that suffer from a lack of participants such as Outlook and Skype, they have acquired a good number of user interactions as can be seen in the Figure above (see Figure 5.3). This is because of their user actions represent the main function of them and they were quite accurate.

Table 5.5 illustrates the number of user interactions in each application for each individual. Users 17 and 22 achieved the minimum numbers of interactions in total with 2586 and 5314 interactions respectively across their usage of applications. While the maximum total numbers of user interactions that have been achieved by user 4 and 15 with 484,332 and 148,538 interactions respectively.

ID	BBC	Dropbox	FB	Google	Outlook	skype	Twitter	Wiki.	YouTube	Total
1	0	0	528	1898	0	0	266	40	7620	10352
2	654	6156	11068	3136	174	44	11494	232	50262	83220
3	28	0	3894	426	276	30	1504	164	22162	28484
4	0	21850	41252	15740	4404	0	1086	0	4.00E+05	484332
5	0	0	1110	17500	46	0	596	0	85356	104608
6	0	9196	386	2752	430	67528	218	0	28318	108828
7	108	0	4344	12590	36	0	3366	114	52550	73108
8	0	0	63104	3628	0	0	214	44	17404	84394
9	124	3164	68764	7248	10902	2922	9950	1384	2.00E+05	304458
10	0	0	2240	3904	3084	7698	1296	54	20050	38326
11	0	0	5350	4594	3162	1208	616	0	26284	41214
12	1630	540	7570	31616	2524	41010	7034	232	105336	197492
13	86	4976	16700	13454	0	0	1192	30	15050	51488
14	146	8900	1440	122	574	9058	1736	30	13346	35352
15	0	12416	43894	2200	46	0	9270	798	79914	148538
16	0	6134	662	860	62	0	106	228	1206	9258
17	0	0	150	992	0	0	38	168	1238	2586
18	278	4156	1672	16908	0	0	784	100	1622	25520
19	316	3978	3310	4892	25412	19282	636	214	23902	81942
20	0	170	22852	6626	32	0	1716	814	88798	121008
21	94	7654	242	3442	342	114	398	0	17390	29676
22	0	0	444	850	0	0	144	0	3876	5314
23	0	0	51642	1294	56	0	406	204	14712	68314
24	320	2204	1922	5006	58032	19986	768	674	16020	104932
25	0	358	1852	3962	0	0	322	0	11892	18386
26	0	0	24872	1320	0	0	322	42	12594	39150
27	28	6658	5460	27430	13330	9800	15910	68	40730	119414

Table 5.5: number of interactions per user

In terms of applications, there are 9 out of 27 users that have acquired interactions across the nine applications. This percentage has significantly increased to 50% of users who utilised at least 6 applications. This provides a strong proof that half of selected Internet services have been experienced by all users during the experiment. The volume of the interactions in each application is varied from one user to another due to the frequent use the applications and nature of the user action signature.

5.1.2 Feature Selection

Every record from the interaction table represents a feature such as a start and finish time, duration, source port number, destination port number, destination IP address, number and size of packets from the client to server and so forth. Selecting a subset of them should be done carefully in order to not miss some of them that have a value. According to Koller and Sahami (1996), "*The goal of feature selection is to select a small subset such that the resulting class*

distribution, given only the values for the selected features, is as close as possible to the original class distribution given all feature values"

Consequently, reducing the number of unrelated features and selecting the ones that are consistent with the general concept will drastically reduce the running time of learning algorithm (Kohavi et.al, 1995). Dash and Liu (1997) emphasized that there are two major steps in feature selection generation procedure and evaluation function. The generation step has three kinds, complete which means select all features, heuristic which means incremental (either decreasing or increasing) up to find the optimal feature and finally, random manner which means optimality the selection of a subset of features based on the resource availability and trying many times up to get the good result.

In this research, all three approaches have been examined and the result was subsequently different based on the utilized approach. Firstly, all 29 (start and finish time dividing to second on the top of the 9 features below) features were applied for all users to explore the outcomes of this approach, the result was quite poor and it consumed a lot of resources and time. The second attempt was carried out by selecting a random number of features, 4 features were selected (start time, duration, source port and service IP address) and again the result was not satisfying where equal error rate was quite high. Finally, the experimental approach was applied by selecting a number of features then adding and removing until the optimized result was reached. These features are the most appropriate because it represents the direction and value of the interaction. These features are as follows:

1. Start time.
2. Finish time.

3. Source port.

4. Service IP.

5. Destination port.

6. A number of the packet from the client to the server.

7. Size of the packet from the client to the server.

8. A number of a packet from the server to the client.

9. Size of the packet from the server to the client.

5.1.3 Normalization

Normalization is a manner used to standardize the range of independent variables or features of data. It is commonly performed throughout the data pre-processing step. Normalization can be executed at the level of the input features or at the level of the kernel (Priti &Vishakha, 2013). In many applications, the available features are continuous values, where each feature is measured in a different scale and has a different range of possible values. In such cases, it is often beneficial to scale all features to a common range by standardizing the data. Sola and Sevilla (1997) recommended prior to train the data, all input data must be normalized in order to obtain good results and accelerate significantly calculation. There are different techniques for normalization step and everyone has owned a particular manner to be applied (Patki and Kelkar, 2013). So, the technique is going to be used in this research is a linear normalization which means dividing every feature by the maximum value of the vector that belongs to.

5.1.4 Data Split

The split sample manner is a commonly used study in classification setting. There is a different approach for splitting data. The best splitting proportions strategies that are widely used ½ or ⅔

training (Dobbin.K and Simon.R, 2011). Accordingly, in this research ½ training and ½ testing data split strategy has been applied and before that, the data was randomly selected.

5.1.5 Classification

Using a machine learning technique to predict a group of membership for the data exemplar is the main theme of classification phase and a supervised learning approach was utilised. The classifier took the normalised data as inputs and applied its algorithm upon it (Sola and Sevilla, 1997). A neural network classifier was chosen as it has previously been proven to be an effective approach that deals with complicated patterns (Chtioui et al., 1997). One of these classifiers is feed-forward multi-layered perceptron (FF-MLP) and it considered as a large scale optimisation problem because it can be used for difficult to complex problems and it is the most popular neural networks that are working effectively and aiding solving a complex issue (Ilonen et al., 2003; Sharma et al.,2013; Metha et al.,2015). One layered neural network was built with nine inputs and one output by using Matlab tools version 2015b. MLP networks with a single-hidden layer and a nonlinear activation function are fundamentally able to distinguish classes with decision boundaries of arbitrary complexity, (Hartman, Keeler, Kowalski, 1990; Haykin, 2003). Within the FF-MLP network, a supervised learning technique called Backpropagation Marquardt was used for training the network and to solve the nonlinear problem. Neuron sizes were varied in order to seek to optimise (10,15, 20, 25 and 30 neurons). Table 5.6 defines the different neural network parameters with varying numbers of neurons utilised to allow for optimisation of the problem.

Classifier Name.	No. Neuron	No. Layer	Activation Function	Training Function	Epochs
FF-MLP	10-30	1	tansig	trainlm	1000 Iterations

Table 5.6 : Neural network setting

5.1.6 Evaluation

The evaluation process has been conducted in order to validate the level of correct classification. In the identification approach, the evaluation can be done by comparing the existing user templet user many temples to compare to what extent this template may match with one of the previous templates that already exist. However, in order to assist scalability issues involved with identification mode, a verification model has implemented with a single class classifier. Thus, the evaluation stage is undertaken using the standard approach for biometric testing which gives the system an ability to reject imposters and match the authorised users, which is initiated for all the participants (Clarke et al., 2004). Indeed, in each endeavour of identifying a particular user, the system has labelled the interactions that belong to this user by number 1 as a meaning of authorised and labelling the reset users interaction by value 0 as an indication of imposters. The experiment is repeated for each participant, in which every participant has to play the role of the authorised user and the remaining acts as the impostors. Thus, dividing each user interactions in each application into two halves; one for training the classifier and the other for testing; is vital because training and testing samples should be separated to ensure that there is no single sample that has been utilised in both parts. Also, in order to ensure there is sufficient data for both training and testing, a minimum of 28 user interactions is set as a threshold: users will be excluded from an application test if they had less than 28 interactions for that application. Ultimately, the evaluation phase of the classifier performance utilises the True Positive Identification Rate (TPIR) - if the biometric system outputs the identities of the top t matches for each user sample, where t represents the rank of accuracy. The following sections will show the result of the evaluation step in details.

5.2 Experimental Results

The results show that concentrating upon the user actions can produce a promising result in terms of user profiling. Table 5.8 illustrates the user's identification rate in the three different ranks and the average users' performance in the three ranks were gradually increased from 50.5% in ***rank 1*** to 71% in ***rank 5***. In ***rank 1***, more than half of the participants have acquired discriminative information which aided the system to create an accurate users profiling, thus, they achieved performance above average. In fact, among these users, there are some users who achieved high level of accuracy in ***rank 1*** such as users 9, 24, 27 where their performance are 68.8%, 79.5% and 88.5% respectively.

User ID	Total No. interactions	Ident. Rate rank 1 %	Ident. Rate rank 3 %	Ident. Rate rank 5 %
1	10352	49	55.3	63.9
2	83220	48.2	70.1	74.4
3	28484	46.3	64.8	74.4
4	484332	65.8	78.3	82.1
5	104608	32.3	38	51.7
6	108828	55.1	71.8	79.8
7	73108	36.9	69.1	80.2
8	84394	60.8	67.6	68.7
9	304458	68.6	75.4	82.6
10	38326	39.4	56.9	63.9
11	41214	51.2	55.1	57.9
12	197492	65.3	75	78.8
13	51488	54.1	63.3	69.5
14	35352	34.9	62.5	72.7
15	148538	59.8	80.1	84.9
16	9258	31.1	53.2	60.1
17	2586	28.4	39.1	43.6
18	25520	64.1	73.6	75.7
19	81942	45.5	60.1	71.2
20	121008	44.7	51.5	64.2
21	29676	50.6	71.8	87
22	5314	19.1	66.4	68.6
23	68314	41	54.4	61.7
24	104932	79.5	84.2	85.8
25	18386	53.7	61.2	68.3
26	39150	50.2	52.9	54.8
27	119414	88.5	91.6	92.8
Average		**50.5**	**64.5**	**71**

Table 5.7 : User identification rate

The overall average of user identification in ***rank 3*** and ***5*** have gradually increased by about 15% in ***rank 3*** and then from ***rank 3*** to ***rank 5*** at about 8%. Subsequently, there are some users that achieved identification more than 85% in ***rank 3,*** such as users 15, 24 and 27 in which their TPIR are 84.9%, 85.5% and 92.8% respectively. This suggests there are some very reliable feature sets amongst the users and applications.

With reference to the applications, there are a number of key results that can be derived from the experiment. Table 5.9 illustrates the TPIR for all users in each application. Applications that have achieved a high performance in ***rank 1*** considered as a high accurate application in terms of user profile because this reflects the level of discriminative information of users which indicate that choosing the classifier was a valid choice. Skype and Outlook applications have

achieved high TPIR from ***rankv1*** with 98.1% and 96.2% respectively which means the system has correctly classified almost all their samples by assigning them to the correct user. Although, BBC application has got achieved a good level of accuracy 81.8% in ***rank 1***, this value has gradually increased to 88.7%, in ***rank 2*** and improved to 95.4% in ***rank 5***. There are some applications that achieved a low TPIR, such as Google and Wikipedia where the system has correctly identified more than 66% of their samples from the first top value (***rank 1***) but this proportion could be considered as a promising value if we knew that they will have been utilized by the majority of the participants as show in Table 5.3 and the level of improvement that was achieved in their accuracy in ***ranks 2*** and ***3***. Although, YouTube application have acquired the lowest level of accuracy where its TPIR is 62.8% in ***rank1***, this proportion is still positive where YouTube interactions were quite large in number and represents more than 55% of whole users' applications usage as illustrated in Table 5.4 above.

App. Name	No. Users	Rank1 %	Rank2 %	Rank3 %	Rank4 %	Rank5 %
Skype	12	98.1	98.2	98.2	98.2	98.2
Outlook	19	96.2	96.8	96.9	96.9	97
BBC	12	81.8	88.7	92.5	93.7	95.4
Google	27	71.7	77	79.4	80.7	82.2
Wikipedia	20	66.9	78.2	83.6	86.6	89.2
Facebook	27	66.7	69.6	70.8	71.3	71.9
Twitter	27	65.3	75.3	79.5	81.9	83.4
YouTube	27	62.8	71.4	74.8	77	78.9
Dropbox	16	57.1	67.5	73.9	78.8	82.8

Table 5.8 : Application Identification Rate

Although, some applications results have acquired enough discriminative information to be correctly identified in ***ranks 1*** and ***2*** such as Skype and Outlook , the other applications have almost achieved significant improvement in their performance from ***rank 1*** to ***rank 5*** between 2% to 25%, for instance BBC, Wikipedia, Twitter, YouTube and Drobox. In addition, applications that have a small number of participants often achieved a high performance, and this is due the number of samples that needs to be compared with it such as Skype, BBC and Outlook.

For individual services, there are promising results that have achieved across all applications. The tables below show the level of TPIR in each application in different ranks. As Table 5.10, Skype application is considered as the top application that has a rich discriminative information based on TPIR result.

User ID	No. Interactions	Rank1 %	Rank2 %	Rank3 %	Rank4 %	Rank5 %
2	22	90.9	90.9	100	100	100
3	15	100	100	100	100	100
6	33764	100	100	100	100	100
9	1461	100	100	100	100	100
10	3849	63.7	63.7	63.7	63.7	63.7
11	604	80.3	83.9	85.2	91.3	94.8
12	20505	99.7	99.7	99.7	99.7	99.7
14	4529	99.9	99.9	99.9	99.9	99.9
19	9641	99.4	99.5	99.5	99.6	99.6
21	57	100	100	100	100	100
24	9993	99.9	99.9	99.9	99.9	99.9
27	4900	100	100	100	100	100

Table 5.9 : Skype Identification Rate

There is 75% of users that have successfully been profiled with almost 100% TPIR from ***rank 1*** such as users 3, 5, 24 and 27. Subsequently, their interactions size varies but the performance remains high which means the level of uniqueness of information that is located on their traffic are precise enough to distinguish the individual, such as users 15 and 21 who have a smaller size of interactions compared to the other users. However, users 2, 10 and 11 have achieved the lowest TPIRs with 90.9%, 63.7% and 80.3% respectively, even though this outcome is not satisfying, their performance has slightly improved by almost 10% in ***rank 5***. Skype application has achieved the best user profile result among the 9 application with almost over 90% TPIR and this is due to the nature and precise of its signature also the port connection is often constant with particular number above of 1024 (Skype, 2016). Ultimately, this feature helps the application to register the lowest application that had data reduction with 69% of original dataset.

In the same view, Outlook application achieved high level of accuracy from ***rank 1*** with over 96% TPIR and this is due to the level of uniqueness of information that exists in the traffic.

Regardless, Outlook traffic represents more than 5% of the entire users traffic, it been utilized from almost 3/4 of participants. In addition, it has been registered as the second lower application in terms of data reduction of 82.5% by applying the proposed user interaction approach on their traffic as illustrated in Table 5.4 and this is due to the nature of user action signature that exists on the traffic, such as attach file and send email as explained in the previous chapter.

User ID	No. Interactions	Rank1 %	Rank2 %	Rank3 %	Rank4 %	Rank5 %
2	87	5.7	597	72.4	72.4	75.8
3	138	60.8	60.8	60.8	60.8	60.8
4	2202	91.8	97.9	97.9	97.9	98.3
5	23	13	13	13	13	13
6	215	85.5	92	92	92	92
7	18	0	33.3	55.5	100	100
9	5451	94.9	95.4	95.4	95.4	95.4
10	1542	83	85.4	86	86	86
11	1581	80.5	80.5	80.6	80.7	80.7
12	1262	80.2	82	82	82.1	82.4
14	287	62.3	64.8	66.2	66.2	67.6
15	23	0	26	73.9	100	100
16	31	35.4	54.8	58	61.2	61.2
19	12706	95	95.5	95.6	95.7	96
20	16	0	0	0	25	56.2
21	171	85.3	85.3	85.3	85.3	85.3
23	28	71.4	71.4	71.4	71.4	71.4
24	29016	99.9	99.9	99.9	99.9	99.9
27	6665	99.8	99.8	99.8	99.9	99.9

Table 5.10 : Outlook True Identification rate

With reference to the table above, the system can successfully be profiling more than ½ of participants from ***rank 1*** with a level of TPIR above 80%. In the meantime, 50% users in this group can be correctly identified by almost their whole traffic from ***rank 1,*** for instance, users 24 and 27. Moreover, in the majority of them, their performance has been improved by 5% from ***rank 1*** to ***rank 5***. Although, some of users have achieved low performance such as 7 and 16 who achieved 0% and 35.4% TPIR in ***rank 1***, this proportion is improved in different ranks such as in ***rank 3*** where more than 50% of user 7 traffic can be identified and in ***rank 5*** as well where the system can correctly assign all user 7 traffic. It seems to be the number of interactions are altered from one user to another and it applies the main pillar in this approach but the type of

signature also has a vital influence, for instance some of participants have acquired a small number of interactions but they achieved a high TPIR, such as users 21 and 23. Thus, there is a strong relation between number of interactions, type of interaction signature and the TPIR. So, the performance of the system relied on variety parameters such as number of user interactions signature, and their precisely signature from each application that can be determined before.

Despite the fact that BBC application has the smallest amount of user's interactions size, it achieved as a third application in terms of performance with average TPIR above 80% (by calculating the results in rank 1 and divided by total number of participants) in ***rank 1*** as shown in Table 5.11. Therefore, from ***rank 1*** the system is able to identify 25% of the participants' traffic with TPIR more than 80%, such as user 12 and 24 where the TPIR above 90% for both as demonstrates in Table 5.11. They have a satisfying amount of interactions size and uniqueness information that contribute to the system in order to distinguish a particular user from another user from ***rank 1***.

User ID	No. Interactions	Rank1 %	Rank2 %	Rank3 %	Rank4 %	Rank5 %
2	327	74.6	83.1	87.1	91.4	92.9
3	14	0	42.8	78.5	100	100
7	54	64.8	64.8	64.8	64.8	64.8
9	62	59.6	61.2	61.2	62.9	75.8
12	815	95	95.4	96.4	97	97.5
13	43	60.4	72	74.4	76.7	76.7
14	73	34.2	58.9	76.7	80.8	84.9
18	139	82	94.2	98.5	100	100
19	316	61.7	88.2	98.1	98.4	99.3
21	47	74.4	74.4	76.6	93.6	100
24	160	91.8	91.8	92.5	94.3	94.3
27	14	71.4	78.5	92.8	100	100

Table 5.11 : BBC True Identification rate

In comparison, there are some users who achieved low TPIR in ***rank 1*** and do not have that much amount of interactions size but their performance has dramatically increased up to correctly assign all entire traffic to the right user, such as users 3 and 27. With reference to Table

5.11 True identification in ***rank 2*** has achieved a good performance rate where more than ½ of users' traffic can be successfully identified with more than 75%.

Google is one of the four applications that have been utilized by all participants as demonstrated in Table 5.8. It has two main signatures as explained in the previous chapter; search and modify document. These two user actions gave it this kind of popularity to be used by the users. In addition, by concentrating upon these users action, data reduction reaches up 89.6% of the whole application traffic.

The results in Table 5.12 show that there are a number of promising results from ***rank 1***. Amongst 27 users, there are 11 users who have acquired a uniqueness of information in their traffic which contributes to the system towards right classification with level of accuracy reached to 70% and more in ***rank 1***. In addition, this percentage is slightly improved ranging between 2% to 8% in each rank such as users1 and 27 with 95.2% and 99.9% respectively.

User ID	No. Interactions	Rank1 %	Rank2 %	Rank3 %	Rank4 %	Rank5 %
1	949	95.2	96.1	96.2	96.2	96.2
2	1568	32.4	52.8	80.1	84.3	87.1
3	213	59.1	69	95.7	100	100
4	7870	91.6	91.8	91.9	92.1	92.2
5	8750	73.6	78.9	79.3	79.6	79.8
6	1376	89.1	90	90.1	90.4	90.6
7	6295	79.2	86	87.4	87.8	88.1
8	1814	42.7	54.5	56.1	58.2	59.1
9	3624	0	0	0	0	13.6
10	1952	56.1	67.4	69.6	72.4	74.2
11	2297	72.7	75.3	76.5	76.9	77.1
12	15808	73	78.2	79.6	80.6	82
13	6727	75.3	84.8	86.2	87.9	88.4
14	61	26.2	70.4	90.1	90.1	100
15	1100	68.4	78.1	82	83.4	83.7
16	430	16.2	43.7	70.2	76.5	81.6
17	496	59	73.5	80.4	88.7	93.5
18	8454	68.3	75.2	77.4	79.8	82.2
19	2446	42.3	51.8	69.8	73.9	75.8
20	3313	63.2	63.9	64	64.1	64.4
21	1721	18.9	52.6	74.1	85.5	93.6
22	425	29.4	48	48.2	48.2	48.9
23	647	0	0	0	0.9	0.9
24	2503	76.5	79.9	81.8	83.7	86.1
25	1981	80.8	81.7	83.3	88.4	91.1
26	660	64.8	65	65	65	65
27	13715	99.9	99.9	99.9	100	100

Table 5.12 : Google True Identification rate

The remaining users who represent about ½ of participants have acquired a fluctuated TPIR as illustrate in Table 5.12 in ***rank 1***. However, this proportion has increased dramatically among each ***rank*** up to 20% or more, for instance user 21 has got TPIR 18.9% in ***rank 1*** and subsequently when the performance increased from ***rank 1*** to ***rank 2*** with more than 30% to score 52.6 % and accordingly from ***rank 2*** to ***rank 3*** with 20%. Ultimately, this application has achieved a satisfying level of discriminative information that aids the system to assign each interaction to the associated user and if it is not from ***rank 1*** it would be possibly be in ***ranks 2*** or ***3*** as the result described above.

In spite of the fact that the data reduction process reduced a vast amount of the data with Wikipedia being the highest application that was exposed (99.5% of the original traffic was

reduced) the remaining data has enough discriminative information which leads to a high level of TPIR of some users as shown in Table 5.13.

User ID	No. Interactions	Rank1%	Rank2%	Rank3%	Rank4%	Rank5%
1	20	100	100	100	100	100
2	116	74.1	78.4	78.4	78.4	78.4
3	82	6	6	31.7	56.1	79.2
7	57	49.1	61.4	61.4	61.4	61.4
8	22	100	100	100	100	100
9	692	93	94.2	94.8	94.9	95.2
10	27	0	3.7	11.1	29.6	29.6
12	116	56.8	65.5	67.2	71.5	75.8
13	15	0	0	20	46.6	53.3
14	15	0	0	0	0	0
15	399	64.6	91.9	99.2	99.5	100
16	114	95.6	95.6	95.6	95.6	95.6
17	84	52.3	67.8	70.2	72.6	73.8
18	50	98	98	98	98	98
19	107	0	0.9	21.5	48.6	72.9
20	407	57.7	85.5	88.9	93.1	94.6
23	102	20.5	38.2	68.6	82.3	91.1
24	337	81.6	91.9	93.7	93.7	93.7
26	21	76.1	76.1	80.9	80.9	85.7
27	34	73.5	73.5	73.5	73.5	73.5

Table 5.13 : Wikipedia Identification Rate

Therefore, 40% of the participants have successfully classified their traffic with more than 73% of accuracy. In fact, the level of uniqueness in their traffic reaches to the extent where the whole user interactions assigned correctly to the associated user from ***rank 1***, such as users 1 and 8. However, there are some users that have achieved low TPIR in ***rank 1*** and ***2*** , however, they achieved a good improvement up to correctly classified almost 70% of their traffic in ***rank 5,*** for instance user 3 has achieved 6% TPIR in ***rank 1*** and this proportion continues to increase in each ***rank*** to attain almost 80% of the traffic has accurately classified.

Social media Internet services are becoming common applications across the world, such as Facebook, Twitter and so forth Wilson et.al (2012). With reference to the findings in the previous chapter, different user actions have been discovered on Facebook application with their signature. This application has achieved 66% TPIR in ***rank 1*** as an average value of all entire users' result as demonstrated in Table 5.8 above. However, 1/3 of participants has

acquired a satisfying level of uniqueness of information on their traffic and achieved TPIR of 63% and above as illustrate in Table 5.14.

User ID	No. Interactions	Rank1%	Rank2%	Rank3%	Rank4%	Rank5%
1	264	0	0	25	39.3	46.2
2	5534	30.9	38	39.6	40.3	41.8
3	1947	45	45.5	45.5	46.4	46.4
4	20626	17.3	17.4	17.4	17.4	17.5
5	555	0.5	5	5	8	24.1
6	193	0.5	2.5	25.9	42.4	55.9
7	2172	19.7	47.7	78.5	84	89.5
8	31552	79.5	81.4	81.5	81.6	81.6
9	34382	92.5	97.1	97.3	97.5	97.7
10	1120	11.8	35	73.2	79.8	91.3
11	2675	48.6	48.8	48.8	48.8	48.8
12	3785	57	57.6	59.7	63.6	67
13	8350	72.7	74.8	74.8	74.8	74.8
14	720	4	30	56.9	73.7	90.9
15	21947	74.5	77.3	77.3	77.3	77.4
16	331	22.3	38	54.9	56.4	57.1
17	75	0	0	0	0	1.3
18	836	30.3	34.8	34.8	34.8	34.8
19	1655	20.6	27.9	39.8	57.5	67.3
20	11426	73.9	75.5	76	76.1	76.2
21	121	3.3	31.4	32.2	32.2	38
22	222	0	63.9	78.3	78.3	78.3
23	25821	75.5	75.6	75.6	75.6	75.6
24	961	73.3	77.2	80.5	82.7	83.9
25	926	45.8	57.6	60.5	64.7	65.5
26	12436	62.4	65.4	66.2	66.4	66.6
27	2730	72.6	72.7	72.7	72.7	72.7

Table 5.14 : Facebook Identification rate

In addition, this proportion has gradually increased in each ***rank*** between 6-10% and in ***rank 5;*** the number of users who achieved TPIR above 65% represents almost 2/3 of existing participants.

With regards to individuals, there are some participants who achieved a good TPIR in ***rank 1***, such as users 9 and 8 where the system was able to correctly identify 92.5% and 79.5% of them as can be seen in Table 5.14.

With reference to the Table above, it seems that users who have acquired a good number of interactions achieved a good TPIR and this is because of this amount of interaction gives the system an ability to find a user pattern that differentiate one user from another. For instance,

users 8 and 9 have achieved a maximum number of interactions in which they achieved the highest TPIR.

In the same vein, and as mentioned above Twitter application has been utilized by all participants throughout the experimental and almost 90% of its original traffic has been reduced after applying the user interaction approach on the traffic as illustrated in Table 5.5. Even though, the overall result gives an indication that the application owns a discriminative information that contributes to the system in order to be able to correctly classify 66.4% from all entire traffic in ***rank 1*** and this proportion has gradually increased to reach up to 83.4% in ***rank 5***.

User ID	No. Interactions	Rank1%	Rank2%	Rank3%	Rank4%	Rank5%
1	133	0	0	1.5	7.5	21
2	5747	60.4	79.9	85.5	88	89.5
3	752	55	56.1	56.2	56.2	56.2
4	543	20.4	49.7	69.6	80.8	87.6
5	298	0.3	10.4	17.1	48.9	66.1
6	109	33	47.7	50.4	51.3	52.2
7	1683	17.1	40.2	53.9	58.1	63.3
8	107	25.2	33.6	40.1	42	42
9	4975	79.5	89.9	94	95.8	96.2
10	648	16	43.3	45.9	47.9	52
11	308	25.3	37.9	39.6	46.1	46.1
12	3517	67	69.3	70	71.2	71.8
13	596	49.1	59.7	62	64.4	65.7
14	868	6.4	32.4	59.6	75.1	83.5
15	4635	70.2	78.8	82.1	84.4	85.3
16	53	0	0	0	1.8	3.7
17	19	0	0	0	0	0
18	392	46.4	66.5	71.1	73.2	74.4
19	318	25.1	29.5	30.8	31.7	34.2
20	858	42.6	53.3	56	58.3	58.3
21	199	79.3	83.9	85.9	87.9	89.4
22	72	43	68	68	68	68
23	203	58.6	66	66.5	66.5	66.5
24	384	44.7	49.7	52.6	53.1	54.1
25	161	14.9	24.8	29.8	37.2	46.5
26	161	0	0	0	0	0
27	7955	97.3	98.1	98.9	99.4	99.7

Table 5.15 : Twitter Identification Rate

In terms of individuals, there is a number of participants who achieved a good TPIR, such as users 27, 9, and 15 where their results are 97.3%, 79.3% and 70.2% respectively as demonstrate in Table 5.15. In fact, these results show that Twitter application has owned enough user action that reveals some uniqueness on their traffic.

It is clear that 1/3 of users has achieved TPIR of 50% and above in ***rank 1***, but this proportion has dramatically risen up to 3/4 of participant has achieved 50% TPIR. To this extent, the majority of them can be correctly identified in ***rank 5*** by more than 70% as can be seen in Table 5.15. However, there are two users; user17 and 26; out of 27 users where the system could not find any discriminative information on their traffic and achieved 0% TPIR across all ranks, although there are some users achieved low correct classification in ***rank 1*** or ***2***, they could achieve a level of accuracy in the remaining ranks, such as users 1 and 5.

In contrast, YouTube application has been exposed to the massive data reduction that omits more than 93% of its traffic after user interaction was applied on the traffic as demonstrated in Table 5.4, it represents more than 50% of the whole users traffic in the new dataset that includes just the valuable information about the users (user interactions). Based on the predefined signature of YouTube application, there are two main functions that are frequently used; upload and download. Hence, it consumes a stream of packets in both tasks, and this is described in the number of interactions column in Table 5.16.

User ID	No. Interactions	Rank1%	Rank2%	Rank3%	Rank4%	Rank5%
1	3810	50	53.5	54.4	55.6	56.7
2	25131	31.2	36.2	37.5	38.6	40.6
3	11081	44.5	48.8	49.9	51.2	52.9
4	2E+05	90.9	97.2	98.3	98.7	99.1
5	42678	74.1	74.8	75	75.3	75.7
6	14159	34.8	56	71.8	78.5	82.4
7	26275	28.8	63.7	81.9	91.2	94.7
8	8702	56.5	59.3	60.1	60.4	60.9
9	1E+05	83.6	90.2	91.6	92.8	93.7
10	10025	45	47.9	48.8	49.8	50.8
11	13142	0	0	0	0	0.1
12	52668	46.3	77	90.2	93.7	95.3
13	7525	40.1	42.6	44.4	45.4	46.3
14	6673	33.1	38.4	42.4	43.4	44
15	39957	71.1	74.7	75.4	76	76.8
16	603	28	28	28.3	29.5	29.8
17	619	30.6	42.4	45	48.3	49.5
18	811	67.2	67.8	68.4	68.6	68.9
19	11951	21.9	31.7	37.1	40.4	44
20	44399	0	0	0.6	4.4	13.9
21	8695	17	32.4	52.6	82.7	96.8
22	1938	4.2	65.6	71.2	76.2	79.2
23	7356	20.2	32.1	44.7	54.7	64.6
24	8010	86	91	93.5	95.3	96.2
25	5946	51.1	53.8	56.3	58.3	62.5
26	6297	47.5	51.3	52.4	54.8	56.9
27	20365	99.9	100	100	100	100

Table 5.16 : YouTube Identification Rate

However, the TPIR is varied from one user to another, and this is due to the level of valuable information that distinguishes a particular user are different. There are more than 37% of participants who achieved TPIR more than 53% and when it was compared with the number of packets that were investigated it represents a big number, for example user 12 has acquired more than 52,000 interactions and the system was able to identify almost ½ of them. In fact, some of them have acquired a large number of interactions and almost 100% have been correctly classified in ***rank 1***, such as users 4 and 27 where the TPIR achieved 90.9% and 99.9%.

In similarity with YouTube application, user interaction of Dropbox application signifies less than 1% of the original application traffic, but it has been utilized by around half of participants as shown in Table 5.17. In addition, the majority of user interaction size is more than a thousand and this is due to the nature of user action signature and how often user used a particular

application. It is clear that there is a level of discriminative information for all participants with different levels of accuracy in ***rank1*** but this amount improved has gradually to score around 40% TPIR as a minimum value in ***rank 5***.

User ID	No. Interactions	Rank1%	Rank2%	Rank3%	Rank4%	Rank5 %
2	3078	34.1	44	50	56.8	63.8
4	10925	82.5	90.7	94.5	96.1	97.6
6	4598	43.1	65.2	72.7	78.8	85.4
9	1582	14.6	30.9	44.6	59.6	75.9
12	270	12.9	24	30.3	35.1	38.1
13	2488	80.9	81	81.3	81.3	81.3
14	4450	48.2	63.5	71.1	77.8	83.8
15	6208	70	71.1	71.2	71.2	71.3
16	3067	20.6	41	65.7	81.1	92
18	2078	56.4	64.3	66.8	69.6	71.5
19	1989	43.7	49.2	49.4	49.7	51.3
20	85	75.2	75.2	75.2	85.8	85.8
21	3827	26.9	48.1	68.1	84.2	92.7
24	1102	61.8	63	63.7	63.8	64.2
25	179	75.9	75.9	75.9	75.9	75.9
27	3329	82.3	85.7	87.2	88.1	89.2

Table 5.17 : Dropbox Identification Rate

In terms of individual, there are a promising results that were achieved from some users, such as users 4, 13 and 27 where the number of interaction exceeded 2000 and the system could classify more than 80% of their traffic to the right user as can be seen in Table 5.17.

5.3 Discussion

Although the system has achieved promising results, some users in different applications did not have sufficient uniqueness of information on their traffic to contribute to the system to find out a pattern in terms of user behavioral in particular applications in order to profile them, hence the TPIR was quite low. However, all participants have experienced the use of at least three applications as shown in Table 5.18. These three applications represent the top three applications that achieved a high level of accuracy from rank 1 for each user.

User ID	First App.		Second App.		Third App.	
	Name	TPIR%	Name	TPIR%	Name	TPIR%
1	Wikipedia	100	Google	95.2	YouTube	50
2	Skype	94.1	BBC	74.6	Wikipedia	74.1
3	Skype	100	Outlook	60.8	Google	59.1
4	Outlook	91.8	Google	91.6	YouTube	90.9
5	YouTube	74.1	Google	73.6	Outlook	13
6	Skype	100	Google	89.1	Outlook	85.5
7	Google	79.2	BBC	64.8	Wikipedia	50
8	Wikipedia	100	Facebook	79.5	YouTube	56.5
9	skype	100	Outlook	95	Wikipedia	93
10	Outlook	83	Skype	63.7	Google	56.1
11	Outlook	80.5	Skype	80.3	Google	72.7
12	Skype	99.7	BBC	95	Outlook	80.2
13	Dropbox	80.9	Google	75.3	Facebook	72.7
14	Skype	100	Outlook	62.3	Dropbox	48.2
15	Facebook	74.5	YouTube	71.1	Dropbox	70.9
16	Wikipedia	95.6	Outlook	35.4	YouTube	28
17	Google	59	Wikipedia	52	YouTube	30.6
18	Wikipedia	98	BBC	82	YouTube	67.2
19	Skype	99.4	Outlook	95	BBC	61.7
20	Dropbox	75.2	Facebook	73.9	Google	63.7
21	Skype	100	Outlook	85.3	Twitter	79.3
22	Twitter	43	Google	29.4	YouTube	4.2
23	Facebook	75.5	Outlook	71.4	Twitter	58.6
24	Outlook	100	Skype	100	BBC	91.8
25	Google	80.8	Dropbox	75.9	YouTube	51.1
26	Wikipedia	76.1	Google	64.8	Facebook	62.4
27	Skype	100	Google	100	YouTube	100
Average		87.4		68.9		61.9

Table 5.18 : Users TPIR in Rank1 Top Three Applications

The results reveal that 1/3 of participants have been correctly identified via their interactions with a level of accuracy of more than 80% in all top three applications. It is also shown that more than 75% of the users have acquired at least one application with 80% TPIR, where the system has an ability to profile 92% of the users from their interactions with TPIR of more than 74%, there is a small proportion where the system could only assign less than 60% of the interactions to the associated participant.

On the level of individual user, the average TPIR in first application for all users on average has achieved 87.4%, this result has proved that it is possible to use user interaction for creating user profiles which is the main aim of this work. Therefore, there are some participants who achieved high level of TPIR across the three applications from ***rank 1*** as shown in Table 5.18, such as users 4, 9, 24 and 27, but user 27 has acquired the best user profile because of the highest level

of accuracy it has acquired in all top three applications where the system was able to completely identified all its interactions with 100% accuracy. Although the majority of participants have achieved a promising result in the first application, Skype application has achieved the best user profile result among the 9 applications with almost over 90% TPIR and this is due to the nature and precision of it signature.

The experiment reveals that the nature of the user interaction derived from application level is unique; thereby using it to build a user behavioral profile is a promising solution to identify a user. Moreover, the experiment evidently shows that by using user interactions the system was able to identify some participants in different applications with 100% level of accuracy (as shown in Table 5.18). Since this high level of accuracy achieved form ***rank 1***, it is a clear indication that there is a discriminative information presence in the user interaction which contributes towards right classification. The explanation of TPIR being high is attributed to the level of information uniqueness of a particular user in a particular application. Therefore, obtaining a precise user actions pattern leads to the creation of an accurate user profile.

There are many attributes that play a vital role in order to have a high TPIR; first one is the precision of user interactions signature of the application, size of user interactions and the number of user actions that can be identified within the application. These attributes have a vital influence on NN algorithm result, because NN is trying to identify a specific user based on learning the behavioral of such a user from the data and if the data does is not accurate enough to reveal what a user is doing, the classifier will find it difficult to assign this kind of traffic to the associated user.

The experimental results proved that applications that have an accurate user actions signatures and good number of user interactions achieved a high level of TPIR such as Skype, Outlook and BBC applications. Skype application is the evidently example of this type because it has a specific signature for the majority of its interactions and most of its user actions were identified. In fact, all of its signatures have a unique pattern, for instance, texting message has a baseline value and it is just one packet that is sent from the client to the Skype application server. In addition, calling by either video or audio has also a precise signature from the values perspective and by using a constant port. Subsequently, the system was able to determine the participants that have used Skype application with 100% level of accuracy.

In terms of users, there are many participants who achieved a positive performance and as mention above the number of applications usage, total number of interactions, and its amount in the accurate applications have an influence to the overall user performance. For example, users 9, 24 and 27 have achieved the highest average performance among the 27 users in ***rank 1*** with 68.6%, 79.5% and 88.5% respectively as shown in Table 5.8. By analyzing the applications usage in the nine applications as illustrates in the table below, all of them have acquired a high total number of user interactions in the nine applications.

ID	BBC	Drop.	FB	Google	Outlook	Skype	Twitter	Wiki	YouTube
9	124	3164	68764	7248	10902	2922	9950	1384	2.00E+05
24	320	2204	1922	5006	58032	19986	768	674	16020
27	28	6658	5460	27430	13330	9800	15910	68	40730

Table 5.19 : users with different rate of interactions

With reference to Figure 5.4 below, User 27 has achieved an excellent performance in the majority of application with TPIR almost 100% in 5 applications out of 9, thus, the overall performance has achieved a good result. This is due to the majority of their interactions in each application was quite sufficient for the system to be capable to create a user profile that distinguishes this particular user from others.

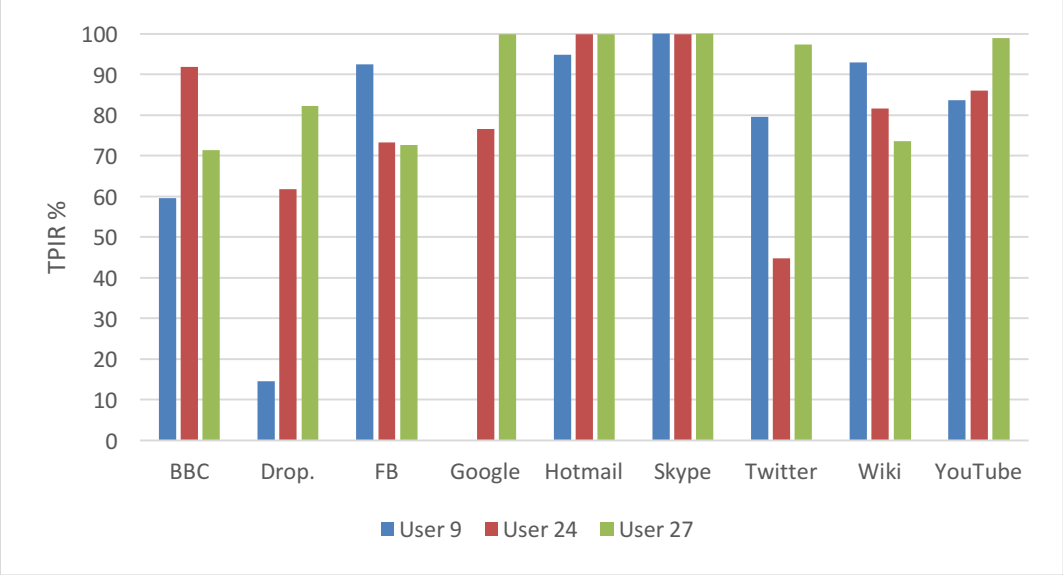

Figure 5.4 : Sample of users with high TPIR

In spite of the fact that user 9 has acquired a number of interactions in Google application as can be seen in Table 5.19, these interactions do not have a uniqueness of information that help the system to be able to build a user profile and the performance has achieved 0% for this application as can be seen in Figure 5.4 above. So, it is evident that user interaction should includ some uniqueness of information; otherwise the system will not able to create a unique profile that differentiate the user.

Nevertheless, BBC application has acquired the smallest number of user interaction but it has a uniqueness of information that is derived from application user actions such as watching video clips and listening to audio which help the system to be able to find a pattern in the user behavioral in order to properly create a user profile as illustrated in Table 5.10.

There are some applications that have acquired an altered user actions signatures which means the accuracy of the signature is slightly low. In this case the overall performance and induvial performance in these applications will be affected accordingly. One of the popular applications that are included in the experiment is Facebook, it has been used by all participants. Its

performance is slightly low and this is due to the nature of user actions were not quite accurate, even though, all of them could be identified, thereby the level of discriminative information is not sufficient to distinct the users.

Concerning the result analysis in terms of the accuracy of user actions signature impact, based on the available data, it can be concluded that a low accuracy of user actions leads to unsatisfying classification performance, although the classifications performance may be considered in a good level compared to large number of interactions that have been classified. For instance, YouTube, Google and Dropbox applications have acquired an altered user actions signature as discussed in the previous chapter and massive numbers of interactions, even though, they achieved a very positive performance with TPIR reached to more than 70% in ***rank 1***.

The result of this experiment proved that using user behavioral profiling based on user interactions that are derived at application level helped to make an accurate user identification system. The technique applied in this study is the identifications model, however, this does not mean it is going to be applied to one, 1000 or more. This is because about 99% of UK organizations as an example are considered as SMEs where the number of employee ranges between (1 to 249 employee) and a few are considered as enterprise organizations with thousands of employees (National Statistic, 2015). In addition, there is a technical solution that may help to in implementing this approach into enterprise organizations. For instance, an appropriate network architecture design to restrict IP pools based on physical location or logical business unit where it is possible to control and monitor sub sets rather than deal with a large number of peoples simultaneously.

5.4 Conclusion

The chapter proved that it is possible to successfully identify users based on their user actions that are derived at the application level. It is envisaged from the results that system performance is largely dependent on accuracy of user interactions, number of user interactions identified in each application and the amount of user interactions. The date gathered from the nine applications showed that there is a level of uniqueness of information within the user interactions. This kind of discriminative information does vary between users as can be seen in Table 5.8. In addition, the chapter presented the top three applications of each user that achieved the best TPIR and it showed that there is a number of participants that achieved a promising result where the system was able to identify their related traffic with accuracy attained 100% as presented in Table 5.19. Subsequently, using user interactions derived from application level for identifying users is a novel solution for user identification which aids towards identifying insider misuse efficiently.

6 Behavioral Fusion and Timeline Analysis

6.1 Introduction

Whilst the previous chapter has shown that it is possible to successfully identify users, the performance does vary between users. Looking to improve the performance and take into account the practical application of this technique, the chapter proposes two approaches. The first is to apply a fusion-based approach to the classification scheme and the second is to focus on the timestamps of interactions to see if decisions being made with stronger performing services can be utilised to identify interactions from poorer performing services (or services where too little data exists to create a classifier).

6.2 Fusion

Fusion is combining the information from different sources and it is a common approach that can be used in different domains. It firstly emerged in the literature in the 1960s, in data manipulation as a mathematical prototypical (Esteban et.al, 2004). Subsequently, it was applied in the US military field and utilized in different aspects for a long time. In the recent years, data fusion has received a significant attention due to the ability to improve the accuracy by observing features that are otherwise difficult to perceive in the first classification. Data fusion is based on user behavioral profiling that aims to solve the problem of some user interactions features in order to improve the performance of the system.

A multi-algorithmic approach would enable utilizing a range of biometrics classification algorithms (each model to focus on differing aspects of the problem) and combine the results through fusion.

Commonly, according to various studies, fusion can occur at different phases of the authentication process; sensor, feature, matching score, and/or decision level (Clarke, 2011; Ross, 2007; Sim et al., 2007) as outlined below.

- Sensor level fusion: The raw biometrics data is consolidated prior to feature extraction these data were captured by multiple sensors or by a single sensor acquiring multiple samples (for example fusing different face images from one or different cameras).
- Feature level fusion: After obtaining multiple samples from one or more biometrics traits, the feature vector is extracted from each sample using a variety of algorithms. These feature vectors are then merged together to be utilized in the following matching phase (e.g. fusing the feature vectors of the face and voice).
- Matching score level fusion: The produced results of multiple biometrics classifiers are joined at this level to produce a new accumulated match score to be utilized for the subsequent decision process.
- Decision fusion: This fusion happens when each incorporated biometrics system has provided its own decision to enable a final authentication decision.

Various studies have utilized fusion model to improve the performance of the biometrics technique in authentication mode. Table 6.1 illustrates a literature review of different studies that reflect the level of improvement occurred when fusion technique is utilized.

Studies	Biometric Modalities	Level of Fusion	Fusion Approach	#Users	Performance	
					FRR %	FAR %
Kumar et al.,2003	Palm print			100	4.5	2
	Hand geometry				8.3	5.3
	Multi-mode	Feature	Concatenation		5.1	2.3
		Match Score	Max rule		1.4	0
Snelick et al ,2003	Face			1000	24.1	0.1
	Fingerprint				17	0.1
	Multi-model	Match score	Min-Max+Simple Sum		5.1	0.1
Rose & Govindarajan,2005	Hand geometry			100	15	0.1
	Face				35	0.1
	Multi-mode	Match Score	Min-Max+Simple Sum		2	0.1
Jain et al,2005	Face			100	32.3	0.1
	Fingerprint				16.4	0.1
	Hand				53.2	0.1
	Multi-Mode	Match Score	Min-Max+Simple Sum		5.1	0.1
Koreman et al, 2006	Voice			30	3.2 EER	
	face				27.6 EER	
	Signature				8 EER	
	Multi-mode	Match Score	Min-MaxGMM		0.8 EER	
Kounoudes et al,2008	Voice			30	4.1	4.2
	Fingerprint				11.4	9.9
	Hand geometry				9.1	9.4
	multi-modal	Decision	majority Voting		0.9	1.2
Soltane et al, 2010	Face			30	0.4 EER	
	Speech				0 EER	
	Multi-modal	Match Score	Adoptive Bayesian		0.1 EER	
Saevanee et al,2014	linguistic				22 EER	
	Keystroke				20 EER	
	behavioral profiling				20 EER	
	Multi-modal	match Score	weighted average		8 EER	
Burriro et al.,2015	Voice			26	12.6	63.9
	arm movement				6.1	26.6
	multi-modal	Decision	weighted average		4.1	11

Table 6.1 : Summary of multi-modal biometric

With reference to the table above, it is evident that making integration between more than one biometrics models has improved the technique performance rather than just using individual biometrics. Match score level is a common Fusion manner that has been utilized in the

aforementioned studies. This is due to the technique feature where the existing and proprietary biometric systems are not affected. In addition, it is widely believed to be the most accurate approach of fusion type and thus it is the most utilized fusion type (Ross et al., 2006). Consequently, the fusion model that is based on match score level is going to be used to improve user identification performance based on user behavioral profiling model in this research as demonstrates in Figure 6.1.

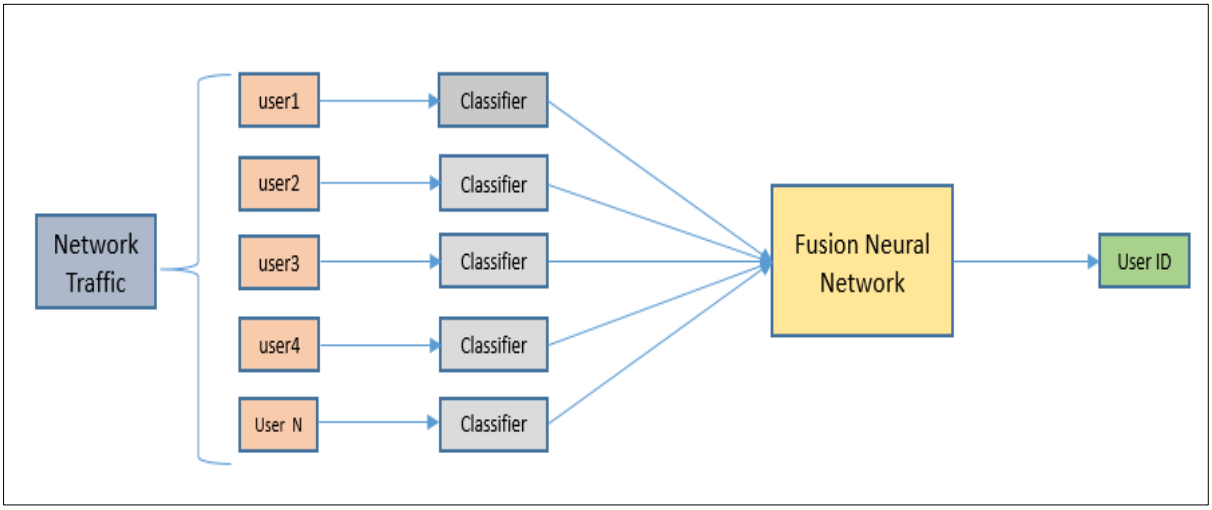

Figure 6.1 : Behavioural profiling fusion mode

6.3 Timeline Analysis

Timeline analysis refers to the process that scheduling information is used to develop a timeline of events that lead to the incident (Araste et al 2007). It is mainly used by forensic investigators because it includes information for instance when the files uploaded, accessed and changed and this information with their sequences are very critical from the forensics point of view perspective (Rocha,2014). Hosmer (1998) stated that the sequence timeline of computer events might be very useful to the forensic investigator. This is because it may provide a critical part of the information that leads to the prosecution of involved persons. Accordingly, using this concept in this study may offer another technique that helps to improve the user profiling, thus, the algorithm that is implemented in this study is working based on this aim in order to link the

interactions to the associated person without solely focusing on the IP address based on the realistic time window.

6.4 Methodology

6.4.1 Fusion Approach

Data fusion approach can take advantage of the concepts of variety and redundancy to improve system performance. Diversity can be utilized to improve system performance via the incorporation of different information. In the same view, redundancy can attain the same target through the re-use of data (Boujelbene et.al,2011). Therefore, the output of the MLP classifier has been re-used as an input to the fusion method here in order to improve the performance. The data divided in ½ for training and same part for testing with using the same FF-MLP classifier setting.

6.4.2 Timeline Analysis Approach

The timeline approach is different from the previous models because it sorts the event as it happens from the user without any change. It is useful in digital forensics for the purpose of identifying when a specific activity occurs on a computer and is fundamentally utilized for data reduction or identifying specific objects that happened (Basis Technology, 2013). As mentioned above, the timestamp is used to schedule the interactions, which means all interactions are sorted descending by using the timestamp. Although IP address is not reliable for a long period of time due to the use of mobile technologies and DHCP mechanism, there is a window of time that the IP address remains constant whilst the connection between the user and the application server does not terminate. Subsequently, this window is trusted enough to be associated with a particular user. Moreover, the technique has made a link between the fusion result and timeline analysis through the selected three applications that have high TPIR

for each user from fusion phase. This result represents the level of confidence that belongs to the associated user. Based on that, the system will seek for a value that can be matched with the predefined output in a particular application. Once the system has found this value, all interactions within a time window (the time windowing can be one of these values 30 Seconds, 60 Second, 120 Second and 240 seconds) are turned to be to the same user regardless of the classifier results as shown in the figure below. This technique is implemented on all participants for the main three applications.

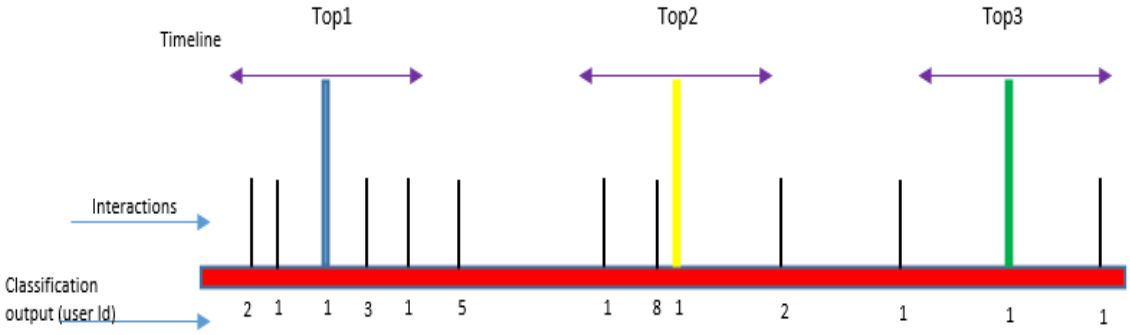

Figure 6.2 : Timeline analysis approach

6.5 Experimental Result

6.5.1 Fusion Result

Generally, the level of performance across all participants has improved by using fusion technique. This improvement does vary between users as can be seen in Table 6.2. The results in Table 6.2 illustrate that the average user identification rate by using fusion in ***rank 1*** has improved by 10% and achieved roughly 60% of accuracy compared to the standard approach where its TPIR achieved 50%, and this is due to giving a machine learning a chance to train and test the output of the first classification step in order to explore any pattern that has not been considered in the first attempt. With reference to the table below, 80% of the entire participants have acquired sufficient discriminative information in fusion mode which aids the system to

correctly classify ½ of their traffic in ***rank 1*** comparing with 70% in standard model for the same performance level. Also, this increase in the performance has been also recorded in ***ranks 3 and 5***. In fact, this 10% difference in fusion performance means that there are correctly classified user interactions by assigned them to the associated person.

User ID	Fusion %	Standard %	Fusion %	Standard %	Fusion %	Standard %
	TPIR Rank 1	TPIR Rank 1	TPIR Rank 3	TPIR Rank 3	TPIR Rank 5	TPIR Rank 5
1	50.5	49	59	55.3	66.9	63.9
2	64.4	48.2	73.9	70.1	79.8	74.4
3	55.4	46.3	68.3	64.8	81.2	74.4
4	80.5	65.8	89	78.3	90.7	82.1
5	45.9	32.3	64.3	38	69	51.7
6	61	55.1	71.9	71.8	77.9	79.8
7	47.6	36.9	60.3	69.1	70.2	80.2
8	70.5	60.8	75.9	67.6	78.9	68.7
9	82	68.6	89.4	75.4	92.8	82.6
10	51.7	39.4	69	56.9	74.6	63.9
11	65.4	51.2	78.1	55.1	84.2	57.9
12	72.3	65.3	82.4	75	85.4	78.8
13	60.1	54.1	71.2	63.3	82	69.5
14	50.7	34.9	64.6	62.5	69	72.7
15	67.8	59.8	85.3	80.1	88.2	84.9
16	40.5	31.1	65.5	53.2	71.4	60.1
17	31.4	28.4	43.9	39.1	46.3	43.6
18	66.9	64.1	80.3	73.6	89.7	75.7
19	55.7	45.5	71.5	60.1	77	71.2
20	64.1	44.7	74.5	51.5	86.4	64.2
21	50	50.6	69.1	71.8	79.3	87
22	36.3	19.1	72.3	66.4	80.9	68.6
23	55.7	41	70.5	54.4	73.3	61.7
24	79.2	79.5	87.3	84.2	94.1	85.8
25	47.2	53.7	55.1	61.2	60.2	68.3
26	63.2	50.2	68.6	52.9	72.4	54.8
27	90	88.5	95.6	91.6	98.6	92.8
Average	59.4	50.2	72.4	64.5	78.5	71

Table 6.2 : User identification rate by using fusion behavioural Model

There are some users that achieved a significant increase in their performance in fusion technique compared to the standard mode. For example, users 4, 9 and 22 have achieved TPIR 65.8%, 68.6 and 19.1% in ***rank 1*** respectively; however, this proportion of the same users has improved by about 15% for each to attained 80.5% in user4, 82% in users 9 and 36.3% in user 22. Although some participants have acquired a satisfying TPIR in the standard approach, their performance has also risen up after applying fusion mode; for instance, user 27 has achieved

88.5% accuracy in standard mode and the performance was improved to 90% in ***rank 1*** by applying the fusion method.

In terms of applications, the results show that there is a key outcome out of this experiment after applying fusion approach. Generally, applications that have achieved a high performance in ***rank 1*** are considered as a high accurate application in terms of user profile and after applying fusion technique the level of applications accuracy has increased and this reflects the level of discriminative information of user which contributes the classifier towards correct choice. Based on the results in Table 6.3, it is evident that all applications in fusion technique have achieved a better performance than the standard approach in all identification ranks. In fact, the fusion technique provided a good opportunity for the system to be able to successfully identify more than 80% of user interactions in five out of nine applications in ***rank 1***. In contrast, in standard model, this kind of performance had been achieved just in three applications as illustrated in Table 6.3. Hence, by utilizing fusion model the level of discriminative information in practice has improved in two applications from less than 80% of their traffic to be up to 80%, thus their overall performance has increased as well, such as Facebook and Google applications.

App. Name	No. Users	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1 %	Rank1 %	Rank3 %	Rank3 %	Rank5 %	Rank5 %
Skype	12	99.8	98.1	100	98.2	100	98.2
Outlook	19	97.3	96.2	98.9	96.9	99.3	97
Facebook	27	83.4	66.7	87.6	70.8	88.6	71.9
BBC	12	83.1	81.8	93.2	92.5	97	95.4
Google	27	82.1	71.7	88.8	79.4	90.4	82.2
Wikipedia	20	72.7	66.9	85.8	83.6	90.3	89.2
Twitter	27	72.4	65.3	85	79.5	89.3	83.4
YouTube	27	71.4	62.8	84	74.8	87.9	78.9
Dropbox	16	66.6	57.1	79.3	73.9	84.5	82.8

Table 6.3 : Average TPIR in different ranks per application

In comparison with standard approach, Skype and Outlook applications have achieved slightly higher TPIR in ***rank1*** with about 99% TPIR which means the fusion system has correctly classified almost all of their samples by assigning them to the associated user, as can be seen in Table 6.3. However, applying fusion technique has dramatically improved the overall performance of Facebook and Google applications with more than 10%, from 66.7% and 71.7% in standard mode, to being considered as part of the top four performance applications with TPIR attained 83.4% and 82.1% in ***rank 1***.

Skype application continues to score the most uniqueness of information in each experiment. It is obvious that more than 90% of its participants has successfully classified the whole of their interactions by the system in ***rank 1*** except user 11 where the TPIR achieved 85.4% as demonstrated in Table 6.4 compared to 80% in standard approach and this is due to using fusion approach which has improved the performance by reconsidering some of invisible features that the user has. In fact, the highest increase in fusion performance has been achieved by user 10 where standard approach correctly identified just 63.7% of the traffic, while in fusion the system has able to successfully identify user 10 traffic with accuracy attained 100% in ***rank 1***. This due to some features were being considered after applying fusion mode.

User ID	# Inter.	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1 %	Rank1 %	Rank3 %	Rank3 %	Rank5 %	Rank5 %
2	22	90.9	90.9	90.9	100	100	100
3	15	100	100	100	100	100	100
6	33764	100	100	100	100	100	100
9	1461	100	100	100	100	100	100
10	3849	100	63.7	100	63.7	100	63.7
11	604	85.4	80.3	100	85.2	100	94.8
12	20505	100	99.7	100	99.7	100	99.7
14	4529	99.9	99.9	99.9	99.9	99.9	99.9
19	9641	99.4	99.4	100	99.5	100	99.6
21	57	100	100	100	100	100	100
24	9993	99.9	99.9	100	99.9	100	99.9
27	4900	100	100	100	100	100	100
Average	**0**	97.9	94.4	99.2	95.6	99.9	96.4
Total # Inter.	**89340**						

Table 6.4 : Skype TPIR

Although the ***ranks 1*** in fusion approach has achieved one user result that has TPIR less than 90%, ***ranks 3*** has overcome this issue and registered high level of accuracy by scoring 90% as a low value for all users. However, in standard approach this kind of issues remain across all ranks such as user 10 where the TPIR achieved 63.7% in ***rank 5*** as can be seen in Table 6.4 above. Ultimately, by implementing fusion technique the system has more ability to profile the user from rank 1 where the minimum value of TPIR was 85.4%, however by normal classification this process might not be an easy task where the system faces some difficulties to profile the user in the most accurate application in this study.

In the same view, Outlook application has achieved average TPIR of more than 97% as shown in Table 6.3 above and this means that the majority of Outlook users traffic have been correctly classified. However, with reference to the Table 6.5 below, more than 2/3 of participants have utilized Outlook application during the experiment and almost 60% of users have achieved TPIR of more than 75% in ***rank 1***; Even though, it seems that users who have achieved promising TPIR have a good number of user interactions.

User ID	# inter.	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1 %	Rank1 %	Rank3 %	Rank3 %	Rank5 %	Rank5 %
2	87	80.4	50.7	82.7	72.4	88.5	75.8
3	138	55	60.8	55	60.8	56.5	60.8
4	2202	91.9	91.8	99.6	97.9	99.8	98.3
5	23	18.5	13	60.8	13	73.9	13
6	215	92	85.5	93.4	92	97.6	92
7	18	0	0	5.5	55.5	33.3	100
9	5451	96.4	94.9	98.9	95.4	99.8	95.4
10	1542	86.4	83	94.4	86	96.9	86
11	1581	95	80.5	97.1	80.6	97.2	80.7
12	1262	78.2	80.2	93.1	82	96.2	82.4
14	287	72.4	62.3	90.5	66.2	90.9	67.6
15	23	4.3	0	73.9	73.9	82.6	100
16	31	0	35.4	80.6	58	83.8	61.2
19	12706	97.5	95	99.1	95.6	99.2	96
20	16	0	0	12.5	0	75	56.2
21	171	66.6	85.3	66.6	85.3	98.2	85.3
23	28	75	71.4	96.4	71.4	96.4	71.4
24	29016	99.9	99.9	99.9	99.9	99.9	99.9
27	6665	99.7	99.8	99.7	99.8	99.7	99.9
Average		65.3	62.6	78.9	72.9	87.6	80.1
Total # Inter.	61462						

Table 6.5 : Outlook TPIR

In fact, within this proportion there is more than ¾ of them who have achieved an excellent level of accuracy of over 90% such as users 19, 24 and 27. In addition, this group has had significant improvement in the performance reach to 25% from ***rank 1*** to ***rank 5***, for instance user 23 has got 75% TPIR in ***rank 1***, but this value has gradually risen up to 96% in ***rank 5***.

Nevertheless, Outlook application has acquired a discriminative information that contributes to the system to record a positive result in terms of profiling a user, few users have not acquired sufficient uniqueness of information in ***rank 1***, however they could be identified in ***rank 2*** or ***3*** with a sufficient level of TPIR such as user 5 where the system fails to determine the traffic in ***rank 1*** but ***rank 2*** and ***3*** have achieved good level of TPIR where more than 60% of their traffic identified.

Facebook application is one of the applications that have significantly improved in its performance with fusion classifier compared to its result from the first experiment that was discussed in the previous chapter. After using fusion technique, it has achieved the third

promising result with average TPIR greater than 83% in ***rank 1*** as can be seen in Table 6.3 above. This application is one of the four applications that have been utilized by all users. Hence, it has discriminative information that contributes to the system to be able to correctly identify more than 40% of users with accuracy level reached to 70% and more in ***rank 1*** as demonstrate in Table 6.6.

User ID	No. inter.	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1 %	Rank1 %	Rank3 %	Rank3 %	Rank5 %	Rank5 %
1	264	0	0	12.1	25	31	46.2
2	5534	46	30.9	60.6	39.6	64.1	41.8
3	1947	66.6	45	70.4	45.5	70.8	46.4
4	20626	82.7	17.3	88.9	17.4	89.8	17.5
5	555	27.3	0.5	39.8	5	42.1	24.1
6	193	21.2	0.5	26.4	25.9	31	55.9
7	2172	28.7	19.7	45.8	78.5	47.3	89.5
8	31552	90.6	79.5	92.7	81.5	92.9	81.6
9	34382	92	92.5	95.5	97.3	95.9	97.7
10	1120	16.9	11.8	61.9	73.2	80.8	91.3
11	2675	71.1	48.6	75.9	48.8	76	48.8
12	3785	65.9	57	76.6	59.7	79.9	67
13	8350	88.1	72.7	89.9	74.8	89.9	74.8
14	720	7	4	20.6	56.9	31.2	90.9
15	21947	89.3	74.5	90.5	77.3	90.7	77.4
16	331	31.4	22.3	64	54.9	84.5	57.1
17	75	0	0	9.3	0	10.6	1.3
18	836	30.9	30.3	59.2	34.8	92.2	34.8
19	1655	27.7	20.6	51.7	39.8	72.8	67.3
20	11426	83.5	73.9	85.2	76	85.3	76.2
21	121	0	3.3	40.4	32.2	40.5	38
22	222	0	0	56.7	78.3	76.1	78.3
23	25821	92.5	75.5	93.2	75.6	93.3	75.6
24	961	78.3	73.3	85.3	80.5	92.9	83.9
25	926	0	45.8	0	60.5	0	65.5
26	12436	85.6	62.4	86.4	66.2	86.5	66.6
27	2730	76.8	72.6	97.3	72.7	99.5	72.7
Average		48.1	38.3	62	54.7	68.4	61.7
Total # Inter.	193362						

Table 6.6 : Facebook TPIR

In spite of the fact that ***rank 1*** has acquired some of the users who achieved low TPIR, this proportion has changed in the positive direction in ***ranks 2*** and ***3*** to score higher TPIR. For instance, the system was able to determine around 16% of 1120 interactions of user 10, and this amount has increased up to 60% in ***rank 3***. Subsequently, the result in Table 6.6 above; there is a significant increase in user level performance between fusion and standard approaches. This change might reach 50% in some cases, such as user 4 where the system was able to identify 17.3% of their traffic in standard approach; however, after applying fusion mode this proportion was sharply increased to achieve 82.7% in ***rank 1*** and the same thing

occurred in different users, for instance for users 5,6 and 26 with vary accuracy between 10% to 25%.

BBC application has not been used by all participants; around ½ of them have utilized BBC application. However, 75% of them has enough distinctive information from user interactions which provides a performance of more than 70% as illustrated in Table 6.7, in comparison with 6 participants that have acquired the same performance using the standard technique.

User ID	No. inter.	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1%	Rank1%	Rank3%	Rank3%	Rank5%	Rank5%
2	327	75.5	74.6	87.7	87.1	99.3	92.9
3	14	14.2	0	35.7	78.5	92.8	100
7	54	81.4	64.8	81.4	64.8	87	64.8
9	62	70.9	59.6	87.1	61.2	93.5	75.8
12	815	97.4	95	98.1	96.4	99.3	97.5
13	43	58.1	60.4	65.1	74.4	65.1	76.7
14	73	49.3	34.2	83.5	76.7	87.6	84.9
18	139	86.3	82	92	98.5	96.4	100
19	316	70.8	61.7	99.6	98.1	100	99.3
21	47	72.3	74.4	74.4	76.6	78.7	100
24	160	84.3	91.8	97.5	92.5	98.7	94.3
27	14	78.5	71.4	85.7	92.8	100	100
Average		69.9	64.1	82.3	82	91.5	90.5
Total # Inter.	2064						

Table 6.7 : BBC TPIR

Although, the system has failed to identify any interaction of user 3 using the standard method, fusion approach has correctly found out that 14.2% of user 3 traffic in ***rank 1*** and it represents the only value that has identified less than 58% TPIR. With reference to Table 6.7, all users that have utilized BBC application have achieved high TPIR in the remaining ranks and with clear and better performance in ***rank 1***of fusion side (roughly 6%) and slightly better also in the fusion side in the remaining ranks. In fact, 10 out of the 12 users have achieved a satisfying TPIR in ***rank 3*** with TPIR of more than 80% in fusion mode; however, this proportion has decreased to almost 50% of the users that have acquired the same performance using standard approach.

Google is one of applications that have improved in fusion phase by more than 10% in ***rank 1*** as can be seen in Table 6.3. This means that using fusion classifier is able to increase the performance through training the data of the user for a second round. Table 6.8 illustrates that there is a significant improve in the average performance level of users in fusion mode by scoring 69.6%, in comparison standard approach has acquired about 57% in ***rank 1***. It also shows that the system was able to successfully classify almost 2/3 of participants' traffic with level of accuracy up to 70% while the other technique 40% of participants could achieve this value.

In spite of the fact that the overall users average of TPIR in Google application in ***rank 1*** was 69.6%, in fusion mode as illustrated in Table 6.8, there is a set of users who achieved TPIR above 85% in ***rank 1***, such as user 1,4 and 27 especially they has large number of interactions

User ID	No. inter.	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1%	Rank1%	Rank3%	Rank3%	Rank5%	Rank5%
1	949	92.7	95.2	95.6	96.2	96.2	96.2
2	1568	52	32.4	56.3	80.1	58.1	87.1
3	213	75.1	59.1	85.9	95.7	86.8	100
4	7870	93.5	91.6	98.1	91.9	98.5	92.2
5	8750	80.1	73.6	87	79.3	88.7	79.8
6	1376	88.3	89.1	92.8	90.1	94.4	90.6
7	6295	86.7	79.2	89.2	87.4	89.6	88.1
8	1814	54.8	42.7	66.5	56.1	71.1	59.1
9	3624	90.9	0	95.4	0	95.5	13.6
10	1952	65.7	56.1	79.5	69.6	80.9	74.2
11	2297	70.1	72.7	86.8	76.5	92.9	77.1
12	15808	85.8	73	91.7	79.6	92.6	82
13	6727	81.2	75.3	89.8	86.2	91	88.4
14	61	0	26.2	0	90.1	0	100
15	1100	77.5	68.4	82.7	82	87	83.7
16	430	71.1	16.2	79.5	70.2	79.7	81.6
17	496	68.3	59	88.9	80.4	91.1	93.5
18	8454	76.7	68.3	88.6	77.4	90.5	82.2
19	2446	56	42.3	68.4	69.8	78	75.8
20	3313	80.9	63.2	88.2	64	89.7	64.4
21	1721	24.4	18.9	54.7	74.1	62.4	93.6
22	425	41.1	29.4	77.8	48.2	89.1	48.9
23	647	31	0	35.8	0	36.3	0.9
24	2503	80	76.5	83.8	81.8	86.1	86.1
25	1981	85.4	80.8	88.9	83.3	90.4	91.1
26	660	72.4	64.8	80.4	65	84.8	65
27	13715	99.7	99.9	99.9	99.9	99.9	100
Average		69.6	57.5	78.9	73.1	81.5	77.6
Total # Inter.	97195						

Table 6.8 : Google TPIR

Google application is considered as one of the top three applications in terms of usage with total number of interactions roughly 200 thousand, the fusion mode has learnt more the discriminative information which contributes to the system to infrequently score TPIR of less than 50% in ***rank 1***; hence, there is high potential to improve the performance in the remaining ranks; for example user 21 where the TPIR in ***rank 1*** was 24% but in ***rank 5*** this proportion attained to 62%.

With reference to the results in Table 6.8, the fusion performance has achieved better average of users in TPIR, in ***rank 1*** there is an improvement in TPIR by more than 10% compared to the standard approach and the same increase happened in the rest ranks with about 5%.

Although Wikipedia application is one of the applications that has a small amount of interactions compared to other applications, it has a sufficient discriminative information. Therefore, some of participants have achieved a promising performance in ***rank 1***. Table 6.9 below shows that more than ½ of users acquired at least 60% TPIR in the fusion mode in comparison with 40% of participants have acquired the same proportion using the standard approach. In fact, there are different users who achieved over 90% TPIR, for instance users 1, 8, 9 and 18 where the system almost identifies all their traffic as can be seen in Table 6.8.

User ID	No. inter.	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1%	Rank1%	Rank3%	Rank3%	Rank5%	Rank5%
1	20	100	100	100	100	100	100
2	116	74.1	74.1	77.5	78.4	78.4	78.4
3	82	6	6	60.9	31.7	97.5	79.2
7	57	50.8	49.1	61.4	61.4	77.1	61.4
8	22	100	100	100	100	100	100
9	692	93.2	93	94	94.8	95.6	95.2
10	27	7.4	0	29.6	11.1	29.6	29.6
12	116	67.2	56.8	72.4	67.2	76.7	75.8
13	15	0	0	33.3	20	86.6	53.3
14	15	60	0	73.3	0	86.6	0
15	399	67.6	64.6	95.2	99.2	98.2	100
16	114	95.6	95.6	95.6	95.6	95.6	95.6
17	84	57.1	52.3	69	70.2	73.8	73.8
18	50	98	98	98	98	98	98
19	107	25.2	0	61.6	21.5	67.2	72.9
20	407	76.4	57.7	89.4	88.9	91.1	94.6
23	102	50	20.5	81.3	68.6	83.3	91.1
24	337	78.9	81.6	91.1	93.7	99.4	93.7
26	21	71.4	76.1	71.4	80.9	71.4	85.7
27	34	73.5	73.5	85.2	73.5	91.1	73.5
Average		62.6	54.9	77	67.7	84.8	77.5
Total # Inter.	2817						

Table 6.9 : Wikipedia TPIR

Despite the fact that using fusion approach has improved the performance, in one case the system has failed to identify the user traffic (user 13), this approach is still better than the standard technique where the system was not able to correctly classify traffic of 4 users. In addition, there has been a gradual increase of TPIR value in fusion mode across all ranks between 12% as in ***rank 1*** and 5% in the rest.

Social network applications are the most popular applications that have been utilized by all participants in this experiment, such as Twitter. Applying the fusion technique has improved the performance of Twitter application by about 8% compared to single classification mode in ***rank 1*** as shown in Table 6.3. With reference to Table 6.10 below, average user identification rate in Twitter application has achieved a significant change between two approaches by more than 10% in ***rank 1*** and between 10% to 20% in ***ranks 3*** and ***5***.

User ID	No. inter.	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1%	Rank1%	Rank3%	Rank3%	Rank5%	Rank5%
1	133	2.2	0	25.5	1.5	45.1	21
2	5747	68.8	60.4	89	85.5	92.6	89.5
3	752	74.6	55	78.3	56.2	78.9	56.2
4	543	37.9	20.4	53.4	69.6	59.8	87.6
5	298	40.2	0.3	49.3	17.1	55	66.1
6	109	32.1	33	44.9	50.4	60.5	52.2
7	1683	36.9	17.1	71.6	53.9	84.9	63.3
8	107	45.7	25.2	53.2	40.1	57	42
9	4975	82.4	79.5	89.9	94	91.9	96.2
10	648	35.9	16	56	45.9	61.8	52
11	308	23.7	25.3	53.2	39.6	82.4	46.1
12	3517	68.7	67	88.9	70	95.7	71.8
13	596	52	49.1	66.1	62	83.3	65.7
14	868	63	6.4	68.3	59.6	69.8	83.5
15	4635	76.5	70.2	87.5	82.1	90.1	85.3
16	53	0	0	43.4	0	49	3.7
17	19	0	0	0	0	0	0
18	392	52.2	46.4	71.9	71.1	82.4	74.4
19	318	36.1	25.1	50.3	30.8	52.8	34.2
20	858	46.7	42.6	69.7	56	80.1	58.3
21	199	78.3	79.3	93.9	85.9	96.4	89.4
22	72	65.2	43	69.4	68	69.4	68
23	203	58.6	58.6	64.5	66.5	65.5	66.5
24	384	44	44.7	58.5	52.6	81.5	54.1
25	161	16.7	14.9	41.6	29.8	61.4	46.5
26	161	37.2	0	44.7	0	45.3	0
27	7955	97.7	97.3	99	98.9	99.5	99.7
Average		47.1	36.7	62.2	51.3	70	58.2
Total # Inter.	35694						

Table 6.10 : Twitter TPIR

This kind of improvement by using fusion approach has aided the system to correctly identify ½ of participants traffic with performance of up to 50%, however, by using the standard approach the system registered such a level of performance for ¼ of participants in ***rank 1*** as can be seen in Table 6.10 above. With regards to the fusion part in table above, there is some users the system could not be able to create a user profile for them, such as user 16 and 17 and this is because the number of interactions were quite low (53 and 19 interactions) and also these interactions did not have any uniqueness of information. However, user 16 was slightly different because the system was able to recognize their interactions and achieved a performance attained 43.4% and 49% in ***rank 3*** and ***5***.

In terms of number of participants and interactions, YouTube application has been utilized by all the participants and acquired a maximum number of user interactions. Table 6.11 below outlines the results of the two approaches; fusion and standard; of all 27 users' that have utilized YouTube application during the selected period. In general, the number of interactions is sufficient enough to give the system an ability to distinguish individual users from others if these interactions have a unique pattern for each user. According to ***rank 1*** result in Table 6.11, the overall average users TPIR shows that fusion system achieved better performance than the standard mode by almost 10%. This kind of positive improvement has a positive influence on user level identification; thus, the system was able to correctly recognize the traffic of 18 users out of 27 with level of performance of up to 50% compared to 10 users of the same value by using the standard approach. These outcomes indicate that although YouTube application suffers from lack of user action at the application level, it still could find out a unique pattern for the majority of participants and the level of uniqueness of information has improved by using the fusion technique as it has been proved. In fact, within this category, there are some users that achieved high improvements in their performance after applying the fusion model, for example users 5, 11, and 20 where their TPIR were 74% for user 5 and 0% for user 11 and 20 from the standard approach and after fusion has been applied the performance has significantly increased for all of them to score 82%,47% and 61.1% respectively as illustrated in Table 6.11 below.

User ID	No. inter.	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1%	Rank1%	Rank3%	Rank3%	Rank5%	Rank5%
1	3810	57.8	50	61.8	54.4	62.5	56.7
2	25131	39	31.2	65.2	37.5	81.7	40.6
3	11081	51.4	44.5	60.4	49.9	66.5	52.9
4	2E+05	89.8	90.9	98.5	98.3	99.2	99.1
5	42678	82.2	74.1	84.9	75	85.4	75.7
6	14159	39.9	34.8	74.3	71.8	81	82.4
7	26275	48.9	28.8	67	81.9	72.2	94.7
8	8702	61.3	56.5	67.3	60.1	73.5	60.9
9	1E+05	84.2	83.6	94.8	91.6	96.5	93.7
10	10025	49.8	45	61.9	48.8	72.1	50.8
11	13142	47	0	55.5	0	57	0.1
12	52668	54.4	46.3	83.2	90.2	90.4	95.3
13	7525	52	40.1	64.1	44.4	67.7	46.3
14	6673	49.3	33.1	54.2	42.4	55.7	44
15	39957	76.9	71.1	83.2	75.4	84.5	76.8
16	603	43.6	28	44.2	28.3	48.7	29.8
17	619	31.6	30.6	52.5	45	56.3	49.5
18	811	66.7	67.2	73.3	68.4	75.5	68.9
19	11951	35.3	21.9	48.1	37.1	54.8	44
20	44399	61.1	0	77	0.6	83.9	13.9
21	8695	16.4	17	57	52.6	84.1	96.8
22	1938	38.9	4.2	85.2	71.2	89.2	79.2
23	7356	27.3	20.2	51.7	44.7	65.4	64.6
24	8010	84.2	86	91.4	93.5	94	96.2
25	5946	58.3	51.1	61.3	56.3	62.1	62.5
26	6297	49.6	47.5	60.1	52.4	73.9	56.9
27	20365	99.3	99.9	100	100	100	100
Average		55.4	44.5	69.5	58.2	75.3	64.1
Total # Inter.	678816						

Table 6.11 : YouTube TPIR

In comparison with other applications, YouTube has a steady performance that has increased in different ranks. Subsequently, ***rank 5*** has registered a maximum rate of TPIR where the majority of users have acquired TPIR of more than 70%. It seems that it is quite rare to find an individual who can acquire performance of less than 50% using the fusion technique.

Similarly, Dropbox application suffers from a lack of user actions at the application level and has been utilized by about 60% of users. However, eventually it's the level of performance that is acquired in ***rank 1*** is promising as can be seen in Table 6.12 below. With reference to the table below, there is about 80% of participants that the system can correctly identify based on their traffic with a level of accuracy attained 50% and more by using the fusion technique and

this number is better than the number that can be obtained by using the standard mode by almost 40%.

User Id	No. inter.	Fusion	Standard	Fusion	Standard	Fusion	Standard
		Rank1%	Rank1%	Rank3%	Rank3%	Rank5%	Rank5%
2	3078	52.8	34.1	55	50	55.8	63.8
4	10925	86.9	82.5	95.8	94.5	97.4	97.6
6	4598	53.6	43.1	71.2	72.7	80.5	85.4
9	1582	28.2	14.6	49.1	44.6	67	75.9
12	270	33.3	12.9	37.4	30.3	37.7	38.1
13	2488	89.5	80.9	90.1	81.3	90.6	81.3
14	4450	55.5	48.2	91.3	71.1	99.1	83.8
15	6208	82.5	70	83.9	71.2	84.2	71.3
16	3067	41.7	20.6	51.5	65.7	58.7	92
18	2078	57.3	56.4	79.3	66.8	93.1	71.5
19	1989	53.4	43.7	65	49.4	68.7	51.3
20	85	100	75.2	100	75.2	100	85.8
21	3827	42.1	26.9	65.7	68.1	74.3	92.7
24	1102	63.3	61.8	78.5	63.7	94.4	64.2
25	179	75.9	75.9	83.8	75.9	87.1	75.9
27	3329	84.5	82.3	93.7	87.2	97.8	89.2
Average		62.5	51.8	74.4	66.7	80.4	76.2
Total # Inter.	49255						

Table 6.12 : Dropbox TPIR

In contrast, the number of users' interactions is a bit high; the system could be able to find out a pattern of each individual. Therefore, numbers of users have obtained high TPIR such as user 4, 13 and 20 where their performance in ***rank 1*** are 86.9%, 89.5% and 100% respectively.

In addition, this level of discriminative information continues to increase across the ranks where the system has correctly classified 11 users out of 16 with TPIR of more than 75% by using the fusion technique.

6.5.2 Top Three Applications

The criteria that used to select the top three applications here is based upon the value of the TPIR, which means the high three applications in TPIR is going to be chosen in this part. It is evident that applying the fusion mode has improved the level of users' average performance in the top three applications by roughly 7% in ***rank 1*** in each of them from the previous standard results. The accuracy of user profiling in first application has improved to 93.3% as an overall

average compared to about 87% in the standard mode as can be seen in Table 6.13 below. In fact, with reference to the results in Table 6.13, there are about 70% of participants that achieved a result above than the average in first top application, in addition, there are 11 participants in the fusion mode comparing with 8 in the standard acquired unique interactions that aided the system to create their profiling in a manner which can be 100% identified among other users. Skype, Outlook and Wikipedia applications are the most popular applications that have discriminative information compared to other applications in fusion. This is due to their user interactions signatures were accurate, thus, their user profiling achieved the best performance. Though, some users acquired a performance below the average, just two users (user 17 and 22) went far from the average and achieved low performance while the rest acquired TPIR within 5-10% below the average.

In contrast, using the standard technique the level of user profiling accuracy based on user interactions derived at application level was not quite sufficient to produce a unique user profiling, hence, there are just about 50% of the participants who achieved performance above the average which is 87.4% in first application and about half of them achieved 100% TPIR. In addition, there are some users that recorded a very low performance comparing with the average such as user 17 and 22 where their performance 59% and 43% but this proportions are better that the performance that achieved in the standard mode.

ID	Fusion		Standard		Fusion		Standard		Fusion		Standard	
	First App.		First App.		Second App.		Second App.		Third App.		Third App.	
	Name	TPIR %	Name	TPIR %	Name	TPIR %	Name	TPIR %	Name	TPIR %	Name	TPIR %
1	Wiki.	100	Wiki.	100	Google	95.2	Google	95.2	YouTube	57.8	YouTube	50
2	Skype	94.1	Skype	94.1	BBC	75.5	BBC	74.6	Wiki.	74.1	Wiki.	74.1
3	Skype	100	Skype	100	Google	75.1	Outlook	60.8	Twitter	74.6	Google	59.1
4	Google	93.5	Outlook	91.8	Outlook	91.9	Google	91.6	YouTube	89.9	YouTube	90.9
5	YouTube	82.2	YouTube	74.1	Google	80.1	Google	73.6	Outlook	40.2	Outlook	13
6	Skype	100	Skype	100	Outlook	92	Google	89.1	Google	88.3	Outlook	85.5
7	Google	86.7	Google	79.2	BBC	81.4	BBC	64.8	Wiki.	50	Wiki.	50
8	Wiki.	100	Wiki.	100	Facebook	90.6	Facebook	79.5	YouTube	61.3	YouTube	56.5
9	skype	100	skype	100	Outlook	96.4	Outlook	95	Wiki.	93.2	Wiki.	93
10	Skype	100	Outlook	83	Outlook	86.4	Skype	63.7	Google	65.7	Google	56.1
11	Outlook	95	Outlook	80.5	Skype	85.4	Skype	80.3	Facebook	71.1	Google	72.7
12	Skype	100	Skype	99.7	BBC	97.4	BBC	95	Google	85.8	Outlook	80.2
13	Dropbox	89.5	Dropbox	80.9	Facebook	88.1	Google	75.3	Google	81.2	Facebook	72.7
14	Skype	100	Skype	100	Outlook	72.4	Outlook	62.3	Twitter	63	Dropbox	48.2
15	Facebook	89.3	Facebook	74.5	Dropbox	82.5	YouTube	71.1	YouTube	76.9	Dropbox	70.9
16	Wiki.	95.6	Wiki.	95.6	Google	71.1	Outlook	35.4	YouTube	43.6	YouTube	28
17	Google	68.3	Google	59	Wiki.	57.1	Wiki.	52	YouTube	31.6	YouTube	30.6
18	Wiki.	98	Wiki.	98	BBC	86.3	BBC	82	Google	76.7	YouTube	67.2
19	Skype	99.4	Skype	99.4	Outlook	97.5	Outlook	95	BBC	70.8	BBC	61.7
20	Dropbox	100	Dropbox	75.2	Facebook	83.5	Facebook	73.9	Google	80.9	Google	63.7
21	Skype	100	Skype	100	Twitter	78.3	Outlook	85.3	BBC	72.3	Twitter	79.3
22	Twitter	65.2	Twitter	43	Google	41.1	Google	29.4	YouTube	38.9	YouTube	4.2
23	Facebook	92.5	Facebook	75.5	Outlook	75	Outlook	71.4	Twitter	58.6	Twitter	58.6
24	Outlook	100	Outlook	100	Skype	100	Skype	100	BBC	84.5	BBC	91.8
25	Google	85.4	Google	80.8	Dropbox	75.9	Dropbox	75.9	YouTube	58.3	YouTube	51.1
26	Facebook	85.6	Wiki.	76.1	Google	72.4	Google	64.8	Wiki.	71.4	Facebook	62.4
27	Skype	100	Skype	100	Google	100	Google	100	YouTube	100	YouTube	100
Avg.		**93.3**		**87.4**		**82.5**		**75**		**68.9**		**61.9**

Table 6.13 : Top three application results

At the user level, there are number of participants who have dramatically increased their performance between fusion and standard techniques and this might be associated with reorganizing the order numbers of applications in terms of performance. For instance, by using standard approach user 10 achieved in the first and second top applications (Outlook and Skype) 83% and 63.7% TPIR respectively; however, in fusion the order of top three applications has changed where skype application moved forward to be the first one with accuracy attained 100% and Outlook application step back to be the second application with a level of increasing in the performance achieving 86.4%. Moreover, there are some participants that have the same number of applications sequences in both techniques but with some differences in their performance in fusion technique, such as user 9, 24 and 27. This is because these applications are considered as top applications that have unique user actions which aided the system to correctly create a user profiling and then achieved high performance in the standard technique.

Similarly, the second and third top applications achieved discriminative information of some participants. In spite of the fact that, the level of performance here is quite low, but fusion technique has a better performance in the individual level and then in the overall average compared to the standard technique. Despite the fact that first application mostly consisted of the accurate applications as listed in Table 6.3, such as Skype and Outlook , there are other applications that are considered as less accurate based on their average performance, such as Dropbox and YouTube but in terms of individual some users have achieved promising results attained 100% of accuracy.

6.5.3 Timeline Analysis Result

Timeline analysis approach is built based on fusion results for each user as mentioned in the methodology section, the event sorted based on the time stamp and the top three application in TPIR that have been utilized as a trust point of the user action. Hence, the system assigns all events within the time windows to the one user. It is evident that using a correlation between top three applications results of fusion technique and timeline analysis have provided more authenticity to assign the interactions to the associated individual, and this hypothesis is supported by the results obtained from this experiment as shown in Table 6.14. This approach also provides a more realistic scenario in practice, as the samples are time synchronized and thus presented in order.

Table 6.14 clearly shows that applying timeline analysis on the user interactions based on the top three applications from fusion approach has significantly increased the overall performance with the first time window of 30 seconds achieved an average user performance of 70.2% and this performance is better than the results registered in ***rank 1*** using the fusion technique with more than 10%. This outcome is ongoing to be improved in each time window until attained

72.8% in 4 minutes as demonstrated in Table 6.14. It is shown that there is a direct relationship between length of time windowing and level of performance

Consequently, the number of user interactions that is labeled to the associated user by using this approach has been improved in the individual performance, thus, improved the overall performance.

User ID	Fusion %			Timeline Analysis %		
	TPIR Rank 1	TPIR Rank 3	TPIR Rank 5	30 Sec	60 Sec	240 Sec
1	50.5	59	66.9	68.6	69.6	71
2	64.4	73.9	79.8	46.6	47.2	50.5
3	55.4	68.3	81.2	55.5	56.6	58.1
4	80.5	89	90.7	93.1	94	96.7
5	45.9	64.3	69	88	88.7	89.5
6	61	71.9	77.9	79.8	79.8	79.8
7	47.6	60.3	70.2	54.9	54.9	55.2
8	70.5	75.9	78.9	91.3	92.1	93.6
9	82	89.4	92.8	86.5	86.9	88
10	51.7	69	74.6	66	67.4	72.1
11	65.4	78.1	84.2	61.1	62.7	67.7
12	72.3	82.4	85.4	77.7	80.1	81.8
13	60.1	71.2	82	81.3	83.1	84.6
14	50.7	64.6	69	64.7	65.4	72.9
15	67.8	85.3	88.2	90.4	91.7	93.2
16	40.5	65.5	71.4	46	46.1	46.8
17	31.4	43.9	46.3	48.6	48.6	49.8
18	66.9	80.3	89.7	72.5	74.3	76.9
19	55.7	71.5	77	76.1	77.2	79.1
20	64.1	74.5	86.4	67.1	67.4	67.7
21	50	69.1	79.3	26.1	26.1	26.4
22	36.3	72.3	80.9	37.6	38.2	38.5
23	55.7	70.5	73.3	84.1	85.5	86.5
24	79.2	87.3	94.1	94.8	94.8	94.8
25	47.2	55.1	60.2	61.3	62.6	64.6
26	63.2	68.6	72.4	79.5	81.1	83.2
27	90	95.6	98.6	97.5	97.6	97.7
Average	**59.4**	**72.4**	**78.5**	**70.2**	**71.1**	**72.8**

Table 6.14 : Timeline results

Table 6.14 above highlights that in the first time window (i.e 30 seconds), there is about ½ of participants that have achieved a promising performance where their TPIR more than the average (70.2%). In fact, amongst this group there are some users that acquired a positive result where their performance is over 90% such as users 4,8,15 and 27. In comparison, fusion technique in ***rank 1*** has achieved just one user achieved performance above 90%. The system shows that more than half of the participants have correctly identified their traffic with accuracy attained 70% and more. By using this technique, forensic investigator can potentially find more traffic associated to a user, for instance users 4, 24 and 27 as illustrated in Table 6.14.

It is clear that by increasing the time window, the outcomes of the system are going to be improved and this is happened here for the majority of participants. The comparison between timeline analysis in second time window (60 seconds) and fusion result in ***rank 3*** show that the performance of the second time window illustrates that about 60% of users achieved performances more than the 71.1% which is the average of time window, in addition, nearly 1/3 of them achieved 90% and more accuracy. While in the fusion model where the average performance is 72.4%, there is about 40% of participants that have acquired TPIR that exceeded this amount of 65% and above.

Similarly, by applying the time window of 240 seconds, the performance has increased. Although, each time window has recorded a positive difference in their performance, time windowing of 240 seconds has obtained the best. There are 3/4 of the entire participants has got TPIR more than 70%. Moreover, the majority of them the system is able to be identified their traffic correctly with performance 75% and more. Subsequently, this approach has investigated the event by using the original time of the event and using the most authenticate result (top three applications) as a fundamental basis for applying the timeline approach in order to raise the accuracy of the system and the beneficial of utilized this kind of correlation technique is clearly notable.

6.6 Discussion

Generally, using behavioral fusion approach has significantly improved the performance comparing with standard behavioral profiling. This kind of improvement was between 8 to 10% in each rank. Also, classifying the dataset twice (fusion technology) has produced an additional feature that leads to more discriminative information, thus the system was capable to build a more accurate user profile based on user interactions. Accordingly, by initiating a correlation between fusion behavioral results of the top three applications and timeline analysis

based on different time windowing an accurate user profile has been produced as demonstrated above in timeline result. Indeed, the timeline analysis results proved that the result that got in 30 second time windowing was almost better than behavioral fusion result in the first three ranks where the average user's performance attained 70.2%.

There are different applications that have achieved a high TPIR using behavioral fusion technique, such as Skype and Outlook. In fact, skype application continues to score the best application performance in both approaches. However, the level of its performance has increased in the fusion mode and achieved almost 100% in all ranks due the nature of its signatures were quite precious and the higher number of user action identified. In contrast, Dropbox application achieved the lowest performance using fusion model with 66.6 % TPIR (but still better than what the system got it from the standard approach by almost 10%). It is arguable that the reason behind this low in the performance because of the level of unique features inside this application was not enough for the system to be able to distinguish its users' traffic.

As can be seen in Table 6.2, fusion behavioral technique has performed better than behavioral profiling in terms of individual perspective. In fact, there are a numbers of participants who achieved a good performance in ***rank 1*** such as user 4, 9, 24 and 27, where their TPIRs are between 80% to 90%. This is due to the fact that it has been utilized almost by the nine services that were included in this experiments and high level of number of interactions especially in the most accurate applications such as skype, Outlook and Google. For instance, user 9 has experienced the nine services and with massive number of interactions attained more than 30,000 interactions as discussed in the previous chapter. Similarly, user 27 has achieved TPIR of 90% and by considering the number of services and interactions, we found that all services have been utilized by the user and the total number of interactions is about 120,000.

The lowest user identification rate in ***rank 1*** was achieved by users 17 and 22 when their results were 28.4% and 19.1%. These outcomes reveal that the uniqueness of information from those users' interactions was not quite enough to differentiate them from others, and in this case there are potential reasons: lack of user interactions in accurate applications or lack of applications used or both of them. As mentioned above the total number of user 17 interactions' was 2586 from five applications and the majority of interactions come from YouTube application while YouTube is considered as the second lowest application in level of accuracy. In the same vein, user 22 has experienced just four of the nine applications (i.e Facebook, Google, Twitter and YouTube) with total number of interactions 5,314 and more than 72% of them obtained from YouTube application as well.

In timeline analysis approach, four-time windows (i.e 30, 60, 120 and 240 seconds) were used in the analysis phase to explore to what extend the system can identify the related interactions in variety time windows as can be seen in Table 6.14 above. In general, the time line analysis approach has achieved a better performance by using 30 seconds time windows; comparing with first three ranks of the fusion approach. In fact, an accurate user profile was built by this approach which helped the system to score 70.2% as an average user identification rate in the first time window and this number has continuously increased each time window until achieved 72.8% in the last one. With reference to Table 6.14, the system could successfully identify some users' interactions with an excellent performance for instance, users 4, 8, 15, 24 and 27 where their accuracy was over 90%. Moreover, the majority of the rest participants' performance was achieved more than 50% in time windowing 30 second and this result has improved in the next time windowing (60 sec) as illustrated in Table 6.14.

To conclude, using these two approaches has produced an accurate performance results. These result means, it is possible to apply these approaches to link the user interactions to the

individual with high accuracy which aids organisations to identify the insider misuse. It should be noted however, the time window can only be set to a period where the underlying assumption that the traffic still belongs to the same user holds true. To pick a period window that is too large will suffer from misclassifications. Due to the nature of the data collection approach used to collect and label the traffic used in this research, it was not possible to increase the time window further to investigate this – as the result would simply have improved further (as the samples don't include different users using the same IP).

6.7 Conclusion

In general, utilizing the fusion behavioral and timeline analysis approaches can achieve a good level of accuracy and provides better classification performance than using behavioral profiling. By employing fusion behavioral profiling technique based on user interactions derived from the application level, the results show that a number of users have achieved average TPIR of more than 80%, in addition to the level of accuracy in different applications attained 100%, such as Skype.

Based on the success of these experiments results, the timeline analysis has been implemented upon fusion results with a specific time windowing in order to link more interactions to the associated individuals. The results show that this technique has significantly improved the average of user identification rate to score a result (70.1% with 30 seconds), better than fusion behavioral profiling. Subsequently, it is evident that it is possible to creating a user profiling from user actions derived from applications level with a high level of accuracy, thus an encouraging solution for attributing the insider misuse to the associated person through their interactions on different Internet services applications and overcome on the limitation that exist in the response model.

7 Conclusion and Future Work

The chapter concludes with the achievements of the research, highlights the limitations and opportunities for further investigation and research. The research aimed to determine whether it was possible to identify individuals from network traffic meta data.

7.1 Achievements of Research

Insider attack has become a serious problem for organizations and recently researchers sought to develop an applicable countermeasure to overcome this issue. The existence proposed solutions have to rely on the anomaly detection technique to analyze network traffic while there are serious limitations in this approach. Consequently, this research has overcome of these limitations by focusing on the user interactions of the user rather than the IP.

The following points are the main achievements of this research:

1. A feature extraction process was developed focusing on identifying and classifying user-level interactions with Internet-based applications. A significant contribution in this research was to only focus upon the metadata and not seek to perform any form of deep packet inspection.
2. An experimental investigation and evaluation of user behavioural profiling that showed that it is possible to successfully identify a good proportion of the users given the pattern in which they interact with Internet-based applications. Importantly, this was achieved by not using the user's IP address in any form. In order to aid scalability, a multi-classifier based approach was devised to complete N two-class problems, where N is the size of the user population.
3. The application of fusion to the biometric multi-classifier in order to aid and refine the classification and decision stages of the biometric system.

4. The creation of a timeline algorithm that utilizes the IP address as a basis for grouping interactions over short periods of time. The final outcome from the technique has shown a significant improvement in performance.

7.2 Limitations of Research

Despite the research objectives have been accomplished, a number of limitations linked to this research can be identified. The following list highlights these limitations:

1. The experimental dataset was limited in terms of participants and the duration of the collection period. Ideally, more participants and a longer profile period would have provided a more reliable measure of performance that could be achieved in practice.
2. Whilst it is not imperative for the proposed technique to have interactions extracted from every web-application (as the timeline experiment showed), having more applications modelled would have provided a richer and more comprehensive set of interaction signatures from which to extract from the dataset.

7.3 Future Research

A number of opportunities exist for further research and/or enhancement. These are outlined below:

- The underlying classifiers utilised in the classification stage were based upon the traditional feed-forward networks. Further exploration and investigation of alternative models and seeking to optimise them could provide for improvements in the recognition accuracy.
- The dataset could be expanded importantly in the number of participants – as this would permit further investigation of the scalability challenges of such an approach in practice. However, it would also be good to increase the duration in order to better understand

and model the changes that exist with regards to the permanence of the reference template.

- Automation of the identification of interactions from Internet-based applications. Whilst not required for the approach to work, the more applications are modelled and the resulting interactions signatures are extracted, the more accurate and reliable the resulting identification is likely to be. It will also provide a basis for creating revised signatures when existing applications are updated.

7.4 The Future of Behavioral Monitoring for User Identification

Internet services applications are becoming part of the daily lives of individuals and organizations, with more than three billion users now using internet services across the world. Unfortunately, with this rise in internet use comes increasing rates of cyber-related crime and insider misuse is considered one of the main threats that many organisations experience. Although many techniques currently exist for detecting or preventing insider misuse, this research highlights the need to go beyond the detection stage and provide a new mechanism that is able to attribute misuse actions to the individual who committed the crime. This research programme has, through an in-depth investigation of different user interactions, drawn from the applications level and built a means of behavioural-biometric profiling capable of identifying users. So, the depending on behavioural monitoring of user interactions is going to be highly required for identifying the identity of the individual.

References

Abramson, M. & Aha, D., 2013. User Authentication from Web Browsing Behavior. *The Twenty-Sixth International FLAIRS Conference*, pp.268–273. Available at: http://www.aaai.org/ocs/index.php/FLAIRS/FLAIRS13/paper/viewFile/5865/6081.

Academy, M. & Point, W., 2012. Biometrics Metrics Report. , (December), p.59.

Amit Mehta, Arjun Singh Parihar and Neeraj Mehta, 2015: Supervised Classification of

Dermoscopic Images using Optimized Fuzzy Clustering based Multi-Layer Feed-Forward Neural Network, IEEE International Conference on Computer, Communication and Control, Indore, 10-12 Sept. 2015

Arasteh, A.R. et al., 2007. Analyzing multiple logs for forensic evidence. *Digital Investigation*, 4(SUPPL.), pp.82–91.

Atallah, L. & Yang, G.Z., 2009. The use of pervasive sensing for behaviour profiling - a survey. *Pervasive and Mobile Computing*, 5(5), pp.447–464. Available at: http://dx.doi.org/10.1016/j.pmcj.2009.06.009.

Aupy, A. & Clarke, N., 2005. User Authentication by Service Utilisation Profiling. *Advances in Network and Communications Engineering 2*, p.18.

AusCert, 2003. Computer Crime and Security Survey 2003.

Banerjee, S.P. & Woodard, D., 2012. Biometric Authentication and Identification Using Keystroke Dynamics: A Survey. *Journal of Pattern Recognition Research*, 7(1), pp.116–139. Available at: http://jprr.org/index.php/jprr/article/view/427%5Cnhttp://www.jprr.org/index.php/jprr/article/view/427/167.

Banfield, R.E. et al., 2007. A Comparison of Decision Tree Ensemble Creation Techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(1), pp.173–180. Available at: http://ieeexplore.ieee.org/document/4016560/.

Bhattacharyya, D. et al., 2009. Biometric Authentication: A Review. *International Journal of Service, Science and technology*, 2(3), pp.13–28.

Blackwell, C., 2010. A Security Architecture to Protect against the Insider Threat from Damage, Fraud and Theft. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, 41 LNICST, pp.102–110. Available at: http://portal.acm.org/citation.cfm?doid=1558607.1558659.

Bolzoni, D. et al., 2006. POSEIDON: A 2-tier anomaly-based network intrusion detection system. *Proceedings - Fourth IEEE International Workshop on Information Assurance, IWIA 2006*, 2006, pp.144–156.

Bradford, P. & Hu, N., 2005. A layered approach to insider threat detection and proactive forensics. *Proceedings of the Twenty-First Annual Computer Security Applications Conference (Technology Blitz)*. Available at: http://www.acsa-admin.org/2005/techblitz/hu.pdf.

Bryce, C., Bryce, C. & Ave, L., 2013. Timeline Creation and Analysis Guides Written by. , (Lcdi).

BBC,2012.Eye scanners at England airports turned off, available on, http://www.bbc.co.uk/news/uk-england-17058448, accessed on 18th Feb 2014

Burge, P. & Shawe-Taylor, J., 1997. Detecting Cellular Fraud Using Adaptive Prototypes. *Proc of AI Approaches to Fraud Detection and Risk Management*, pp.9–13. Available at: http://eprints.soton.ac.uk/259660/.

Burnap, P. et al., 2015. Real-time classification of malicious URLs on Twitter using machine activity data. *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2015*, (August), pp.970–977.

Buschkes, R., Kesdogan, D. & Reichl, P., 1998. How to increase security in mobile networks by anomaly detection. *Proceedings 14th Annual Computer Security Applications Conference (Cat. No.98EX217)*, 4, pp.3–12. Available at: http://dblp.uni-trier.de/db/conf/acsac/acsac1998.html#BuschkesKR98.

Cappelli, D.M. & Trzeciak, R.F., Best Practices For Mitigating Insider Threat: Lessons Learned From 250 Cases.

Cappelli, D. et al., 2009. Common sense guide to prevention and detection of insider threats 3rd edition–version 3.1. *Cert*, (January), pp.1–88.

Caputo, D.D., Maloof, M.A. & Stephens, G.D., 2009. Detecting insider theft of trade secrets. *IEEE Security and Privacy*, 7(6), pp.14–21.

Castelluccia, C., 2012. European Data Protection: In Good Health? , pp.21–34. Available at: http://link.springer.com/10.1007/978-94-007-2903-2.

CERT, 2013. The CERT Insider Threat Center. *Carnegie Mellon Software Engineering Institute*. Available at: www.cert.org/insider_threat/.

CERT Australia, 2012. *Cyber Crime & Security Survey Report 2012*,

Chan, P.K. et al., 1999. Disfribufed Dam Mining in Credit Cud Fmud Defwdon.

Charistof Stormann, 1997. Fraud management tool : evaluation report. , pp.1–30.

Check, O., 2011. Check-in summary. , p.5.

Chee, O.C., 2007. Information Leakage, Detection, and Prevention. *The Global Voice of Information Security*, (December), pp.37–39.

Chiang, C.Y. et al., 2011. Seizure prediction based on classification of EEG synchronization patterns with on-line retraining and post-processing scheme. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp.7564–7569.

Choi, B. & Cho, K., 2012. Detection of Insider Attacks to the Web Server. *Journal of Wireless Mobile Networks, Ubiquitous ...*, pp.35–45. Available at: http://isyou.info/jowua/papers/jowua-v3n4-3.pdf.

Clarke, N. et al., 2004. Application of keystroke analysis to mobile text messaging. *... of the 3rd Security Conference, Las Vegas, ...*. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Application+of+keystroke+analysis+to+mobile+text+messaging#0.

Clarke, N. et al., 2017. Insider Misuse Identification using Transparent Biometrics. *Proceedings of the 50th Hawaii International Conference on System Sciences*, pp.4031–4040.

Costa, J. et al., 2012. Improved Blind Automatic Malicious Activity Detection in Honeypot Data. *Proceedings of the Seventh International Conference on Forensic Computer Science*, pp.46–55. Available at: http://www.icofcs.org/2012/papers-published-008.html.

Covavisaruch, N. et al., 2005. Personal Verification and Identification Using Hand Geometry. *Ecti Transactions on Computer and Information Technology*, 1(2), pp.134–140.

Crawford, M. & Peterson, G., 2013. Insider Threat Detection Using Virtual Machine Introspection. *2013 46th Hawaii International Conference on System Sciences*, pp.1821–1830. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6480061.

CSIEM, 2013. Cisco Security Information Event Management Deployment Guide.

CSO magazine, 2010. 2010 CyberSecurity Watch Survey – Survey Results Conducted by CSO magazine in cooperation with the U . S . Secret Service , Software Engineering Institute CERT Program at Carnegie Mellon University and Deloitte. , (August 2008), pp.1–17.

CYBERARK, 2016. Global Advanced Threat Landscape Survey. , (June). Available at: http://www.cyberark.com/resource/2016-global-advanced-threat-landscape-survey/.

Dainotti, a., Pescape, a. & Claffy, K.C., 2012. Issues and future directions in traffic classification. *IEEE Network*, 26(1), pp.35–40. Available at: http://ieeexplore.ieee.org/ielx5/65/6135845/06135854.pdf?tp=&arnumber=6135854&isnumber=6135845%5Cnhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6135854&tag=1.

Das, R., 2006. An introduction to biometrics A concise overview of the most important biometric technologies. *Keesing Journal of Documents & Identity*, (17), pp.3–5. Available at:

http://www.biometricnews.net/Publications/Biometrics_Article_Introduction_To_Biometrics.pdf.

Datardina, M. & Leung, K., 2009. Information Leakage & Data Loss Prevention. *IT Assurance & Governance Information*.

Department for Business, I. and S., 2015. Statistical Release Business Population Estimates for the Uk and Regions 2015. *Ghareeb*, pp.1–16. Available at: https://www.gov.uk/government/statistics/business-population-estimates-2015.

Deris Stiawan, 2012. Intrusion threat detection from insider attack using learning behavior-based. *International Journal of the Physical Sciences*, 7(4), pp.624–637.

Detica, 2011. THE COST OF CYBER CRIME. , p.4.

Dorigo, S., 2012. Security Information and Event Management. , p.75. Available at: www.ru.nl/publish/pages/769526/thesissanderdorigo.pdf.

En, P.N., 2013. Implementation of a timeline analysis software for digital forensic investigations.

EventLog Analyser, 2017.SIEM. Available at: https://www.manageengine.com/products/eventlog/manageengine-siem-whitepaper.html. Accessed on 15th Jan 2017.

Environment, D., 2016. SIEM : Keeping Pace with Big Security Data.

Fawcett, T. & Provost, F., 1999. Activity monitoring: Noticing interesting changes in behavior. *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, 1(212), pp.53–62. Available at: http://portal.acm.org/citation.cfm?id=312195.

Fawcett, T. & Provost, F., 1997. Adaptive Fraud Detection. *Data Mining and Knowledge Discovery*, 316(1), pp.291–316.

Flegel, U., Vayssière, J. & Bitz, G., 2010. A State of the Art Survey of Fraud Detection Technology. *Insider Threats in Cyber Security*, 49, pp.73–84. Available at: http://link.springer.com/10.1007/978-1-4419-7133-3%5Cnhttp://link.springer.com/chapter/10.1007/978-1-4419-7133-3_4.

Fridman, L. et al., 2016. Active Authentication on Mobile Devices via Stylometry, Application Usage, Web Browsing, and GPS Location. *IEEE Systems Journal*, pp.1–10.

Gessiou, E., Vu, Q.H. & Ioannidis, S., 2013. IRILD : an Information Retrieval based method for Information Leak Detection.

Giles Hogben, 2010. ENISA Briefing : Behavioural Biometrics. *Computational Intelligence*.

Gordon, L. a et al., 2006. Computer Crime and Security Survey. *Computer Crime and Security Survey*, (11), pp.0–28.

Great Britain. Department for Business Innovation and Skills, 2014. 2014 Information Security Breaches Survey: Technical Report. , p.22.

Greitzer, F.L. et al., 2009. Predictive modeling for insider threat mitigation. *Contract*, (April), pp.1–14. Available at: http://www.pnl.gov/cogInformatics/media/pdf/TR-PACMAN-65204.pdf.

Grosser, H., Britos, P. & Garcia-Martinez, R., 2005. Detecting fraud in mobile telephony using neural networks. *Innovations in Applied Artificial Intelligence*, 3533, pp.613–615.

Hall, J., Barbeau, M. & Kranakis, E., 2005. Anomaly-based intrusion detection using mobility profiles of public transportation users. *WiMob'2005), IEEE International Conference on Wireless And Mobile Computing, Networking And Communications, 2005.*, 2, pp.17–24. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1512845.

Haller, J. et al., 2011. Best Practices for National Cyber Security: Building a National Computer Security Incident Management Capability, Version 2. 0., (April), p.40. Available at: http://www.sei.cmu.edu/reports/11tr015.pdf.

Hargreaves, C. & Patterson, J., 2012. An automated timeline reconstruction approach for digital forensic investigations. *Digital Investigation*, 9(SUPPL.).

Hargreaves, C. & Patterson, J., 2011. DIGITAL FORENSIC RESEARCH CONFERENCE An Automated Timeline Reconstruction Approach for Digital Forensic Investigations An automated timeline reconstruction approach for digital forensic investigations. , 44(0). Available at: http:/dfrws.org.

Herrmann, D. et al., 2010. Analyzing Characteristic Host Access Patterns for Re-Identification of Web User Sessions.

Hilas, C.S. & Sahalos, J.N., 2005. User profiling for fraud detection in telecommunication networks. *5th International Conference on Technology and Automation*, pp.382–387.

HM Government, 2015. 2015 Information Security. , p.49. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/432412/bis-15-302-information_security_breaches_survey_2015-full-report.pdf.

Hollmén, J., 2000. *User profiling and classification for fraud detection in mobile communications networks*, Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.79.6058&rep=rep1&type=pdf.

Hosmer, C., Time-lining computer evidence. *1998 IEEE Information Technology Conference, Information Environment for the Future (Cat. No.98EX228)*, pp.109–112. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=713393.

Hu, Y. et al., 2011. Profiling file repository access patterns for identifying data exfiltration activities. *IEEE SSCI 2011: Symposium Series on Computational Intelligence - CICS*

2011: 2011 IEEE Symposium on Computational Intelligence in Cyber Security, pp.122–128.

Huang, G.-B. et al., 2012. Extreme learning machine for regression and multiclass classification. *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics*, 42(2), pp.513–29.

Imhanwa, S., Greenhill, A. & Owrak, A., 2015. Relevance of Cloud Computing: A Case for UK Small and Medium Sized Tourism Firms. *GSTF Journal on Computing*, 4(3), pp.1–10.

Innovation, B., 2013. 2013 INFORMATION SECURITY BREACHES SURVEY Commissioned by: Conducted by: In association with: Information security: Cover image:

Institute, P., 2012. 2012 Cost of Cyber Crime Study: United States., (October), p.30. Available at: http://www.ponemon.org/local/upload/file/2012_US_Cost_of_Cyber_Crime_Study_FINAL6 .pdf.

International Chamber of Commerce, 2005. Intellectual Property: Source of innovation, creativity, growth and progress. *Growth (Lakeland)*, pp.1–20.

ISACA,2010. Data leak prevention, available on, http://www.isaca.org/Search/Pages/DefaultResults.aspx?k=DLP&s=Site%20Content&start1=0&ct=Site&cs=About%20ISACA&scopes=People,Site%20Content,Conversations, accessed on 17th Feb 2015.

Internet Society, 2009. IP Address Affinity. , (June), pp.1–4.

J. Eom, M Oark, S Park, T.C., 2013. A Framework of Defense System for Prevention of Insider's Malicious Behaviors. *Journal of Chemical Information and Modeling*, 53(9), pp.1689–1699.

Jain, A.K. & Nandakumar, K., 2012. Biometric authentication: System security and user privacy. *Computer*, 45(11), pp.87–92.

Jain, A.K., Ross, A. & Prabhakar, S., 2004. An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), pp.4–20.

Jakobsson, M. et al., 2009. Implicit authentication for mobile devices. *HotSec*, p.9.

Jin, Y. et al., 2012. A Modular Machine Learning System for Flow-Level Traffic Classification in Large Networks. *ACM Transactions on Knowledge Discovery from Data*, 6(1), pp.1–34.

Kamesh & Sakthi Priya, N., 2014. Security enhancement of authenticated RFID generation. *International Journal of Applied Engineering Research*, 9(22), pp.5968–5974.

Kancherla, R., 2008. Behavioral Fraud Mitigation through Trend Offsets. , pp.1–11.

Karlzén, H., 2009. An Analysis of Security Information and Event Management Systems. *Analysis*, (January).

Kaspersky, 2012. Global IT Security Risks: 2012. , pp.1–21.

Keeney, M. & Rogers, S., 2010. 2010 Cybersecurity Watch Survey: Cybercrime Increasing Faster Than Some Company Defenses. , pp.1–21. Available at: http://www.cert.org/insider-threat/research/cybersecurity-watch-survey.cfm?

Khan, H. & Hengartner, U., 2014. Towards application-centric implicit authentication on smartphones. *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications - HotMobile '14*, pp.1–6. Available at: http://dl.acm.org/citation.cfm?doid=2565585.2565590.

Khan, M.N.A., Chatwin, C.R. & Young, R.C.D., 2007. A framework for post-event timeline reconstruction using neural networks. *Digital Investigation*, 4(3–4), pp.146–157.

Kind, A., Stoecklin, M.P. & Dimitropoulos, X., 2009. Histogram-Based Traf fi c Anomaly Detection. , 6(2), pp.110–121.

Lafayette, W., 2011. CERIAS Tech Report 2011-25 Ensemble Classification for Relational Domains by Hoda Eldardiry Information Assurance and Security.

Li, F., Clarke, N. & Papadaki, M., 2009. Intrusion detection system for mobile devices: Investigation on calling activity. *8th Security Conference*.

Li, F. et al., 2010. Behaviour profiling on mobile devices. *Proceedings - EST 2010 - 2010 International Conference on Emerging Security Technologies, ROBOSEC 2010 - Robots and Security, LAB-RS 2010 - Learning and Adaptive Behavior in Robotic Systems*, (2010), pp.77–82.

Li, F. et al., 2011. Misuse Detection for Mobile Devices Using Behaviour Profiling. *International Journal of Cyber Warfare and Terrorism*, 1(1), pp.41–53.

Li, Y. & Guo, L., 2007. An active learning based TCM-KNN algorithm for supervised network intrusion detection. *Computers and Security*, 26(7–8), pp.459–467.

Liu, B. et al., 2014. DECAF: Detecting and Characterizing Ad Fraud in Mobile Apps. *Nsdi 2014*. Available at: https://www.usenix.org/conference/nsdi14/technical-sessions/presentation/liu_bin.

Joseph Lewis,2002, University of Maryland, Bowie State University, "Biometrics for secure Identity Verification: Trends and Developments" January 2002. (journal style)

Liu, S. & Kuhn, R., 2010. Data Loss Prevention. *IT Professional*, 12(2), pp.2008–2011.

Maes, S., Tuyls, K. & Vanschoenwinkel, B., 1993. Credit Card Fraud Detection Using Bayesian and Neural Networks. *Maciunas RJ, editor. Interactive image-guided neurosurgery. American Association Neurological Surgeons*, pp.261–270.

Mahoney, M. & Chan, P.K., 2001. PHAD: Packet header anomaly detection for identifying hostile network traffic. *Florida Institute of Technology technical report CS-2001-04*, (1998), pp.1–17. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.15.4041&rep=rep1&type=pdf.

Mahoney, M. V., 2003. A Machine Learning Approach to Detecting Attacks by Identifying Anomalies in Network Traffic. *Computer*. Available at: http://www.cs.fit.edu/~mmahoney/dist/diss.pdf

MRSYS (2016): Automated Fingerprint Identification System (AFIS),[online] http://www.m2sys.com/automated-fingerprint-identification-system-afis/,[accessed on] 10th Jun 2016

Mahoney, M. V. & Chan, P.K., 2003. Learning Rules for Anomaly Detection of Hostile Network Traffic. *International Conference on Data Mining (ICDM)*, pp.601–604. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.6.6512%5Cnhttps://www.thc.org/root/docs/intrusion_detection/nids/Learning-Rules-for-Anomaly-Detection-of-Hostile-Network-Traffic.pdf.

Meier, S., Fuchs, L. & Pernul, G., 2013. Managing the Access Grid A Process View to Minimize Insider Misuse Risks. *Wi*, (March), pp.1–15.

Moskovitch, R. et al., 2009. Identity theft, computers and behavioral biometrics. *2009 IEEE International Conference on Intelligence and Security Informatics, ISI 2009*, pp.155–160.

Myers, J., Grimaila, M.R. & Mills, R.F., 2009. Towards insider threat detection using web server logs. *Proceedings of the 5th Annual Workshop on Cyber Security and Information Intelligence Research Cyber Security and Information Intelligence Challenges and Strategies - CSIIRW '09*, p.1. Available at: http://portal.acm.org/citation.cfm?doid=1558607.1558670.

Ngai, E.W.T. et al., 2011. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), pp.559–569. Available at: http://dx.doi.org/10.1016/j.dss.2010.08.006.

Nicolett, M. & Kavanagh, K.M., 2011. Gartner research: Magic quadrant for security information and event management evaluation criteria definitions. , (May), pp.1–32.

Of, C. & Acm, T.H.E., 2000. B Iometric. *Communications of the ACM*, 43(2).

Ogwueleka, F., 2009. Fraud Detection In Mobile Communications Networks Using User Profiling And Classification Techniques. *Journal of Science and Technology (Ghana)*, 29(3), pp.31–42.

Ogwueleka, F., 2009. Fraud Detection In Mobile Communications Networks Using User Profiling And Classification Techniques. *Journal of Science and Technology (Ghana)*, 29(3), pp.31–42.

Oh, I. & Suen, C., 2002. A class-modular feedforward neural network for handwriting recognition. *Pattern Recognition*, 35(1), pp.229–244. Available at: http://www.sciencedirect.com.ezproxy.cqu.edu.au/science/article/B6V14-4477X94-P/2/b11d5073d65a95041f227deeb9b9292f%5Cnhttp://www.sciencedirect.com.ezproxy.cqu.edu.au/science?_ob=MImg&_imagekey=B6V14-4477X94-P-4P&_cdi=5664&_user=409397&_pii=S00313203000018.

Oh, J., Lee, S. & Lee, S., 2011. Advanced evidence collection and analysis of web browser activity. *Digital Investigation*, 8(SUPPL.).

Olsson, J. & Boldt, M., 2009. Computer forensic timeline visualization tool. *Digital Investigation*, 6(SUPPL.), pp.S78–S87. Available at: http://dx.doi.org/10.1016/j.diin.2009.06.008.

Pannell, G. & Ashman, H., 2010. Anomaly Detection over User Profiles for Intrusion Detection. , (November).

Patcha, A. & Park, J.M., 2007. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks*, 51(12), pp.3448–3470.

Perman, L. & Reznik, L., 2003. Anomaly Intrusion Detection Based on User ' s Behavior Profiling.

Peter Helms, 2012. Security Information & Event Management (SIEM). , (February). Available at: http://www.esl.dk/media/35846/McAfee Security Information & Event Management (SIEM).pdf.

Petrovska-Delacrétaz, D., Chollet, G. & Dorizzi, B., 2009. Guide to biometric reference systems and performance evaluation. *Guide to Biometric Reference Systems and Performance Evaluation*, pp.1–382.

Plate, J., 2009. *Internet Technologies*,

Ponemon Institute, 2013. 2013 Cost of Cyber Crime study: United States. *Ponemon Institute Research Report*, (October), pp.1–30. Available at: http://media.scmagazine.com/documents/54/2013_us_ccc_report_final_6-1_13455.pdf.

PricewaterhouseCoopers International Limited, 2014. Managing Cyber risks in an interconnected world: Key findings from the global state of onformation security survey 2015. , (September 2014), p.42. Available at: http://www.pwc.com/gsiss2015.

PWC, 2010. Information Security Breaches Survey 2010 Technical Report. , pp.1–22. Available at: http://www.google.co.uk/url?sa=t&rct=j&q=information security breaches survey 2010 technical report&source=web&cd=1&cad=rja&ved=0CC0QFjAA&url=http://www.pwc.co.uk/au

dit-assurance/publications/isbs-survey-2010.jhtml&ei=RwG_UY2WGoTWPPn8gdAP&usg=AFQjCNH-1nwv.

PwC, 2014. Managing cyber risks in an interconnected world. , (1).

Priti &Vishakha (2013): Classification using Different Normalization Techniques in Support Vector Machine, [online] http://research.ijcaonline.org/icct/number2/icct1313.pdf, accessed on 29th Jun 2017.

Qayyum, S. et al., 2010. Fraudulent call detection for mobile networks. *Information and Emerging Technologies (ICIET), 2010 International Conference on*, pp.1–5.

Ratha, N.K., Connell, J.H. & Bolle, R.M., 2001. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 40(3), pp.614–634.

Rexer, K., 2011. 5th Annual Data Miner Survey-2011 Survey Summary Report. *Rexer Analytics, Winchester*. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:5th+Annual+Data+Miner+Survey+–+2011+Survey+Summary+Report#0%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:5th+Annual+Data+Miner+Survey-2011+Survey+Summary+Report#0.

Ross, A. & Govindarajan, R., 2004. Feature level fusion in biometric systems. *In proceedings of Biometric Consortium Conference*, 2(c), pp.2–3. Available at: http://www.nws-sa.com/biometrics/ear/featureFusion.pdf.

Y. H. Cho, S. Navab and W. H. Mangione-Smith, “specialized hardware for deep network packet filtering", proceeding of the 12th international conference, pp 452-461, montpellier, france, september 2–4, 2002, in press.

Sandvine, “Global Internet Phenomena Report 1H 2014”, [Online], https://www.sandvine.com/downloads/general/global-internet-phenomena/2014/1h-2014-global-internet-phenomena-report.pdf, date accessed 15 January 2015.

Tsang-Long Pao; Po-Wei Wang, "NetFlow based intrusion detection system," in Networking, Sensing and Control, 2004 IEEE International Conference on , vol.2, no., pp.731-736 Vol.2, 2004 doi: 10.1109/ICNSC.2004.1297037

Crotti, M.; Gringoli, F.; Pelosato, P.; Salgarelli, L., "A statistical approach to IP-level classification of network traffic," in Communications, 2006. ICC '06. IEEE International Conference on , vol.1, no., pp.170-176, June 2006 doi: 10.1109/ICC.2006.254723.

Song, Sui; Li Ling; Manikopoulo, C.N., "Flow-based Statistical Aggregation Schemes for Network Anomaly Detection," in Networking, Sensing and Control, 2006. ICNSC '06. Proceedings of the 2006 IEEE International Conference on , vol., no.,pp.786-791, doi: 10.1109/ICNSC.2006.1673246

Muraleedharan, N.; Parmar, A.; Kumar, M., "A flow based anomaly detection system using chi-square technique," in Advance Computing Conference (IACC), 2010 IEEE 2nd International , vol., no., pp.285-289, 19-20Feb.2010doi: 10.1109/IADCC.2010.5422996

Braga, R.; Mota, E.; Passito, A., "Lightweight DDoS flooding attack detection using NOX/OpenFlow," in Local Computer Networks (LCN), 2010 IEEE 35th Conference on , vol., no.,pp.408-415,10-14Oct.2010 doi: 10.1109/LCN.2010.573575

Winter, P.; Hermann, E.; Zeilinger, M., "Inductive Intrusion Detection in Flow-Based Network Data Using One-Class Support Vector Machines," in New Technologies, Mobility and Security (NTMS), 2011 4th IFIP International Conference on , vol., no., pp.1-5, 7-10 Feb. 2011
doi: 10.1109/NTMS.2011.5720582

Tegeler, Florian; Fu, Xiaoming; Vigna, Giovanni; and Kruegel, Christopher (2012) "BotFinder: finding bots in network traffic without deep packet inspection", proceeding of the 8th international conference on Emerging networking experiments and technologies, page 349-360 ACM New York, NY, USA

Saat, R. et al., 2013. Insider threat Prediction Model on Information Leakage in Cloud Computing Environment II . Insider threat prediction model. , 7, pp.26–32.

Saevanee, H., Clarke, N.L. & Furnell, S.M., 2011. Behavioural Biometric Authentication For Mobile Devices. *Proceedings of the 2011 Collaborative European Research Conference - CERC*, 376, pp.175–184.

Saevanee, H., Clarke, N. & Furnell, S., 2011. SMS linguistic profiling authentication on mobile device. *Proceedings - 2011 5th International Conference on Network and System Security, NSS 2011*, pp.224–228.

Salem, M.B. & Stolfo, S.J., 2009. Masquarade attack detection using a search-behavior modeling approach. , pp.1–17.

Salem, M.B. & Stolfo, S.J., 2011. Modeling User Search-Behavior for Masquerade Detection. *The Innovator*, (Ml), p.16. Available at: http://www.bits.org/publications/doc/InnovatorFeb2011.pdf#page=16.

Salem, M.M.B., Hershkop, S. & Stolfo, S.J.S., 2008. A Survey of Insider Attack Detection Research. *Advances in Information Security*, 39, pp.69–90. Available at: http://link.springer.com/chapter/10.1007/978-0-387-77322-3_5.

Samfat, D. & Molva, R., 1997. IDAMN: An intrusion detection architecture for mobile networks. *IEEE Journal on Selected Areas in Communications*, 15(7), pp.1373–1380.

Symantec,2015, Data Loss Prevention Solution, available on, https://www.symantec.com/content/dam/symantec/docs/data-sheets/data-loss-prevention-solution-en.pdf, accessed on 1st Mar 2015.

SANS, 2012. Host-Based Detection and Data Loss Prevention Using Open Source Tools.

SANS, 2003. SSL and TLS: A Beginners Guide This. *Information Security*, p.18. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Intrusion+Detection+Is+Dead+.+Long+Live+Intrusion+Prevention!#0.

Sharma et al (2013): Extreme Machine Learning: Feed Forward Networks,[online] https://www.ijarcsse.com/docs/papers/Volume_3/8_August2013/V3I8-0470.pdf,[accessed on] 15th July 2017

Silowash et al. (2012): Common Sense Guide to Mitigating Insider Threats 4th Edition,[online on] http://resources.sei.cmu.edu/asset_files/TechnicalReport/2012_005_001_34033.pdf,accessed on 15th July 2017

Senator, T.E. et al., 2013. Detecting insider threats in a real corporate database of computer usage activity. *KDD - International conference on Knowledge discovery and data mining*, p.1393. Available at: http://dl.acm.org/citation.cfm?doid=2487575.2488213.

Shi, E. et al., 2011. Implicit authentication through learning user behavior. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6531 LNCS, pp.99–113.

Saed.F,2016. One in 50 employees could be a malicious insider. Available on, *https://betanews.com/2016/09/16/one-in-50-employees-could-be-malicious-insider/* Accessed on, 15th Dec 2016.

Sibai, F.M. & Menascé, D.A., 2012. Countering network-centric insider threats through self-protective autonomic rule generation. *Proceedings of the 2012 IEEE 6th International Conference on Software Security and Reliability, SERE 2012*, pp.273–282.

Sibai, F.M. & Menascé, D.A., 2011. Defeating the insider threat via autonomic network capabilities. *2011 3rd International Conference on Communication Systems and Networks, COMSNETS 2011*.

Silowash.G et al. (2012): Common Sense Guide to Mitigating Insider Threats 4th Edition ,[online on], http://www.sei.cmu.edu/reports/12tr012.pdf,[accessed on] 15 Jan 2017

Soysal, M. & Schmidt, E.G., 2010. Machine learning algorithms for accurate flow-based network traffic classification: Evaluation and comparison. *Performance Evaluation*, 67(6), pp.451–467. Available at: http://dx.doi.org/10.1016/j.peva.2010.01.001.

Sperotto, A. et al., 2010. An overview of IP flow-based intrusion detection. *IEEE Communications Surveys and Tutorials*, 12(3), pp.343–356.

Splunk, 2011. Splunk , Big Data and the Future of Security.

Stolfo, S.J. et al., 2000. Cost-based modeling for fraud and intrusion detection: Results from the JAM project. *Proceedings - DARPA Information Survivability Conference and Exposition, DISCEX 2000*, 2, pp.130–144.

Stolfo, S. et al., 2010. Behavior Profiling of Email. *Intelligence and Security Informatics*, 2665, pp.960–960. Available at: http://www.springerlink.com/content/l0w399c8nn56863g.

Strasburg, C. et al., 2010. Masquerade Detection in Network Environments. *Applications and the Internet (SAINT), 2010 10th IEEE/IPSJ International Symposium on*, pp.38–44. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5598177.

Sun B; Chen, Z.W.R.Y.F., 2006. Towards adaptive anomaly detection in cellular mobile networks. *CCNC 2006. 2006 3rd IEEE Consumer Communications and Networking Conference, 2006.*, 2, pp.666–670. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1593121.

Sun, B. et al., 2004. Mobility-based anomaly detection in cellular mobile networks. *Proceedings of the 2004 ACM workshop on Wireless security - WiSe '04*, p.61. Available at: http://dl.acm.org/citation.cfm?id=1023646.1023658.

Symantec, 2012. Internet Security Threat Report 2011 Trends. , 17(April), pp.1–52.

Symantek, 2012. Managing Security Incidents in the Enterprise.

SYRIS Technology Corp, 2004. About FAR, FRR and EER. *Technical document*, pp.0–4. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Technical+document#9.

Trustwave Holdings, 2013. Trustwave SIEM for Government.

Van Renterghem, P. et al., 2013. Statistical discrimination of steroid profiles in doping control with support vector machines. *Analytica Chimica Acta*, 768(1), pp.41–48. Available at: http://dx.doi.org/10.1016/j.aca.2013.01.003.

Verizon, 2016. 2016 Data Breach Investigations Report. *Verizon Business Journal*, (1), pp.1–65. Available at: http://www.verizonenterprise.com/resources/reports/rp_data-breach-investigations-report-2013_en_xg.pdf.

Waizumi, Y., Sato, Y. & Nemoto, Y., 2012. A Network-Based Anomaly Detection System Based on Three Different Network Traffic Characteristics. *Journal of Communication & Computer*, 9(7), p.805. Available at: http://connection.ebscohost.com/c/articles/79895857/network-based-anomaly-detection-system-based-three-different-network-traffic-characteristics.

Walker, T., 2008. Practical management of malicious insider threat - An enterprise CSIRT perspective. *Information Security Technical Report*, 13(4), pp.225–234. Available at: http://dx.doi.org/10.1016/j.istr.2008.10.013.

Wang, K., Cretu, G. & Stolfo, S.J., 2006. Anomalous Payload-Based Worm Detection and Signature Generation. *Recent Advances in Intrusion Detection*, 3858, pp.227–246. Available at: http://link.springer.com/10.1007/11663812_12.

Wang, K. & Stolfo, S., 2004. Anomalous payload-based network intrusion detection. *Recent Advances in Intrusion Detection*, pp.203–222. Available at: http://link.springer.com/chapter/10.1007/978-3-540-30143-1_11.

Wang, S., 2010. A comprehensive survey of data mining-based accounting-fraud detection research. *2010 International Conference on Intelligent Computation Technology and Automation, ICICTA 2010*, 1, pp.50–53.

Weber, D., Girardi, B. & Martin, D., 2012. TRANSFORMING TRADITIONAL SECURITY STRATEGIES INTO AN EARLY WARNING SYSTEM FOR ADVANCED THREATS- Big Data Propels SIEM into the Era of Security Analytics. *RSA Security Brief*, (September). Available at: http://www.emc.com/collateral/software/solution-overview/h11031-transforming-traditional-security-strategies-so.pdf.

Wood, B., 2000. An insider threat model for adversary simulation. *SRI International Research on Mitigating the Insider Threat to Information Systems*, 2, pp.1–3. Available at: http://www.csl.sri.com/users/bjwood/Insider_threat_model_v02.pdf.

Wüchner, T. & Pretschner, A., 2012. Data loss prevention based on data-driven usage control. *Proceedings - International Symposium on Software Reliability Engineering, ISSRE*, pp.151–160.

Xu, K. et al., 2010. Real-Time Behaviour Profiling for Network Monitoring. *Int. J. Internet Protoc. Technol.*, 5(1/2), pp.65–80. Available at: http://dx.doi.org/10.1504/IJIPT.2010.032616.

Yampolskiy, R. V. & Govindaraju, V., 2008. Behavioural biometrics: a survey and classification. *International Journal of Biometrics*, 1(1), pp.81–113.

Yang, Y., 2010. Web user behavioral profiling for user identification. *Decision Support Systems*, 49(3), pp.261–271.

Yaseen, Q. & Panda, B., 2009. Knowledge Acquisition and Insider Threat Prediction in Relational Database Systems. *2009 International Conference on Computational Science and Engineering*, 3, pp.450–455.

Yazji, S. et al., 2009. Implicit user re-authentication for mobile devices. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5585 LNCS, pp.325–339.

Yazji, S. et al., 2011. Protecting Private Data on Mobile Systems based on Spatio-temporal Analysis.

Young, W.T. et al., 2013. Use of domain knowledge to detect insider threats in computer activities. *Proceedings - IEEE CS Security and Privacy Workshops, SPW 2013*, pp.60–67.

Zanero, S., 2005. Analyzing TCP traffic patterns using Self Organizing Maps. *Image Analysis and Processing - Iciap 2005, Proceedings*, 3617, pp.83–90.

Zeadally, S. et al., 2012. Detecting Insider Threats: Solutions and Trends. *Information Security Journal: A Global Perspective*, 21(4), pp.183–192.

ZOHO Corp, 2007. Analyzing Logs For Security Information Event Management.

Zou,2006. Elements of Statisitics, available at http://www.ams.sunysb.edu/~jasonzou/ams102/notes/notes3.pdf.

Appendix

Appendix A: Published Papers

- Alotibi G, Li F, Clarke N, Furnell SM (2015): **Behavioural Based Features Abstraction from Network Traffic**, 10th International Conference on Cyber warfare and Security, Kruger National Park, South Africa,24-25 March, pp 1-9, ISBN 978-910309-97,2015.

- Li F, Clarke NL, Alotibi G, Joy D (2015): **Forensic Investigation of Network Traffic: A Study into the Derivation of Application-Level features from Network-Level Metadata**, 6th Annual International Conference on ICT: Big data, Could and Security (ICT-BDCS 2015), 27-28 July, ISSN: 2382-5669, pp68-73, 2015.

- Alotibi G, Clarke NL, Li F, Furnell SM (2016): **User profiling from network traffic via novel application-level interactions,** 11th International Conference for Internet Technology and Secured Transactions (ICITST), pp 279-285, Barcelona, Spain, 2016.

Appendix B: Ethical Approval