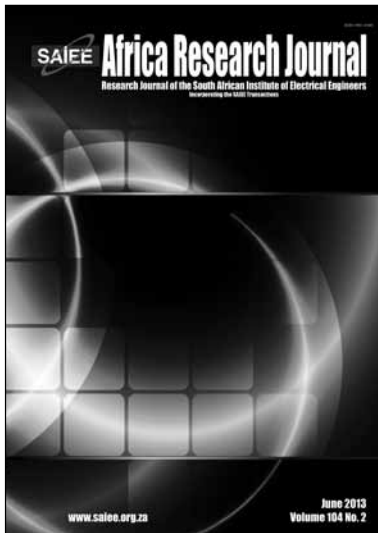


VOL 104 No 2  
June 2013

# SAIEE Africa Research Journal



SAIEE AFRICA RESEARCH JOURNAL EDITORIAL STAFF ..... IFC

A source analysis of the conficker outbreak from a network telescope <i>by B. Irwin</i> .....	38
Towards ensuring scalability, interoperability and efficient access control in a multi-domain grid-based environment <i>by N.A. Azeez and I.M. Venter</i> .....	54
Ignorance to awareness: towards an information security awareness process <i>by T. Gundu and S.V. Flowerday</i> .....	69
Multi-agent augmented computer vision technologies to support human monitoring of secure computing facilities <i>by M. Potgieter and J. Van Niekerk</i> .....	80

## INFORMATION SECURITY SOUTH AFRICA (ISSA) 2012

This special issue of the SAIEE Africa Research Journal is devoted to selected papers from the Information Security South Africa (ISSA) 2012 Conference which was held in Johannesburg, South Africa from 15 to 17 August 2012. The aim of the annual ISSA conference is to provide information security practitioners and researchers, from all over the globe, an opportunity to share their knowledge and research results with their peers. The 2012 conference focused on a wide spectrum of aspects in the information security domain. The fact that issues covering the functional, business, managerial, human, theoretical and technological aspects were addressed emphasizes the wide multi-disciplinary nature of modern-day information security.

With the assistance of the original reviewers, ten conference papers that received good overall reviews were identified. At the conference, I attended the presentation of each of these ten papers and based on the reviewer reports and the presentations I selected six of these papers for possible publication in this Special Edition. The authors of these six selected papers were asked to rework their papers by expanding and/or further formalizing the research conducted. Each of these papers was subsequently reviewed again by a minimum of two international subject specialists. In some cases, where conflicting reviews were received, further reviews were requested. In one case five reviews were requested to enable objective and quality decisions to be made. In all cases the reviews were conducted by members of the Technical Committee (TC) 11 of the International Federation of Information Processing (IFIP) or some subject experts suggested by them. Thus, in all cases the reviews were conducted by reputable subject experts and enough reviews were received to make a confident decision.

In the end four papers were selected to get published in this Special Edition after the reviewer comments were attended to satisfactorily. The four papers cover various aspects of information security. The one paper addresses some technical aspects related to modern malware protection. A second paper focuses on the access control problems experienced in grid-computing systems of today. The third paper attends to some of the human aspects associated with information security and the fourth paper provides an example how computer vision and related systems can be utilized to assist modern security systems. This Special Edition includes four very diverse papers in the discipline of information security, giving a true reflection of the multi-disciplinary nature of this field of study.

Lastly, I would like to express my appreciation to IEEE Xplore, who originally published the ISSA conference papers, for granting permission that these reworked papers can be published in this Special Edition.

Prof Rossouw von Solms  
Guest Editor

## A SOURCE ANALYSIS OF THE CONFICKER OUTBREAK FROM A NETWORK TELESCOPE

B. Irwin

*Security and Networks Research Group, Department of Computer Science, Rhodes University, Grahamstown 6140, South Africa. E-mail: b.irwin@ru.ac.za*

**Abstract:** This paper discusses a dataset of some 16 million packets targeting port 445/tcp collected by a network telescope utilising a /24 netblock in South African IP address space. An initial overview of the collected data is provided. This is followed by a detailed analysis of the packet characteristics observed, including size and TTL. The peculiarities of the observed target selection and the results of the flaw in the Conficker worm's propagation algorithm are presented. An analysis of the 4 million observed source hosts is reported, grouped by both packet counts and the number of distinct hosts per network address block. Address blocks of size /8, 16 and 24 are used for groupings. The localisation, by geographic region and numerical proximity, of high ranking aggregate netblocks is highlighted. The observed shift in geopolitical origins observed during the evolution of the Conficker worm is also discussed. The paper concludes with some overall analyses, and consideration of the application of network telescopes to the monitoring of such outbreaks in the future.

**Keywords:** Conficker, Zotob, malware, network telescope, geolocation.

### 1. INTRODUCTION

This paper explores the application and value of the use of a network telescope [1–3] in the tracking and monitoring of a global malware outbreak over the period from August 2005 to September 2009. During this period, the volume of traffic observed arriving at the research telescope destined for port 445/tcp grew dramatically, particularly over the last 14 months of the dataset, reaching a peak of nearly two orders of magnitude higher than the previously observed traffic baseline. Much of this increase can be attributed to the prevalence of the Conficker worm [4], also known as Kido and DownAdUp [5], and associated exploitation of the vulnerability in the Microsoft RPC/DCOM stack.

Conficker and other associated malware exploits a vulnerability in the Microsoft RPC stack detailed in the Microsoft MS08-067 [6] security bulletin released on 23rd October 2008. The vulnerability exploited is similar to those discovered in July 2003 detailed in MS03-026 [7] and later in MS03-039 [8] and subsequently exploited by the Blaster and Welchia worms in August 2003 [9]. A further vulnerability in the RPC stack was exploited by Sasser in April 2004, taking advantage of the vulnerability disclosed in MS04-011 some seventeen days previously [10]. The problems with the RPC/DCOM stack in Microsoft Windows Family operating systems continued and MS06-40 released in September 2006 [11] patched a further vulnerability that was exploited by various malware, such as Mobox. Given this history of vulnerability, and the widespread adoption of the Windows operating platform, as well as the rapid development of code exploiting these vulnerabilities, researchers were justifiably concerned when the MS08-067 vulnerability was announced. A detailed analysis of the Conficker malware is beyond the scope of this research. For details on the actual origins, and analysis

from a payload and reverse engineering perspective, readers are encouraged to consult in particular the work done by SRI [12] and Symantec [13] on reverse engineering and documenting the initial spread of the malware.

This paper presents a discussion of how the spread of this malware was observed from the perspective of the network telescope system, using data as described in Section 2. An overview of the evolution of the worm is presented along with a time-line of the major points in the evolution of this software in Section 3. This is shown to match fairly accurately with the observed changes in traffic presented in Section 4. An analysis of the observed network traffic is presented in Section 5. A focus of on the traffic distribution across the target addresses is discussed in Section 6. These latter two sections present evidence for the strong likelihood of the majority of the 445/tcp traffic being Conficker related. An analysis of the network sources observed is carried out in Section 7. This is then followed by an analysis of the geopolitical distribution and observed shift in the source origins in Section 8. Considerations of the implication and application of this research, and the network telescope sensor as a malware analysis tool is contained in Section 9. The paper concludes with a reflection on the application of a network telescope to the monitoring of this kind of event, and the views of traffic as observed by other sensors.

### 2. DATA SOURCE

This research was carried out using TCP/IP packet data collected over a 50 month period from August 2005 to September 2009, using a /24 IPv4 network address block, within South African address space. The address space was connected to a passive network telescope [3] operated by the researcher [14]. The above period in

which data was collected includes the period of the outbreak of the Conficker worm. The analysis of traffic relating to this malware outbreak late 2008 is the focus of the remainder of this paper.

### 2.1 Network Telescopes

Network telescopes are a means of monitoring 'Internet Background Radiation' – network packet data which arrives at a destination network unsolicited. This method was popularised by Moore et al [3]. Specifically a network telescope makes use of unallocated IP addresses which are not being used for running services. Based on this any incoming traffic recorded can be viewed as unsolicited, as ideally no traffic would be received, as no clients or servers are operating using these addresses. Care is taken to filter traffic so as to ensure that no response traffic is sent so as to appear to remote hosts as indistinguishable from an unallocated address. A more detailed discussion of the varying modes of operation for network telescopes and related analysis methods can be found in Irwin [14]. Greynets [1, 2] are a related implementation using smaller slices of address space than have traditionally been used for the operation of network telescopes.

What is important to bear in mind when analysing the data collected using the Rhodes University system, is that one of the shortcomings of the current network telescope setup is that only the first packet of the potential TCP 3-way handshake is actually captured. Since the handshake, by design, cannot complete, no data payload can be captured. Due to this limitation it can only be inferred, albeit with a high level of certainty, that the increase in observed traffic is directly related to the Conficker malware. It is believed, based on analysis of the data, that the majority of the recorded connection attempts are automated connections from Conficker, but there is certainly a component which is scanning activity from other sources looking for operational hosts which may also be vulnerable to the MS08-067 issue and subsequently targeted for exploitation.

### 2.2 Data Processing

Data was collected using tcpdump and piped through an analysis framework [14]. Passive operating system fingerprinting was performed using the p0f tool in order to classify the likely origin operating system. This is a passive (in the sense it does not use live interrogation such as implemented by nmap) operating system fingerprinting tool developed by Michal Zalewski. The passive technique allows for it to be used on recorded traffic. The network telescope host itself ensured that no return packets could be sent, and was located outside of the organisational firewall. Geolocation tools were also used to relate a given IP address (or address block) to a country of allocation.

## 3. CONFICKER EVOLUTIONARY TIME-LINE

The evolution of the threat posed by the Conficker can be traced back to the release of the MS08-067 advisory on 23<sup>rd</sup> October 2008 as an emergency, out of sequence, malware patch by Microsoft after exploitation of the vulnerability was observed in the wild. One of the issues to be aware of when analysing Conficker and research around the threats, relates to the two different naming conventions used by Microsoft and the Conficker Working Group (CWG). The former appears to be in more widespread use. These differences are shown Table 1. In this document the Microsoft naming conventions are used. When analysing the traffic, inflexion points can be seen relating to the version changes in the Conficker malware, as seen in Section 4.2. A more detailed timeline of the evolution of Conficker is maintained by the CWG<sup>1</sup>.

Table 1: Conficker Naming

Date	Microsoft	CWG
20 Nov 2008	Conficker.A	Conficker.A
28 Dec 2008	Conficker.B	Conficker.B
20 Feb 2009	Conficker.C	Conficker.B++
4 Mar 2009	Conficker.D	Conficker.C
8 Apr 2009	Conficker.E	-

## 4. TELESCOPE TRAFFIC OBSERVATIONS

Observed network traffic destined to 445/tcp makes for an interesting case study on a number of fronts. Firstly, packets targeting 445/tcp are the single most significant contributor to the total of traffic observed, both in terms of the number of packets and source addresses observed. Secondly, it is used by the Microsoft Windows family of operating systems for RPC/DCOM communications, including file sharing, and is usually enabled on such systems. The popularity of the deployment of these systems makes this an inviting target when vulnerabilities are found, with historically widespread exploitation. Furthermore, this port is generally firewalled by most organisations, and often by home users as well, although usually only for inbound traffic.

### 4.1. Overview

Traffic destined to port 445/tcp as a whole, can be seen to be fairly persistent over the entire duration of the network telescope observation under consideration, having been observed on all but one of the 1 429 days having data (and 98.1% of hourly observations) within the dataset. Over the period 445/tcp was consistently ranked in the top ten targeted ports observed, by both month and year. During the observation period, packet counts for traffic destined to port 445/tcp was the top ranked in 10 of the 17 quarters under study, with its lowest positions being 4th in Q1 2007 and Q4 2008. Figure 1 shows the prevalence of this traffic over the observation period.

<sup>1</sup> <http://www.confickerworkinggroup.org/wiki/pmwiki.php/ANY/Timeline>

Data shown in this figure reflects only that TCP traffic destined to 445/tcp that has the SYN flag set, and hence can be considered 'active' (in terms of this traffic could potentially elicit a valid response from a target). This traffic also accounted for 41.4% of the traffic overall, consisting of some 16 893 920 packets.

The rapid increase in traffic from approximately 1 000 packets/day (ppd) in October 2008 through to nearly 100 000 ppd by the end of September 2009 can be clearly seen. The benefit of the long collection baseline was realised in this comparison, in terms of being able to quantify just how large the increase in traffic was.

The spike in observed activity in the early portion of Figure 1 is probably attributable to the Zotob worm [15] exploiting a vulnerability disclosed on 9th August 2005 in MS05-039 [16, 17], either from the worm itself or related scanning in response to this event by individuals looking for vulnerable hosts. Traffic levels had, however, decreased and became largely normalised by November 2005 and continued to drop through to mid-October 2008. This gradual decrease is likely due to the increased uptake of automated patching of systems through the Microsoft Windows Update mechanism, the release of Service Packs for Windows XP (SP3 - April 2008) and Windows Vista (SP1 - March 2008, SP2 - April 2009) resulting in the remediation of vulnerabilities in the RPC service. More significantly, the lack of any significant vulnerability affecting this protocol during the observation period would have reduced the incidence of scanning. The rapid increase in traffic observed from October 2008 onwards can be attributed to activities surrounding the exploitation of the MS08-067 vulnerability in Microsoft Windows operating systems, most notably by the Conficker worm. The remainder of this paper focuses on activity observed from October 2008 onwards.

#### 4.2. Conficker Related

The Conficker worm was first observed on the 20th /21st November 2008 (depending on time zone), and after almost a year, over 7 million hosts were observed as still infected<sup>2</sup>. After the 20th November, traffic destined to 445/tcp constituted 70% of traffic observed. Over the entire observation period, 4 002 119 unique hosts (86% of the total) were observed sending packets to a destination port of 445/tcp. Of these addresses, 95% were observed after the 20th November 2008, with only 5 544 (0.14%) having been observed prior to the 1st November 2008. Only 0.3% of the IP addresses identified as having targeted 445/tcp had sent any traffic at all prior to the beginning of November 2008. This is indicative of a marked change in observed traffic patterns.

This is not the first work to have been done on looking at Conficker from the perspective of a network telescope

with some detailed analysis having been performed by CAIDA researchers [18], and the subsequent release of a portion of their 3 Days of Conficker dataset [19]. What is novel in this paper is the fine level of detail at which analysis has been performed along with the use of data from a single /24 network telescope with a continuous long term baseline. Previously work [18] on the spread of Conficker relating to network telescopes has made use of the CAIDA /8 telescope, and a further dataset gathered on a /16 netblock.

The discussion in the remainder of this paper primarily focuses on the traffic considered Conficker related from mid-October 2008 through to the end of the dataset at the end of September 2009, in effect covering nearly a year of activity related to the MS08-067 vulnerability, and the spread of the Conficker worm through its evolutionary phases. This long uninterrupted temporal baseline in relation to other research allows for some insight to be gained into the changing behaviour of the malware over an extended period.

An overview of the total traffic observed by the telescope system is shown in Figure 2 as the calculated average number of packets received per IP addresses in the monitored range. A number of distinct spikes in the traffic are noticeable, along with the general increase in traffic over time. The increase is, however, not nearly as rapid as that observable in the latter part of Figure 1. Particularly notable events are the large spike on 28th October 2008, followed by a rapid climb on the 21st November 2008. A second rapid increase can be seen on 1st January 2009, with a consistent increase in traffic rates observed through to mid-February, and a large increase in activity on the 28th February. This is followed by a sharp drop-off mid-March and a small spike prior to 1st April. From this point the traffic continues to increase, other than two dips which were caused by network outages. On initial observation, these periods seem to coincide to those outlined in the evolution timeline in Table 1.

Looking a little deeper, and analysing the composition of the traffic by protocol, one can see that the spikes observed cannot be correlated directly to activity on port 445/tcp. This detailed breakdown of the same dataset and time period as previously shown in Figure 2 can be seen in Figure 3. In the detailed plot, ICMP and UDP traffic have been shown along with the contribution made by traffic destined to 445/tcp on the sensor network. Several points in Figure 3 are worth highlighting:

- Although the spike shown at point A ties in with the release of the MS08-067 security bulletin, it is not related to it, but rather is the result of a burst of classic back-scatter packets originated from a UNIX system located in Jordan.

<sup>2</sup> <http://www.shadowserver.org/wiki/pmwiki.php/Stats/Conficker>

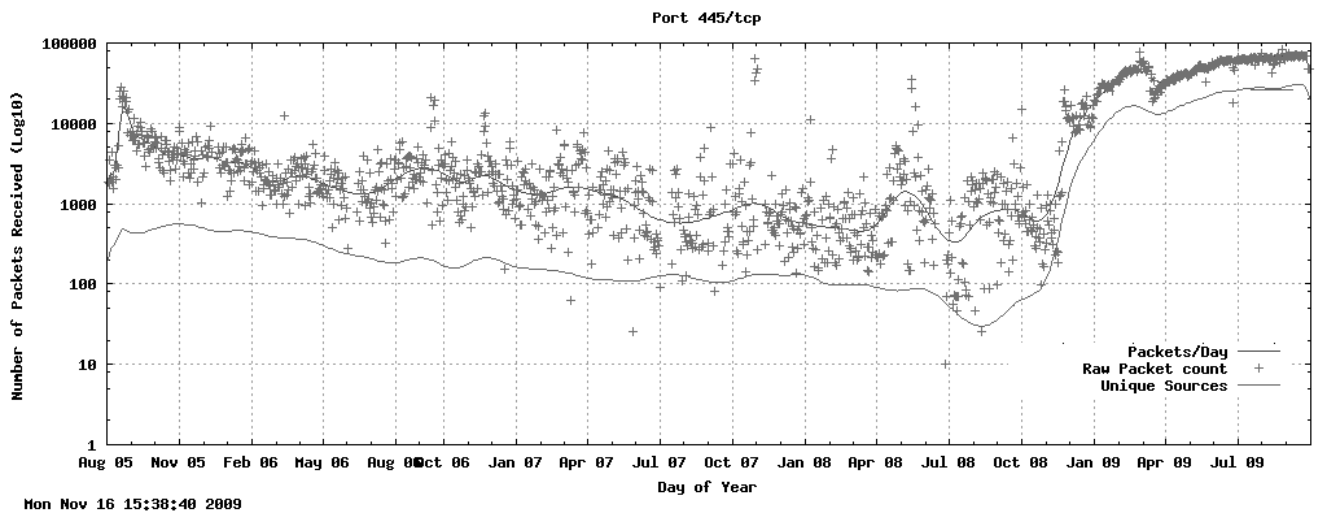


Figure 1: TCP packet received on port 445 by day.

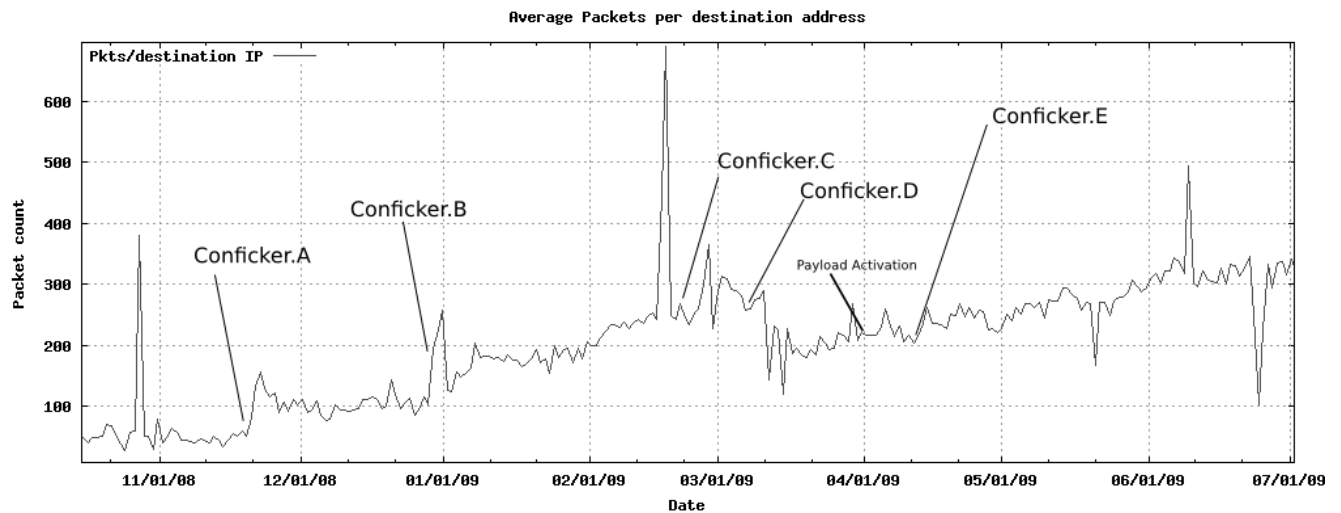


Figure 2: All Traffic November 2008 - July 2009

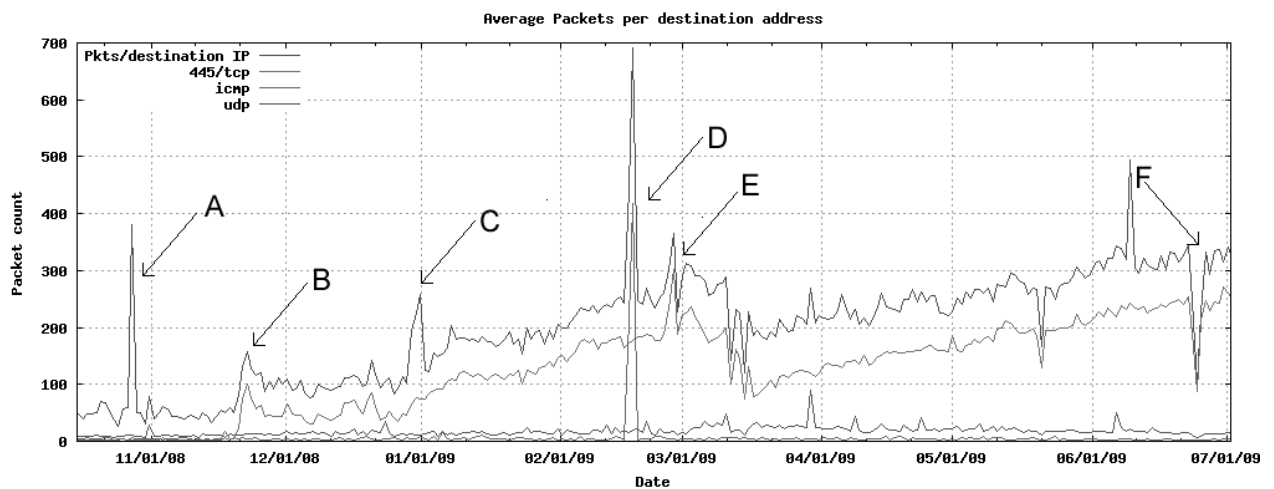


Figure 3: Traffic Protocol Breakdown (November 2008 - July 2009)

- The ‘birth’ of Conficker on the 21 November 2008 (point B) can be seen by the sharp rise in 445/tcp traffic as a portion of the whole. Conficker.B was released on the 28th December 2008 and although there is a spike in total traffic (point C) this was not due to Conficker, but a reflected back-scatter SYN-ACK packets from a web server located in Costa Rica.
- Point D indicates a further anomaly in the traffic pattern, with the spike caused by a flood of 159 000 ICMP TTL expired messages, received from a host in China on the 17th and 18th February 2009.
- The spike in late February 2009 (point E) can be attributed to the release of Conficker.C on the 20th February.
- Point F is worth noting in that the drop in recorded traffic was due to a series of extended network outages adversely affecting Rhodes University’s upstream Internet connection over the period 24th - 26th June 2009.

#### 4.2. Conficker Outbreak

Analysis of the data in the first few days of the Conficker outbreak revealed some interesting trends. The first of these is illustrated in Figure 4, which plots data relating to traffic received on 445/tcp during the period 21st -24th November 2008. Times noted in the Figure are SAST (GMT+2:00). What is immediately noticeable is that while the packets follow a rough circadian rhythm, this trend is even more noticeable when the number of distinct sources for each hour interval is plotted. Similar results were found with the data processed in the CAIDA Conficker report [18]. Considering the 24 hour period from midnight on the 21st November, the number of observed sources per hour can be seen to climb rapidly, from fewer than ten at 05h00 to over 250 by midnight the following day. What is interesting is that there is a large

increase in packets observed around 06h00, yet only twenty source hosts had been observed at this stage. From this point the packet rate per hour dropped dramatically and the host count started to climb — in essence, more hosts were sending relatively fewer datagrams. By 16h00 traffic had reached a low point and subsequently started to increase again, with a rapid growth in the number of sources observed; a high sustained rate being maintained for nearly ten hours before dropping back. This is pattern can be seen to be repeated over the next two days.

An interesting anomaly was found in the data for 445/tcp, with a significant spike in scanning activity detected over a period from late on the 30th September, through to the evening of the 1st October 2008. This increase in scanning was particularly noticeable due to there being almost no traffic targeting 445/tcp in the days leading up to this and very little afterwards. A plot of the relevant traffic is presented in Figure 5. The traffic is also seen to originate from a relatively high number of sources, with 3 476 IP addresses being logged between 03h00 and 18h00 on the 1st October, having sent some 14 808 packets targeting 445/tcp. The top five geopolitical sources as determined by geolocation tools were Brazil (BR), France (FR), the USA (US), Japan (JP) and the Philippines (PH). Together these accounted for nearly half (46%) of the traffic sources observed during this period, a summary of which is shown in Table 2. Brazilian and US hosts achieved an almost complete coverage of the network telescope range. Sources of traffic targeting 445/tcp were observed from 100 countries, although only 30 of these had more than twenty distinct sources. This would indicate a fairly low spread of dispersion of hosts involved in the scanning. This could in turn point to localised networks of previously compromised systems being used.

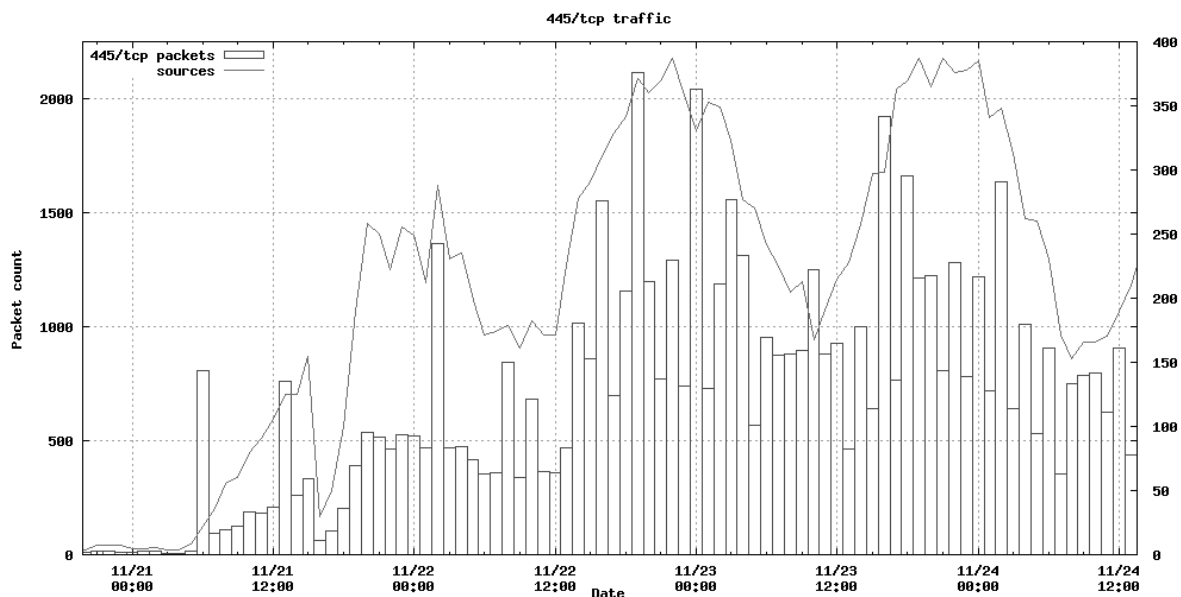


Figure 4: Early days of Conficker (21st - 24th November 2008)

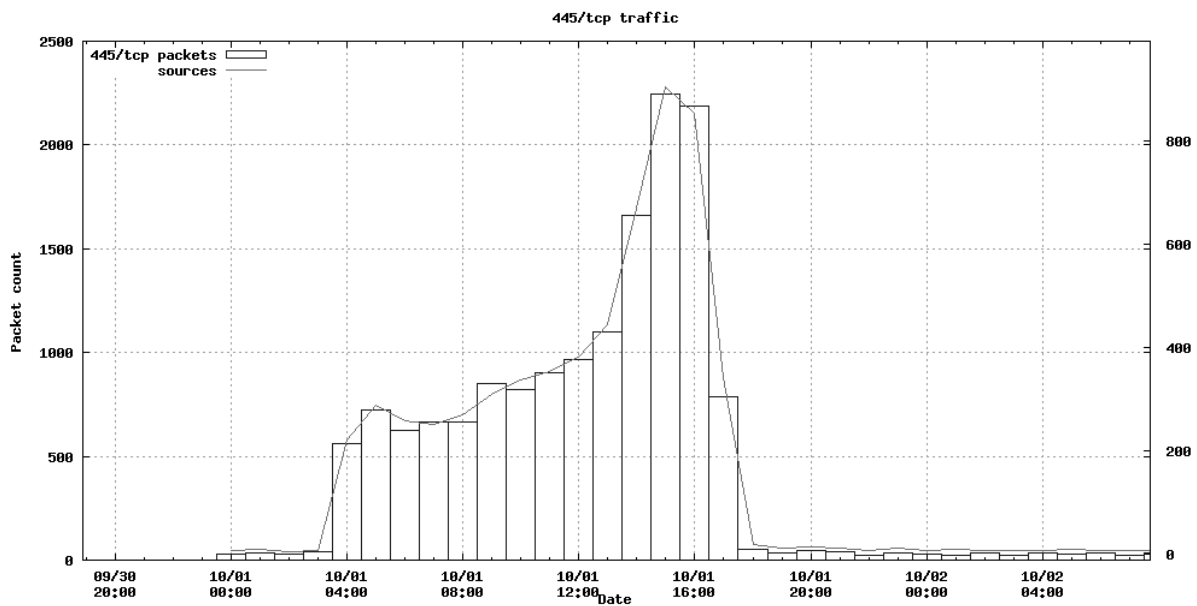


Figure 5: Early reconnaissance (30th September - 2nd October 2008)

Table 2: Top 5 countries (1st October 2009)

Rank	CC	Count <sub>packets</sub>	Count <sub>Source</sub>	Count <sub>Dest</sub>
1	BR	2203	625	249
2	FR	1551	388	231
3	US	1877	279	249
4	JP	654	167	177
5	PH	573	155	157
Total		6858	1614	

Without payloads, it is impossible to determine the exact nature of this traffic, or the means of its generation. While it could plausibly be activity from the Gimmiv trojan [20], it was noted that there were problems with its replication mechanisms. Another possibility is that it could be some custom malware utilising the Chinese exploit kit based on the Gimmiv trojan. It is nonetheless interesting to note that there is a fairly even distribution across the entire monitored IP range, in contrast to what is seen with the actual Conficker malware post release, as explored in Section 6. The majority of the sources (99%) scanned less than ten target addresses, with 1 832 (52%) only probing one host. Only three hosts scanned more than 100 addresses, while the majority of sources sent two packets which were in relatively quick succession, before disappearing. This type of activity would indicate some form of optimised scanning identifying live networks, and possibly using a distributed scan to minimise detection.

It is the researcher's hypothesis that there is a strong likelihood of this having been a distributed scanning attempt, with multiple sources scanning to look for vulnerable hosts possibly for further targeted exploitation, or as a means of seeding initial distribution points for later malware release. The fact that such a high number of hosts only probed a single target points to a well-

coordinated, distributed scan, or to these addresses possibly being used as a decoy scan. The three top hosts (by packet count) were determined to be in Taiwan (TW) and the USA, and are most likely the real hosts. This is supported by the fact that hosts that are geolocated as originating from locales such as the Philippines (PH), Croatia (HR), Turkey (TR) and Austria (AT) all have TTL values above 240. This is highly unlikely given that Rhodes University's Internet connectivity which had, at the time of collection, at least ten hops to get to international peering points in London. Further examination of the TTL values of traffic targeting 445/tcp shows that a significant number of the hosts have the same TTL values despite being geolocated to vastly different parts of the globe, further strengthening the likelihood of packet forgery. The fact that only two packets are sent is also interesting as generally most TCP/IP stacks send three SYN connection attempts before timing out. This could point to the fact that custom code was being used with a short time-out, or that packets were being constructed using 'raw' sockets.

## 5. PACKET DATA

This section evaluates aspects of the packets received on 445/tcp by the network telescope, considering the observed TTL, packet structure, packet retransmission, and source operating system.

### 5.1. Time to Live

An analysis of the TTL values recorded for all incoming traffic destined to 445/tcp showed a very narrow banding where it was observed that the values were, with few exceptions, between 50 and 100. This range covers default TTL settings for both Windows and UNIX platforms, having default base TTL values of 128 and 64 respectively. This banding is further evident when plotted against packet counts for TTL values received overall as

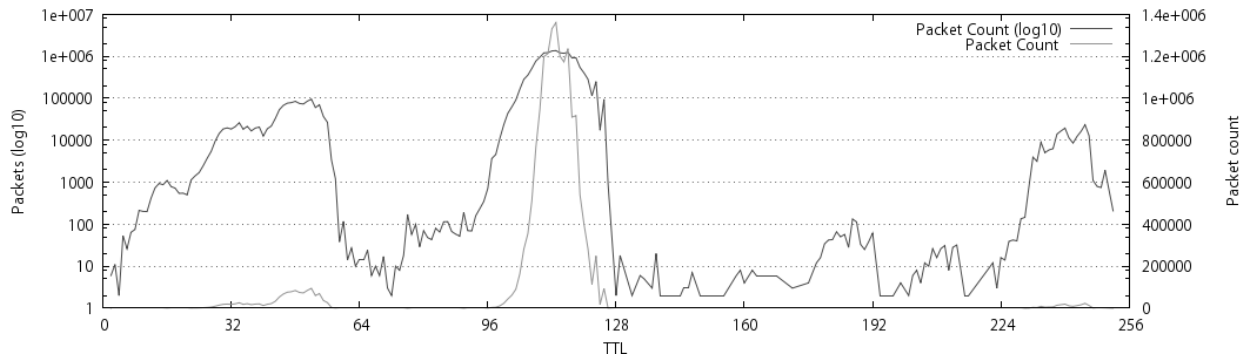


Figure 6: Packet counts by TTL for 445/tcp

presented in Figure 6. In this Figure the TTL values for 445/tcp packets can be seen to be largely grouped in the 96-128 range, with very few packets recorded in the 32-64 and 224-255 ranges.

This provides strong evidence towards Microsoft Windows platforms being the origin of the packets. This was confirmed by the passive operating system fingerprinting that was performed (Section 5.4). This would in turn lend weight to the supposition that when considering the number of distinct sources, the packets observed were actually generated by the automated scanning modes of the Conficker worm, rather than other automated tools.

### 5.2. Packet Structure

Table 3: Size of packets targeting 445/tcp

Size	Count <sub>packet</sub>	Packet %	Count <sub>Source</sub>	Source %
62	12 207 669	86.18	3 248 492	85.16
66	1 281 603	9.04	443 325	11.62
78	423 043	2.98	153 076	4.01
60	150 549	1.06	25 211	0.66
74	100 043	0.70	31 278	0.81
Total		99.96		102.26

$$N_{\text{Packets}} = 14\,165\,097 \quad N_{\text{Sources}} = 3\,814\,447$$

Source % total is >100% as hosts may have sent different size packets

Based on an analysis of the datagrams received, the majority were found to be of 62 bytes in size. A summary of packet sizes is given in Table 3. In this Table, it can be seen that more than 85% of both the hosts and packets match this sizing. This is reached by having a TCP packet with no data and encapsulated inside an IP and finally an Ethernet datagram. In order to reach the value of 62 bytes, rather than the default of 60, TCP options are set. The combination found set most often was “MSS:1460, NOP NOP TCP SACK:True”, which accounted for 8 bytes. These options enabled the Selective Acknowledgement (SACK) on the connection being established, along with a maximum sent size of 1 460 bytes. These settings were found to match captured Conficker propagation traffic. Thus there is a fairly high

probability that the TCP SYN packets being sent to addresses on the telescope actually originated from the Conficker malware.

This provides an example of how, despite being somewhat handicapped by the lack of payloads in a network telescope dataset, comparative data from honeynet or other systems with higher levels of interaction can be used to augment the analysis process. While absolute certainty is not possible without payload analysis, researchers can attain a high level of confidence in their analyses.

### 5.3. Transmission

A further interesting characteristic observed in the 445/tcp traffic after the advent of Conficker, is that it has a very noticeable signature ‘on the wire’ in terms of the way connection attempts are made. Most operating system TCP/IP stacks will send at least three TCP SYN packets in an attempt to establish a connection. By contrast, in the majority of the 445/tcp traffic received after the 20th November 2008, one observes only two connection attempts. An example of this is shown in Figure 7, where source addresses (in the 4th column) make two connection attempts approximately three seconds apart. Similar behaviour has been observed by Aben [18]. This was further validated by the researcher with captures of live propagation traffic obtained from hosts with confirmed Conficker infections.

Considering the total number of sources observed and the total number of packets targeting 445/tcp after 20th November, these are in a ratio of approximately 1:4, indicating that on average most sources scanned two hosts at most. This is discussed further and an analysis of the target addressing is provided in Section 6.

### 5.4. Operating System Fingerprinting

Operating System attribution was performed using p0f. While it is recognised that this method is not flawless, and may be skewed by the use of NAT and dynamic address allocation, it nevertheless provides a useful measure. The results are presented in Table 4. Microsoft



```

1 01:16:21.685360 IP 190.50.x.80.2725 > 196.x.y.3.445: S 2062323770:2062323770(0)
2 01:16:23.509228 IP 77.28.x.55.4853 > 196.x.y.34.445: S 1323192692:1323192692(0)
3 01:16:24.677814 IP 190.50.x.80.2725 > 196.x.y.3.445: S 2062323770:2062323770(0)
4 01:16:26.514630 IP 77.28.x.55.4853 > 196.x.y.34.445: S 1323192692:1323192692(0)
5 01:16:27.693010 IP 79.0.x.248.1731 > 196.x.y.18.445: S 1786561877:1786561877(0)
6 01:16:29.808481 IP 189.101.x.133.2499 > 196.x.y.3.445: S 3114908412:3114908412(0)
7 01:16:30.696890 IP 79.0.x.248.1731 > 196.x.y.18.445: S 1786561877:1786561877(0)
8 01:16:32.751635 IP 189.101.x.133.2499 > 196.x.y.3.445: S 3114908412:3114908412(0)

```

Figure 7: Example of the two packet repetitions

Windows family of operating system platforms accounted for 99.7% of the sources that could be attributed, which is unsurprising given that the Conficker malware targets these platforms, and the TTL data as seen in Section 5.1.

Table 4: Traffic to 445/tcp by attributed Operating System

Rank	Protocol	Number	%
1	Windows	16336052	99.671
2	Proxyblocker	19401	0.118
3	MacOS	10066	0.062
4	FreeBSD	7114	0.043
5	Linux	4361	0.026
6	NetBSD	3981	0.024
7	Cisco	3230	0.019
8	Solaris	1910	0.011
9	Checkpoint	1343	0.008
10	NMAP	1258	0.007
			99.992

N=16 389 887 % of Packets attributable to an IP address with an identified OS N is calculated as packets received after 2008-11-20 00:00 GMT+2

What is surprising is that machines are being observed as infected, despite the patch having been out since 23rd October 2008, and the Microsoft malware removal tool having had functionality to clean the Conficker malware off a system since January 2009. The removal tool is automatically run by the patch update procedure that occurs as part of Microsoft's monthly 'Patch Tuesday' patch cycle. One of the actions taken by the Conficker malware as a means of self-preservation is to disable the automatic update and patching mechanism provided by Microsoft operating systems. What can be concluded from this is that these machines remain infected due to one of the following reasons:

**User unaware** — The user is unaware of the infection.

**User unable to update** — The user is unable to update because of the defence mechanisms put in place by the malware.

**System unable to update** — The user is unable to apply updates due to the system

platform likely being pirated and therefore unable to update<sup>3</sup>.

## 6. TARGET ANALYSIS

Changing focus away from the sources of the traffic to the addresses being targeted in the network telescope address space, a very uneven distribution pattern is observed. The lower half of the monitored space, i.e. 196.x.x.0/25, is targeted substantially more than the upper half (196.x.x.128/25). Particularly heavily targeted is 196.x.x.1, closely followed by other addresses in the lower 16. The first eight addresses in the address block all received more than 100 000 distinct sources. This bias is shown in Figure 8, which considers the number of distinct sources rather than packets observed for each IP address in the monitored range.

The strong bias towards the lower portion of the address space can be seen clearly. Notably, the last address in the monitored range (196.x.x.255) received a much higher coverage than other IPs in the upper /25 portion. The reason for this bias is most likely a naive scanning optimisation, which attempts to probe one or more addresses on a range and, if no response is received, moves onto another range. The probes to the last address on the range may serve a similar purpose. By convention default gateways on most IP networks either make use of the first or last address in a given subnet.

The packet counts received for each address, presents a markedly different picture, as shown in Figure 9. A significant change in the levels can be seen occurring at 196.x.x.128, with a drop in recorded traffic of nearly two orders of magnitude. Figure 9 contains the number of sources, and the packet count observed for each address. A third value is plotted on the right y-axis, being a ratio of the number of packet to distinct sources. This value also makes a change as the target address moves into the upper half of the range, doubling from just over 2 to 4. The dip at the beginning of the graph is due to some traffic having been recorded to 196.x.x.0, which is not normally targeted as it would be regarded as the 'network address', rather than a host address.

<sup>3</sup> Interestingly this is a policy that Microsoft changed with the advent of the Windows 7 platform. Even pirated copies of this OS will be able to receive security updates.

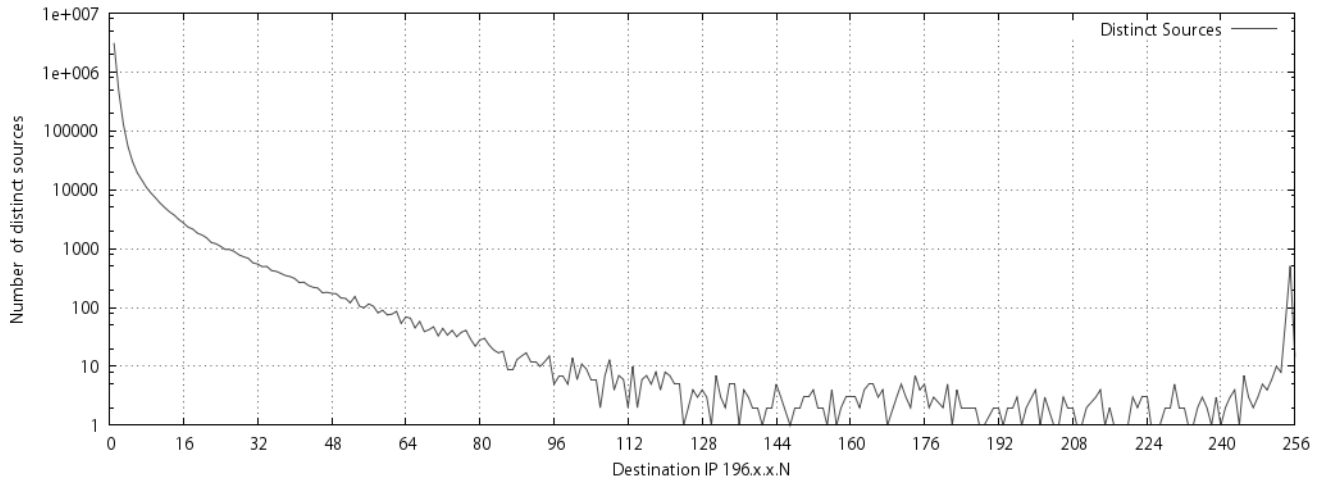


Figure 8: 445/TCP traffic - distinct sources per sensor IP

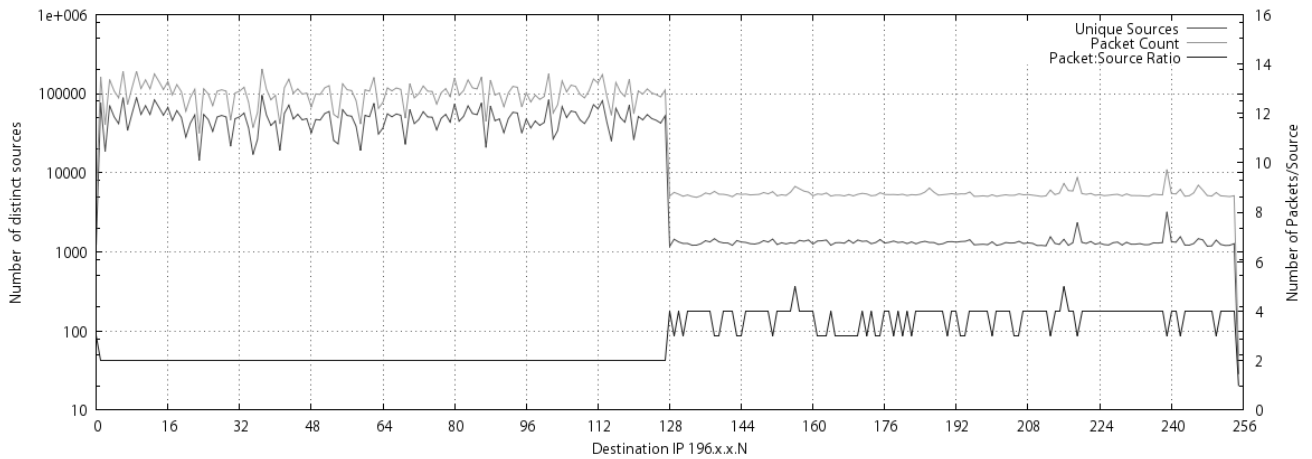


Figure 9: 445/TCP total traffic received per IP in the sensor network

Table 5: Top 445/tcp origins by packet count

Rank	/8	count	%	/16	count	%	/24	count	%	/32	count	%
1	196.0.0.0	1 118 121	7.921	196.21.0.0	287 270	2.035	196.21.218.0	266 790	1.890	196.21.218.x	258 615	1.832
2	189.0.0.0	600 718	4.255	196.20.0.0	234 092	1.658	196.20.164.0	22 734	0.161	196.20.17.x	21 611	0.153
3	190.0.0.0	586 827	4.157	78.106.0.0	91 845	0.650	196.20.17.0	21 655	0.153	196.14.169.x	21 576	0.152
4	92.0.0.0	565 721	4.007	93.80.0.0	91 248	0.646	196.14.169.0	21 576	0.152	196.20.13.x	15 992	0.113
5	95.0.0.0	539 401	3.821	93.81.0.0	83 169	0.589	196.20.165.0	20 762	0.147	196.21.125.x	8 848	0.062
6	89.0.0.0	494 255	3.501	95.28.0.0	76 423	0.541	196.38.187.0	20 077	0.142	196.21.218.x	8 175	0.057
7	78.0.0.0	463 784	3.285	89.178.0.0	75 996	0.538	196.20.167.0	18 283	0.129	59.162.166.x	6 549	0.046
8	79.0.0.0	387 708	2.746	95.24.0.0	68 342	0.484	196.20.166.0	17 758	0.125	196.32.152.x	6 302	0.044
9	93.0.0.0	367 343	2.602	190.51.0.0	54 041	0.382	196.20.13.0	15 992	0.113	196.15.239.x	5 958	0.042
10	201.0.0.0	353 829	2.506	196.205.0.0	53 587	0.379	196.20.140.0	14 668	0.103	196.34.217.x	5 841	0.041
			38.801				7.902				3.115	2.542

$$N_{\text{Packets}} = 14\,115\,791$$

This observed bias towards the lower 128 hosts in the network is due to a bug in the scanning algorithm implemented by Conficker [21, 22]. Due to the way the pseudo random number generator works, a 15-bit value is generated and is then used in a 16-bit field, resulting in the most significant bit of the 2nd and 4th octets (bytes) in an IPv4 address to being zero; in effect limiting these values to the range 0-127. The side effect of this is that it significantly reduces the portion of the Internet that can possibly be scanned [18, 23], limiting it to only 25% of total address space — although it will attempt to scan all the /8 blocks and half of all /16 blocks. Taking this into account, one can use the traffic to the upper 128 address to quantify the other scanning activity present on the port that is not attributable to the Conficker worm propagation, but rather other sources.

Looking at the sources of the traffic directed to the lower half of the monitored range, 26% (1 049 562) had a 2nd octet greater than 127, and 46% (1 842 784) with the last octet having a value greater than 127. Common to both these groupings were 515 834 hosts. What is interesting is that these could not have been infected by direct scanning, although given dynamic addressing they could have had other addresses previously, or have changed networks (as is common with mobile systems).

Infection is still possible though other measures such as via the Windows 'autorun' mechanism on removable media, which was used by the earlier Conficker variants. These sources are analysed further in Section 7.

It is worth noting that the differences in traffic observed between the upper and lower ranges are substantially more than the three times differential described by Wustrow et al [23] and attributed to this activity. This bug most likely accounts for the fact that the second telescope operated by the researcher recorded minimal traffic destined to 445/tcp, since both its 2nd and 4th octets are greater than 127. In addition to this bug in the way random IPs are generated, IP ranges known to belonging to security research firms, and Reserved IPv4 Address space are also avoided by the malware. This is an interesting case which shows the value in having distributed IP address space for monitoring global trends. It is also an argument for chunks of contiguous space rather than smaller fragmented blocks as advocated with the use of greynets [1].

## 7. SOURCE ANALYSIS

Analysing the source address data for the Conficker provides an interesting insight, especially when comparing the top sources ranked by packet count and host count. These are presented in Tables 5 and 6 respectively, and discussed in the sections following. The value in discriminating between these two means of ranking data is that the packet count quantified the volume, and to some extent the persistence of the infection. In contrast, ranking and subsequent analysis by the number of distinct sources (hosts) observed provides

a means for assessing how widespread the infection and related scanning activity was.

### 7.1. Packet Count

Considering the rank order of /8 netblocks, 196.0.0.0/8, managed by AfriNIC, can be seen to have the highest packet count by a significant margin. It is worth taking note that a single host within this block accounted for 23% of the packet count in this netblock. This host was located numerically close to the network telescope address block, and provided a sustained source of traffic over an extended period. The top ten /8 blocks accounted for over 38% of the total packet count can be observed. The numerical sequencing, of the top ten /8 netblocks have very close numerical groupings. The allocations within 189.0.0.0/8, 190.0.0.0/8 and 201.0.0.0/8 are controlled by LACNIC, with the remainder of the top ten being under the control of RIPE. Three adjacent groups of netblocks are observed. This is most likely due to the scanning and propagation mechanisms used by the Conficker malware.

This numerical closeness is also evident when considering the /16 rankings, with two contiguous address blocks in positions one and two. This can also be seen with the two blocks in 93.80.0.0/15. In the /24 rankings four contiguous blocks can be observed in 196.20.164.0/24, contributing to the high rankings of both the 196.20.0.0/16 and subsequently 196.0.0.0/8 netblocks. However, even combined, these four blocks comprising 1 024 addresses account for less than a third of the volume the top observed host. The traffic attributable to individual hosts shows a rather dramatic decrease, with the top ranked host of 196.21.218.x accounting for more than two and half-times the sum of the remainder of the top ten hosts. A second host from the same /24 netblock appears in 6th position.

A comparison of the percentage contribution for each of the netblock aggregation levels decreases sharply from the level covered by the /8 netblocks. However there is relatively little difference between the /24 and /32 levels, largely due to the contribution by the hosts in 196.21.218.0/24. It is still significant that the top host rankings accounted for 2.5% of the total, despite there being over 3.8 million hosts observed, even more so the percentage of traffic attributable to the top ranked hosts. The nearly 80% decrease in composition moving from /8 to /16 is an indicator of how widely spread the scanning activity was, although the proportion covered (7.9%) is less than the 12% observed in the overall dataset, with the values for /24 and host level proportions being much closer.

### 7.2. Source Count

Rankings of the number of distinct sources observed as origins for traffic destined to 445/tcp, aggregated by network block is presented in Table 6. This presents a somewhat different picture to that discussed in the

Table 6: Top origin netblocks for 445/tcp by source count

Rank	/8	Sources	%	/16	Sources	%	%	/24	Sources	%
1	189.0.0.0	209 921	5.511	78.106.0.0	30 407	0.798	46.397	196.1.232.0	233	91.015
2	92.0.0.0	204 970	5.381	93.80.0.0	29 643	0.778	45.231	79.114.134.0	229	89.453
3	190.0.0.0	191 752	5.034	93.81.0.0	25 766	0.676	39.315	79.114.135.0	226	88.281
4	95.0.0.0	190 630	5.004	95.28.0.0	25 703	0.674	39.219	89.179.78.0	224	87.500
5	79.0.0.0	151 942	3.988	89.178.0.0	25 677	0.674	39.179	89.179.104.0	222	86.718
6	78.0.0.0	143 154	3.758	95.24.0.0	23 763	0.623	36.259	89.179.105.0	215	83.984
7	201.0.0.0	113 530	2.980	190.51.0.0	19 982	0.524	30.490	89.179.106.0	213	83.203
8	89.0.0.0	113 106	2.969	77.28.0.0	17 173	0.450	26.203	89.179.107.0	211	82.421
9	59.0.0.0	111 005	2.914	59.93.0.0	15 446	0.405	23.568	93.81.128.0	209	81.640
10	87.0.0.0	110 858	2.910	77.29.0.0	14 024	0.368	21.398	93.81.132.0	209	81.640
			40.449	Total			5.970			

$$N_{Source/P} = 3809104 \quad N_8 = 194 \quad N_{/16} = 14896 \quad N_{/24} = 495602$$

previous section. This ranking provides an indication of how widespread the activity was within a particular netblock, and can be used as a means of determining infection rates. These are not absolute values, and may well be influenced by the use of NAT and dynamic addressing of endpoints or even a combination of these—as is particularly common in modern broadband networks.

At the /8 aggregation level, the first eight ranked netblocks also occur in the top ten by packet count, with 196.0.0.0/8 and 93.0.0.0/8 being omitted. These placed 54th (with 14 914 sources) and 72nd (4 268 sources) respectively. The top four aggregated blocks had a fairly equitable portion of the total hosts served, at just over 5% each. The top ten accounted for over 40% of the total hosts. Again sequential netblocks are seen to occur indicating a strongly likelihood of propagation activities favouring ‘near’ or ‘local’ addresses, although this could have been influenced by some of the other propagation mechanisms used. This sequential locality is repeated again with the /16 aggregation. A strong geopolitical bias can be seen at this level, with five of the top 6 ranks netblocks being under the control of Corbina Telecom (AS8402), head-quartered in Moscow, Russia.

Finally, when considering the top ten /24 netblocks, the top source network (196.1.232.0/24) originates from the Sudan (SD) operated by SudaTel (AS15706). This is significant as 233 individual IP addresses have been observed out of a possible 254 operable host addresses, providing a coverage factor of over 90%. Further sequential address blocks can be seen with 79.114.134.0/23, and the aggregate block of 89.179.104.0/22. Other /24 blocks in the top also exhibit both numerical and topological (being part of the same ASN in global routing tables) closeness. Of interest is that despite 89.179.0.0/16 containing five of the top ten /24 netblocks, it is not in the top ten when aggregation is performed by /16, appearing only in 39th place in these rankings.

The final aspect to be considered regarding sources related to the construction of the observed addresses. This is of interest given the flawed random number generation issue previously discussed in Section 6. Evaluating the data presented in Tables 5 and 6, one can observed that the majority of the /16 networks had a 2nd octet with a value of less than 128. Regarding /24 networks, only the 89.179.104.0/22 as this octet greater than 127. Considering the individual hosts in the top ten, six have a 4th (least significant) octet value > 127. Further analysis of the top 100 000 IP addresses, showed 28% had a second octet greater than 127, with 40% having a 4th octet meeting this criteria.

One explanation for the observed behaviour is that the flaw only affects the random IP selection and scanning phase of Conficker’s propagation. It also implements localised scanning on infected hosts, scanning those addresses nearby numerically. Thus given that IP addresses may be sparsely allocated when one considers the last octet, the influence of the second octet may be seen to be more significant. Considering the total sources identified sending traffic to 445/tcp after the advent of the Conficker worm, just over a quarter (27.5%) of the hosts originated from networks with the second octet greater than 127. Such sources could be considered to be under represented, given an even statistical distribution. These hosts accounted for 3 920 612 packets representing 27.77% of the total traffic directed at 445/tcp, again a somewhat lower contribution than would be expected. Repeating this analysis, for those having both the 2nd and 4th octets greater than 127, resulted in coverage of 27.99% of the traffic, again highlighting the significance of the 2nd octet.

## 8. GEOPOLITICAL ANALYSIS

The last aspect of the case study of Conficker related traffic, is that of the geopolitical origins of the traffic, and particularly how these changed over time. The origins have to some extent been touched on in Section 7., but rather than focusing on specific network address blocks and IP addresses, this section takes an aggregate view at a

country level<sup>4</sup>. This approach is taken in order to provide a bigger picture of the spread and evolution of the worm on a global scale. Country codes as assigned in ISO 3166 are used in tables for brevity.

### 8.1 Pre-Conficker

Prior to the advent of Conficker in November 2008, some 2.7 million datagrams were observed targeting 445/tcp, from 198 thousand source addresses. The geopolitical origins of this traffic are shown in Table 7, which provides a top ten ranking of countries by the number of sources observed, and the number of packets received. Notable in both rankings is the high proportion of the whole with in excess of 70% of both IP addresses and packets being covered. This was observed to change dramatically following the publication of the MS08-067 vulnerability, and subsequent widespread exploitation.

Table 7: Top Sources for 445/TCP - pre Conficker

CC	<i>SrcIP<sub>count</sub></i>	<i>SrcIP</i>	Rank	CC	<i>Packet<sub>count</sub></i>	<i>Packet<sub>%</sub></i>
TW	40 072	20.187	1	ZA	981 349	35.326
US	39 728	20.014	2	US	275 276	9.909
JP	12 772	6.434	3	EG	251 975	9.070
FR	12 436	6.264	4	TW	185 946	6.694
DE	11 574	5.830	5	MU	82 372	2.965
CN	7 307	3.681	6	JP	78 462	2.824
GB	5 902	2.973	7	CN	68 403	2.462
PL	5 449	2.745	8	DE	60 138	2.165
EG	5 217	2.628	9	FR	52 297	1.882
IT	3 492	1.759	10	NG	39 578	1.425
		72.517	Total			74.724

$$N_{SourceIP} = 198503 \quad N_{Packets} = 2777950 \quad N_{Countries} = 190 \\ Avg_{Packets/Src} = 13.994$$

Comparing the two rankings, seven countries are common to both of the top ten rankings. The United Kingdom (GB), Poland (PL) and Italy (IT), while having fairly significant host counts ranked in positions 14, 17 and 19 when packet counts were considered, significantly still within the top twenty. Conversely, South Africa (ZA), while the top ranked source by packet count, was only placed 16th by the source ranking, with 2 592 hosts observed. This serves as an example, where a relatively small number of hosts were responsible for a fairly significant volume of traffic – on average these host sent 378 packets each, or 1.479 packets per address monitored. This rate can be seen to drop off significantly, and should be compared to similar rates after the advent of Conficker.

Consideration needs to be given to the potential bias in the observation towards ZA IP address space, due to the placement of the network telescope sensor. A significant

portion of the observed traffic originating from ZA was from hosts numerically close to the netblock being monitored. This localised numerical bias is often due to poorly implemented scanning algorithms in malware. Several hosts with high packet counts were located at higher education institutions within South Africa. This bias was noted, and was deemed to be less significant as traffic volumes increased post the Conficker outbreak.

### 8.2 Post-Conficker

The situation observed after the widespread exploitation of the MS08-067 vulnerability in Microsoft Windows family systems changed dramatically. The top ten ranked attributable countries of origin by source and packet count are presented in Table 8. While the relative percentage coverage for top ten rankings both by sources and count decreased, the impact on the latter was most notable. Although only a 55% representation of total packet count was achieved by the top ten source countries, this was far more evenly distributed than that given in Table 7.

Table 8: Top Sources for 445/TCP - post Conficker

cc	SrcIP <sub>count</sub>	SrcIP	Rank	cc	Packet <sub>count</sub>	Packet%
RU	577 261	15.154	1	RU	1 791 475	12.691
BR	357 982	9.398	2	BR	1 062 521	7.527
IT	267 660	7.027	3	US	802 232	5.683
CN	265 809	6.978	4	TW	706 241	5.003
TW	233 693	6.135	5	IT	669 286	4.741
DE	215 123	5.647	6	CN	633 169	4.485
AR	181 141	4.755	7	DE	592 773	4.199
IN	149 420	3.922	8	ZA	552 573	3.914
JP	119 299	3.131	9	RO	548 204	3.883
RO	109 987	2.887	10	AR	519 605	3.681
		65.038	Total			55.810

$$N_{SourceIP} = 3809104 \quad N_{Packets} = 14115791 \quad N_{Countries} = 214 \quad Avg_{Packets/Src} = 3.706$$

Comparing the top sources as ranked, shows that although India and Japan had a significant number of sources identified, they ranked much further down when looking at their contribution to the total traffic, ranking respectively in positions: 13 and 14. The remainder of the top ten by number of sources are all present in the top ten by packet count, joined by US, ZA. These two countries ranked 14th and 60th, by the number of identified sources. Hosts from the Russian Federation (RU) and Brazil (BR) maintained their first and second placings in both rankings.

Looking at the average traffic contribution by host for these countries, Russia and Brazil achieved values of 3.1 and 2.96 respectively. This is significantly lower than the average packet contributions observed before the Conficker worm. This measure, while crude, can be used to determine the level of activity. The values themselves would be expected to tend towards 2.0 where only two packets were observed before a host stopped, as discussed

<sup>4</sup> Readers can explore the individual attributions of the netblocks described previously using several tools such as whois and online geolocation services.

in Section 5. The observed values being higher than this can be attributed to a relatively small number of hosts which scanned the entire monitored range, often on a number of occasions.

These rankings can be broken down and evaluated smaller time frames. This is explored in the next section which provides an overview of the dynamics of the spread of the Conficker malware by geographic region.

### 8.3 Conficker Evolution

The final geopolitical analysis of the Conficker Traffic, considers the changing composition of the traffic as the malware evolved. Table 9 shows the lifetime of the Conficker worm segmented into five phases designated A to E. These phases correlate to the five worm variants that were identified (using the Microsoft naming scheme). While it is acknowledged that the actual traffic observed may have originated from multiple variants of the malware, the dates on which new variants were identified serve as a useful means by which to segment the traffic. The dated bounding the periods, as well as the percentage contribution (by packets count and host) to the total traffic directed to 445/tcp after the advent of Conficker is given at the bottom of the table. The first consideration is that of the ranking by packet count as shown in the upper portion of Table 9. The increasing prevalence of the malware on a global scale can be seen by the dilution shown in the overall percentage of traffic covered by the top ten countries in each of the periods, from a high of 70% in A to 57% in D and E. The low of 53.79% in period B would appear to be anomalous.

Looking at the changing composition of top geopolitical sources, countries consistently appearing in all periods are Taiwan (TW), the United States of America (US), China (CN) and Brazil (BR), which are highlighted in bold font in Table 9. The absence of Russia (RU) from the top rankings in period A may be due to a condition in the original variant of the malware that checked for Cyrillic keyboard types, resulting in Russia appearing in 14th and the Ukraine in 42nd place, leading many researchers to initially believe that these countries were the possible origin of the worm. South Africa ranked highly in all but the last period, where the dilution due to the in excess of 10 million packets received, resulted in a rank of 16. For each of the periods, with the exception of B, the top three ranks represent a significant portion of the traffic.

Re-evaluating the same dataset, but ranking countries by the number of hosts observed provides a slightly different picture (as seen in Sections 8.1 and 8.2). Five countries were found to constantly maintain a top ten ranking across the five periods. These were Argentina (AR), Taiwan (TW), China (CN), Brazil (BR) and Russia (RU), and have been boldfaced in the lower half of Table 9. One possibility for these countries having a persistent ranking is that they are regarded as having fairly high

incidence of software piracy<sup>5</sup>. A general dilution of the percentage of hosts covered by the top ten can be seen, and is similar to that observed with packet counts. The proportion of hosts covered drops from 74% in the starting period to only 61% in period B, this then increased up to 66% by period E. Interestingly the Ukraine (UA) makes an appearance in period B, possibly as a flood of hosts were infected with Conficker.B which removed the restriction on not infecting host with Cyrillic, and particularly Ukrainian keyboard settings.

## 9. SIGNIFICANCE AND FUTURE WORK

The dataset used in the research can still be further analysed, particularly from the point of an extended temporal or geopolitical analysis, such as that performed in [14]. Since the Conficker outbreak, there has not been another significant Internet scale event similar to this. As such, further exploration of the dataset on which this work is based, and other subsequently collected datasets may provide better insight into the spread of malware and related malicious activity on a global scale, as well as how to better monitor and defend against these threats.

The information produced by a network telescope can be used in conjunction with existing network security technologies to allow for a means of shunning or otherwise managing potentially hostile hosts, and protecting clients inside a network. This could be achieved through a variety of means, as appropriate for an organisation, ranging from route blackholing to blacklist population.

### 9.1 Significance

The research presented here is significant in relation to existing analysis work in that it has been able to verify findings and observations made by others using much larger network telescopes. The marked difference in traffic targeting the two halves of the monitored network range is important for considering the diversity of placement of network telescope sensors in the future – particularly those with relatively small address space being utilised, something likely to become more prevalent as IPv4 address space becomes more scarce. The work done relating to geopolitical changes observed over the evolution of the malware is also of interest.

## 10. CONCLUSION

This focused analysis of traffic destined to 445/tcp has covered two distinct global malware threats — that of Zotob in August 2005, and Conficker in November 2008. In the intervening period traffic levels remained consistent, and can be attributed to remnants of the Zotob malware and other similar software, and scanning by individuals for hosts having services on 445/tcp exposed to the Internet at large in order to potentially exploit their

<sup>5</sup> <http://portal.bsa.org/idcglobalstudy2007/>

Table 9: Changing Geopolitical sources by Evolutionary Phase

Period				
A	B	C	D	E

## Packet Count

	cc	%	cc	%	cc	%	cc	%	cc	%
1	MU	18.53	RU	9.50	ZA	16.84	RU	13.03	RU	13.84
2	<b>TW</b>	11.67	<b>US</b>	6.00	RU	10.31	ZA	9.77	<b>BR</b>	8.51
3	AR	10.40	<b>BR</b>	5.53	KR	6.07	<b>BR</b>	5.74	<b>US</b>	5.69
4	ZA	6.94	<b>TW</b>	5.51	<b>US</b>	4.62	<b>US</b>	5.39	IT	4.92
5	<b>US</b>	6.32	ZA	5.08	<b>BR</b>	4.51	KR	4.64	<b>TW</b>	4.86
6	<b>CN</b>	5.64	IT	5.02	<b>CN</b>	4.27	IT	4.50	DE	4.44
7	ES	3.16	<b>CN</b>	4.97	IT	4.02	<b>CN</b>	4.03	<b>CN</b>	4.41
8	CL	2.85	KR	4.56	<b>TW</b>	3.60	<b>TW</b>	3.46	RO	4.28
9	CO	2.71	DE	4.29	EG	3.53	RO	3.44	AR	3.59
10	<b>BR</b>	2.50	AR	3.33	DE	3.52	DE	3.27	IN	3.34
		70.72		53.79		61.29		57.27		57.88

## Source Count

	cc	%	cc	%	cc	%	cc	%	cc	%
1	<b>AR</b>	19.57	<b>RU</b>	13.74	<b>RU</b>	15.89	<b>RU</b>	18.27	<b>RU</b>	15.97
2	<b>TW</b>	18.98	IT	7.66	<b>CN</b>	6.54	<b>BR</b>	7.25	<b>BR</b>	10.25
3	<b>CN</b>	9.26	<b>CN</b>	7.34	IT	6.23	IT	6.59	IT	6.87
4	CL	5.19	<b>BR</b>	6.66	<b>BR</b>	6.21	<b>CN</b>	5.99	<b>CN</b>	6.61
5	ES	4.67	<b>TW</b>	6.11	KR	6.02	DE	4.33	TW	5.92
6	US	4.61	DE	5.78	DE	5.03	<b>TW</b>	4.32	DE	5.67
7	CO	3.93	AR	4.56	<b>TW</b>	4.75	IN	4.29	<b>AR</b>	4.46
8	<b>BR</b>	2.97	KR	3.29	RO	3.69	<b>AR</b>	3.95	IN	4.21
9	<b>RU</b>	2.96	UA	3.05	<b>AR</b>	3.54	KR	3.84	JP	3.32
10	DE	2.43	IN	3.03	IN	3.14	RO	3.43	RO	2.96
Total		74.57		61.22		61.04		62.26		66.24

Range	Start	End	Packets	%	Hosts	%
A	20 Nov 2008	28 Dec 2008	482 311	3.41	93 103	2.44
B	28 Dec 2008	20 Feb 2009	1 751 334	12.40	484 931	12.73
C	20 Feb 2009	4 Mar 2009	640 640	4.53	180 631	4.74
D	4 Mar 2009	8 Apr 2009	1 194 125	8.45	348 844	9.15
E	8 Apr 2009	30 Sep 2009	10 047 561	71.17	2 912 630	76.46

vulnerability. Over the period of the Dataset, and particularly in the last 14 months, traffic destined to 445/tcp made a significant contribution to the whole. Given this, it is important to investigate the nature and origins of the datagrams.

While the analysis carried out in this paper is by no means complete, it provides an good example of the kind of focused analysis that can be done with a network telescope, even when considering the limitations of not having packet payloads.

The evolution of the Conficker worm is plotted. The problem with the random scanning and propagation algorithm identified in the reverse engineering of the malware can be clearly observed, and this is seen to be a plausible explanation for the significant difference in traffic observed by the researcher between the dataset being considered for this work, and others utilising different address space. Furthermore, the work presented shows how a network telescope can be used to track the spread and distribution dynamics of widespread Internet worms in the future.

#### ACKNOWLEDGEMENTS

This work was performed in and funded by the Telkom Centre of Excellence in Distributed Multimedia at Rhodes University. Funding was also received from the National Research Foundation Thutuka Program Grant number 69018 and the Rhodes University Research Committee.

#### REFERENCES

- [1] F. Baker, W. Harrop, and G. Armitage, "IPv4 and IPv6 Greynets." RFC 6018 (Informational), Sept. 2010.
- [2] W. Harrop and G. Armitage, "Defining and evaluating greynets (sparse darknets)," in *LCN '05: Proceedings of the The IEEE Conference on Local Computer Networks 30th Anniversary*, (Washington, DC, USA), pp. 344–350, IEEE Computer Society, 2005.
- [3] D. Moore, C. Shannon, G. M. Voelker, and S. Savage, "Network telescopes," tech. rep., CAIDA, 2004.
- [4] Microsoft, "Virus alert about the Win32/Conficker worm (KB962007)." Online, August 18 2008. Last Review: December 1, 2010 - Revision: 10.0.
- [5] Microsoft, "Win32/conficker." Online, 8 Jan 2009. Updated: Nov 10, 2010.
- [6] Microsoft, "MS08-067: Vulnerability in Server Service Could Allow Remote Code Execution (KB958644)," tech. rep., Microsoft, Oct 23 2008.
- [7] Microsoft, "MS03-026: Buffer Overrun In RPC Interface Could Allow Code Execution (KB823980)," tech. rep., Microsoft, July 16 2003. Originally posted: July 16, 2003 Revised: September 10, 2003.
- [8] Microsoft, "MS03-039: Buffer Overrun In RPCSS Service Could Allow Code Execution (KB824146)," tech. rep., Microsoft, September 10 2003.
- [9] Microsoft, "Virus alert about the Nachi worm (KB826234)." Online, August 18 2003.
- [10] Microsoft, "MS04-011: Security Update for Microsoft Windows (KB835732)," tech. rep., Microsoft, April 13 2004. Updated: August 10, 2004.
- [11] Microsoft, "MS06-040: Vulnerability in Server Service Could Allow Remote Code Execution (KB921883)," tech. rep., Microsoft, September 12 2006.
- [12] P. Porras, H. Saidi, and V. Yegneswaran, "An analysis of conficker's logic and rendezvous points," tech. rep., SRI International, 4 February 2009. Last Update 19 March 2009.
- [13] B. Nahorney, "The downadup codex." Online, March 2009.
- [14] B. Irwin, *A framework for the application of network telescope sensors in a global IP network*. PhD thesis, Rhodes University, Grahamstown, South Africa, 2011.
- [15] B. Schneier, "The Zotob Storm," *IEEE Security and Privacy*, vol. 3, pp. 96–, November 2005.
- [16] Microsoft, "Microsoft Security Bulletin MS02-039: Vulnerability in Plug and Play Could Allow Remote Code Execution and Elevation of Privilege (899588)." Online, August 9 2005.
- [17] D. White, "MS05-039 and the Zotob summary." Online, 18 August 2005. Last accessed 2010-12-01.
- [18] E. Aben, "Conficker/Conflicker/Downadup as seen from the UCSD Network Telescope." Online, CAIDA Network Telescope Project - Backscatter, February 2009.
- [19] P. Hick, E. Aben, D. Andersen, and K. Claffy, "The CAIDA UCSD Network Telescope "Three Days Of Conficker" (collection)." Online,



CAIDA Network Telescope Project - Backscatter, 2009. Support for the UCSD Network Telescope "Three Days of Conficker" Dataset and the UCSD Network Telescope are provided by Cisco Systems, Limelight Networks, the US Department of Homeland Security, the National Science Foundation, and CAIDA Members.

- [20] Microsoft, "Malware Protection Center: Win32/Gimmiv," tech. rep., Microsoft, Oct 28 2008. Updated: Apr 17, 2011.
- [21] M. Richard and M. Ligh, "Making fun of your malware." Conference Presentation Defcon 17, Las Vegas USA, August 2009.
- [22] Carnivore.IT, "Conficker does not like me?." Online Blog, 3 November 2009. Accessed 21 November 2010.
- [23] E. Wustrow, M. Karir, M. Bailey, F. Jahanian, and G. Huston, "Internet background radiation revisited," in *Proceedings of the 10th annual conference on Internet measurement*, IMC '10, (New York, NY, USA), pp. 62–74, ACM, 2010.

# TOWARDS ENSURING SCALABILITY, INTEROPERABILITY AND EFFICIENT ACCESS CONTROL IN A MULTI-DOMAIN GRID-BASED ENVIRONMENT

Nureni A. Azeez\* and Isabella M. Venter\*\*

\*Department of Computer Science, University of the Western Cape, Private Bag X17, Bellville, 7535, South Africa Email: 3008814@uwc.ac.za

\*\*Department of Computer Science, University of the Western Cape, Private Bag X17, Bellville, 7535, South Africa Email: iventer@uwc.ac.za

**Abstract:** The application of grid computing has been hampered by three basic challenges: scalability, interoperability and efficient access control which need to be optimized before a full-scale adoption of grid computing can take place. To address these challenges, a novel architectural model was designed for a multi-domain grid based environment (built on three domains). It was modelled using the dynamic role-based access control. The architecture's framework assumes that each domain has an independent local security monitoring unit and a central security monitoring unit that monitors security for the entire grid. The architecture was evaluated using the Grid Security Services Simulator, a meta-query language and Java Runtime Environment 1.7.0.5 for implementing the workflows that define the model's task. In terms of scalability, the results show that as the number of grid nodes increases, the average turnaround time reduces, and thereby increases the number of service requesters (grid users) on the grid. Grid middleware integration across various domains as well as the appropriate handling of authentication and authorisation through a local security monitoring unit and a central security monitoring unit proved that the architecture is interoperable. Finally, a case study scenario used for access control across the domains shows the efficiency of the role based access control approach used for achieving appropriate access to resources. Based on the results obtained, the proposed framework has proved to be interoperable, scalable and efficiently suitable for enforcing access control within the parameters evaluated.

**Keywords:** authorisation, grid, role-based access control, scalability, interoperability, access control, security, multi-domain environment

## 1. INTRODUCTION

Grid computing is an environment that provides unhindered access to computational infrastructure across various domains in academia and industry. It allows the porting, running, sharing and distribution of applications [1]. Since grid computing involves many users from different organizations and domains, sensitive and classified information may be vulnerable if no control policy for regulating and securing all the domains on the grid, is present [2], [3].

The concept of a grid system is analogous to a "water grid system". The facilities of a water grid system make it possible for anyone in his home to open a tap to collect water without knowing exactly where such water is being processed [4]. Similarly grid computing is able to provide endless and ubiquitous access [5] to high quality computing resource without having to know exactly where the data is being processed [1].

Buyya [4], defined a grid as follows: The "grid is a type of parallel and distributed system that enables the sharing, selection, and aggregation of resources distributed across multiple administrative domains based on their (resources) availability, capability, performance, cost, and users' quality-of-service."

The South African Grid (SAGrid) is a typical example of a functional grid. It is a group of South African tertiary institutions (Universities, laboratories and also the Meraka Institute) that are collaborating in the sharing of resources [6].

### 1.1 Why secure a grid?

To prevent sensitive and important information from being copied, altered, divulged to unauthorized users or manipulated has brought about the need for security on a grid system [7]. Without security a grid cannot be considered to be dependable. However, security models on the grid are difficult to implement and to sustain, due to the complexity of the grid environment [8]. Traditional access-based control models are based on recognized inadequacies and there is thus a need to replace them with more flexible [9] models which are relevant to distributed environments [10].

### 1.2 Security challenges

**Scalability:** Scalability caters purposely for future expansion [11]. For a grid environment to be scalable, a centralized administration as well as regular update of the security policies is necessary [12]. In other words, scalability simply means the capability of a grid system

such that it can efficiently handle both a small or large number of nodes and users [13].

**Interoperability:** This can be simply defined as the ability of various systems on the grid to exchange, share and utilize information across platforms. It is a security challenge due to disparate and unequal security policies. The characteristics of an interoperable grid-based environment include:

- the presence of a central authority for security and trust;
- heterogeneous resources, service discovery and management as well as;
- the interdependence of security infrastructures [14], [15].

**Efficient Access control (EAC):** is intended to enforce control over whom (agent) can interact with resources on the grid network. The EAC can be achieved through different means such as authentication and authorisation with the aid of an appropriate access control model. EAC remains a challenge in grid computing mainly because a large number of users are involved. The users are often considered to be dynamic in their requests. This could be attributed to the fact that each domain on the grid has its own policies and the domains are autonomous [33].

To secure a grid based environment without compromising accessibility, interoperability and scalability the following questions can be asked:

- How should a common security policy for various domains on the grid be determined? and
- How should the security of the grid be managed to ensure accessibility of resources in an interoperable and scalable grid based environment?

To achieve the aim of EAC, it was concluded that regulation is required. To regulate and find a solution to the factors which impact EAC within the grid platform, a role based access control (RBAC) model was designed, a prototype built and the prototype was tested with the G3S simulator. The RBAC model is based on three primary rules: role assignment; role authorization and transaction authorization. It was found that the proposed framework is interoperable (in terms of resources; grid middleware, operating system and authorisation), scalable and suitable for enforcing access control within the parameters evaluated.

The remaining part of the paper is organised as follows. In Section II, a summary of related work is presented. A brief analysis of the various security requirements on the grid is explained in Section III. Section IV gives a stratum of the proposed architecture with Subsections A and B presenting the stages of the architectural model. Section V provides a comprehensive overview of the components of the architecture. Section VI gives an

operational overview of the model while Section VII gives an approach for evaluating security in a triple-domain grid-based environment (3DGBE). Section VIII deals with the implementation and evaluation. Finally, the paper is concluded in Section XI.

## 2. SECURITY REQUIREMENTS IN A GRID ENVIRONMENT

The security requirements defined by the International Organization for Standardization (ISO) and the International Telecommunication Union (ITU) are ITU-T Provision X.805 and X.800 [22].

### 2.1 Authorization

For any organization to allow its resources to be jointly shared between all parties involved there is a need for authorization: who should have access to any particular resource and who should not [23][18]. Globus Toolkit Gridmap files [24], Community Authorization Service (CAS) and Virtual Organization Membership Service (VOMS) are authorization measures usually adopted in grid computing [25].

### 2.2 Authentication and Access Control

Impersonation has been identified as a threat [11] in grid environments. Authentication is thus important to prevent illegal access [26]. The main purpose of authentication is solely to confirm that the user is who he claims to represent and not any other person. In both the shared and personal computer system, authentication is usually carried out with the use of a password and username. It has been established that when a password is used to log onto the system [4], the authenticity of a user is usually fully guaranteed. However a password can be stolen hence the information on the system can be vulnerable. Digital certificates, verified by a Certificate Authority [26], are taken as the best way to ensure authentication on the Internet.

### 2.3 Data Confidentiality

The purpose of data confidentiality is to protect data from being divulged to the wrong or an unintended party [27]. Two processes can be used to achieve data confidentiality: data encryption and data decryption. Also, two main types of cryptography can be used to provide data confidentiality [28], i.e. symmetric and asymmetric.

## 3. RELATED RESEARCH

Research done in terms of securing the grid can be divided into three main categories: security-policy aggregation, access control and reliability in grid security.

### 3.1 Security-policy aggregation

In a bid to ensure aggregated security policies across different domains Tari and Fry proposed Global Access Control. A distributed object kernel security service was provided for enforcing and aggregating local and general security policies on the grid. In order to allow control of data aggregation, they provided a security framework Federated Logic Language (FELL) and a logic-based language [16]. The security constraint was enforced by mapping state-transition graphs which model different nodes on the grid. This approach is good and enforces various security measures but it is not scalable since it does not allow more nodes to be added to the grid [6]. Security-policy aggregation in terms of scalability and interoperability still needs to be addressed.

### 3.2 Access control

In the work of Yanxiang et al. a model was developed based on a public key and double-identity authentication on a grid. The model was developed to ensure both authenticity and confidentiality. For the implementation of this model, they applied an RSA (Ronald Rivest, Adi Shamir, and Leonard Adleman) cryptosystem. Furthermore, a double identity authentication approach was employed, to include a time parameter on the server side. Finally, both the server and client produce passwords which change over time. However, this model is not scalable and dynamic as provision was not made for adding users [17].

Some Attribute-Based Access-Control systems such as Akenti and PERMIS have been in use for several grid applications [18]. These authorization systems apply their own rules. As a result, a dynamic attribute based access control is required for the grid computing environment [19]. In this model, there is no room for interoperability across various domains on the grid.

John McLean [20] came up with a framework in which Mandatory Access Control (MAC) models, allow for changes in security to be formalized. He employed algebra to construct his model that paves the way for the discretionary access control for  $n$  persons. This model is good but does not handle the problem that emanates from the separation of duties and cyclic redundancy as a result of roles and hierarchy among participants on the grid.

### 3.3 Reliability in grid security

Laccetti and Schmid [21] came up with a framework for reliable grid security infrastructures using Grid Security Infrastructures (GSI) and Community Security Policy (CSP). Their analysis captured the policies and rules upon which GSI and CSP were based. Trust relationship based on a cryptographic key was used as a guiding principle. It was finally revealed that authentication implemented at grid levels develop a trust relationship that is transitive which is not the case when authentication is used at

operating system tier. Formal model algebra was adopted in developing the security of the grid [21]. This model is not flexible as it has limited application.

## 4. STRATUM OF THE PROPOSED ARCHITECTURE

The proposed architecture constitutes two stages, each of which involves two phases (see Figure 1).

1. The first phase involves various domains. Each of the domains is characterised by a user and a local security-monitoring unit (LSMU).
2. In the second phase, the central security-monitoring unit (CSMU) interacts directly with all the domains of phase.
3. The third phase is a processing phase. All activities that result in the granting of resources are carried out in this phase.
4. The fourth phase is a grid environment phase where many resources are available. A user is allowed to access this phase based on a decision made in the third phase.

### 4.1 Stage 1 of the architecture

This stage involves the interaction between various users and the domains' LSMU with the CSMU. The architecture in Figure 2 and Algorithm 1 give comprehensive information with respect to this interaction and message passing between grid entities. In Figure 2, a theoretical framework of the interaction between the user and the LSMUs of three domains, as well as its interaction of the three domains and the CSMU is depicted.

To explain the process of the architecture presented in Figure 2, let us assume the following scenarios:

1. Adam, a grid user (GU) in Domain A, forwards his request to his domain's LSMU, where his authorisation is verified and confirmed. Adam's status (eligibility as a user) is thus determined. This phase makes Adam's access right to the intended domain known.
2. The LSMU then sends Adam's request to access a resource in any intended domain to the CSMU to reconfirm his authorisation right in his own domain and his rights to access resources of any other domain. The CSMU verifies whether Adam qualifies to access the required resource. There are two outcomes: YES (acceptable) or NO (not acceptable).
3. If NO, the process (request) terminates and the feedback message is communicated to the user.
4. If YES, a "clearance" certificate will be given to the user (Adam) by the LSMU of the intended domain and the user can proceed to stage 2.

5. If there is a successful processing in stage 2, the user will proceed to access resources in the grid environment.

#### 4.2 Stage 2 of the architecture

This stage deals with the interaction between the processing phase and grid environment. This stage comes into play if and only if there is a positive feedback during Stage 1 (See Figure 3 and Algorithm 2).

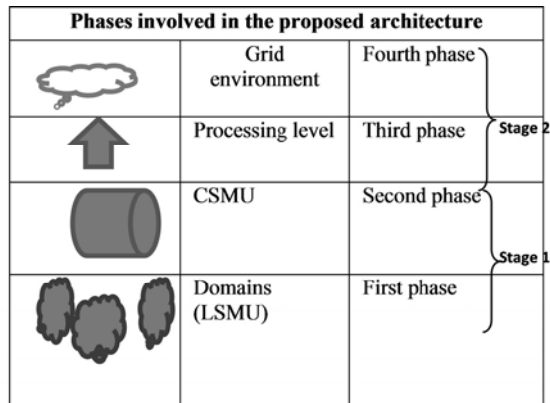


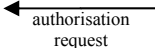
Figure 1: Phases involved in the proposed architecture

Algorithm 1: Algorithm describing the working relation of components in Figure 2

Required by Domain A, Domain B, Domain C, LSMU, CSMU

Begin:


feedback [authorisation] = "Yes or No";


GU {Domain A, B, C}: 


LSMU:

if authorisation = "No"  
then : terminate (process)

else:

if authorisation = "YES"  
Then: LSMU  CSMU

CSMU  { (GU (role)) }:  
If CSMU [permission (decision)] = "yes"

then: CSMU  stage 2;

Stop

The operation of the architecture is presented in Figure 3:

1. Through the grid entry link, the GU requests access (with the role authorisation-certificate) from the grid user authentication service (GUAS). The request is either granted or not.
2. If the feedback is negative, the entire process will be terminated immediately and the request will cease to continue.
3. However, if the feedback is positive (YES), then the request will be forwarded to the policy information point (PIP) (a protocol of XACML (eXtensible Access Control Markup Language)

for access control). This is to source detailed information about the user.

4. The request will further be directed to the policy decision point (PDP), which is another XACML protocol for access control. The PDP is responsible for making a decision on whether the user may access the requested domain. The feedback of the PDP will either be positive (YES) or negative (NO). If the feedback is negative, the entire process stops.
5. If the feedback is YES the request is conveyed to the PEP.
6. The PEP will demand an updated version of the user permission certificate from the PDP (grid virtual organisation (VO)-PDP).
7. A certificate validation/update will be transferred to the centralized resource database server (CRDS) from the PDP (grid VO PDP).
8. Finally, a message will be sent to the user to proceed and access resources on the grid.

The procedure is applicable from either of the domains available on the grid i.e. either Domain A to Domain B or from Domain A to Domain C, and vice versa.



In order to ensure a smooth and efficient access control mechanism on the grid and also to improve the performance of the architecture, the LSMU works with the CSMU. That is, there is smooth correspondence between the local security units of all the domains with the central security unit for the entire grid. They both communicate and work hand-in-hand to achieve a flexible, interoperable and scalable grid environment.

Algorithm 2: Describing the working scenario of the architecture presented in Figure 3

Require: role, user, PIP, PEP, PDP, GUAS, CRDS

Feedback: [yes/no]:

Begin: from stage 1:


request:  GU role certificate [GEL]  
then: GUAS  Role (GU);

Else: if feedback (GUAS) = "No"


Then: terminate (process);

If feedback (GUAS) = "YES"

Then: request  PIP;


PIP:  PDP; //request for appropriate decision


If feedback = "YES"


proceed:  PEP;

Else if feedback = "No"

then: stop (process)

Getupdate:  PDP-VO;

update (certificate):  CRDS;

finalDecison:  Pass to (VO (Grid))

Begin [GU] :access [resource]

Stop

## 5. OVERVIEW OF THE BASIC COMPONENTS OF THE ARCHITECTURE

In the proposed model, each of the domains available in the virtual organisation (VO) has an LSMU saddled with the responsibility of the domain's local security access control and management. The CSMU is an advanced access control and management system that handles access control and authorisation for the various grid entities across the three domains of the model. For any access request by a grid user, the LSMU would verify the user's access privilege. The model is based on the adoption of the XACML's (eXtensible Access Control Markup Language) request-response protocol which makes use of four basic components. The components are: PEP, PDP, PIP and PAP. However, in this model, only PEP, PDP and PIP are used because of their relevance, usefulness and application in the proposed architecture.

### 5.1 Assumptions

1. A user from domain A (Adam) may intend to access a resource in domain B and a user in domain B (Ben) may also be interested in accessing resources from domain A;
2. A user in domain A (Adam) may wish to access resources in domain C while a user that is in domain C (Charles) may equally be interested in resources of domain A.

These are two possible scenarios when a three domain based architecture is being considered. Scenario 1 is illustrated in Figures 2 and 3 and it is equally applicable to other scenarios. Adam, Ben and Charles are users in the domains A, B and C respectively. Each of them is bound with the security and access framework in their respective domains. There are six ways in which access could be requested: request can come from Domain A to Domain B, from Domain A to Domain C, from Domain B to Domain C, etc.

## 6. OPERATIONAL OVERVIEW OF THE MODEL

The security of each individual domain is quite dependable and efficient; because each domain has its own access control and monitoring policy which is monitored by the LSMU. If a user, however, wishes to access resources in another domain, the user from the designated domain will first need to be verified by his domain.

This is achieved by translating the certificate of his domain to the domain in which he wishes to access resources. The translation (or conversion) targets the access privileges and the identities in other domains on the grid. CSMU is mainly in charge of monitoring and overseeing access and security relationship from one domain to another domain depending on where an entity requires access. Also, CSMU is equally responsible for

maintaining the information for mapping interactions between domains (see Figure 2 as well as 3).

## 7. DETERMINATION OF SECURITY IN A 3-DOMAIN GRID VIRTUAL ORGANISATION

### 7.1 Definition of simulation parameters

In order to evaluate the effectivity of the security of the domains; the following parameters defined below were taken into consideration.

*Definition 1:* Let  $DSR(A,B)$  and  $DSR(A,a)$ , denote the *direct security rate* which is determined and evaluated when the CSMU finds and grants permission and access privilege to a user from domain B to domain A or from an entity  $a \in$  domain A to domain A depending on from where the access is requested.  $DSR(A,B,C)$  denotes the DSR between the three designated domains.

*Definition 2:* Similarly let  $SR(A,B)$  or  $SR(A,a)$  denote the security rate for accesses from domain B to domain A or for an access from entity  $a \in$  Domain A to domain A.  $SR(A,B,C)$  denotes the security rate between the three designated domains.

*Definition 3:* Let  $Assess(a_i \dots a_j)^m$  denote assessment for entities  $a_i \dots a_j$  when  $a_i \dots a_j$  terminate at time step  $m$ , and  $-1 \leq Assess(a_i \dots a_j)^m \leq 1$  shows either rejection or satisfaction during the assessment of the entities involved. While '-1' indicates the rejection, which will reduce the value of SR, '+1', however, indicates satisfaction, which will increase the value of SR.

*Definition 4:* Let  $DSR(a_i \dots a_j)$  stands for "Direct Security Rate" in a grid for entities  $a_i \dots a_j$ .

*Definition 5:* Let  $Rep(A, a)$  denote *reputation* and *status* of entity  $a$  in Domain A on a grid.

*Definition 6:* Let  $Approv(a_i \dots a_j)^m$  stand for the *approval* in the service request for  $a_i \dots a_j$  after  $m$  time steps.

## 8. SECURITY EVALUATION IN A 3DGBE

Determining or evaluating the security rate in a multi-domain grid-based environment is completely different from what is obtainable in a single-domain environment. The main reason for this is the interaction and relationship between the grid entities involved. Unlike in a single-domain environment, a multi-domain grid environment has more entities from one domain to another to interact with. Hence, to handle the complexities that arise from the user's accessibility to different domains resources, the security rate (SRs) for the entities of each domain is useful for quick and accurate evaluation of the security within different domains. The approach adopted for determining the inter-domain security rate value is simple and provides the benefit of feedback that is flexible and dynamic in nature.

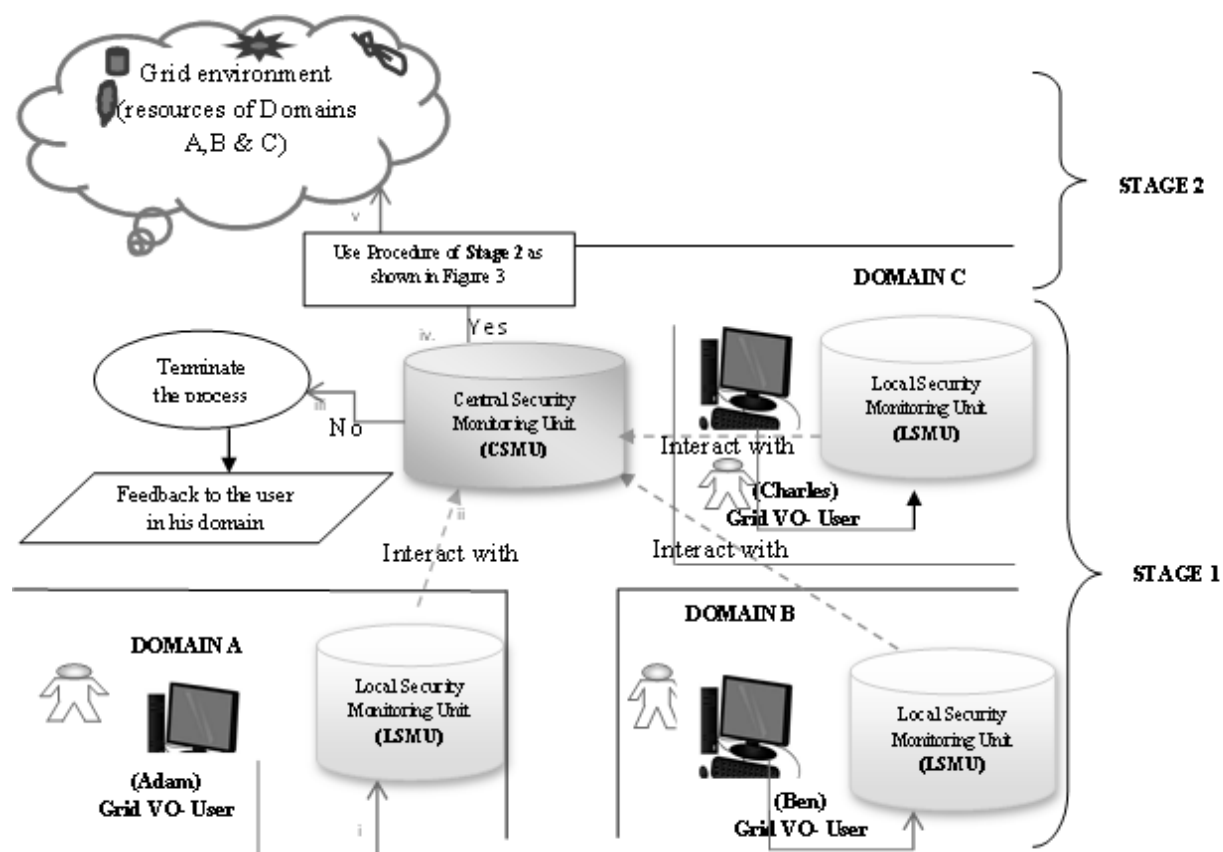


Figure 2: A 3-domain role based access control architecture showing interaction between users, CSMU and LSM

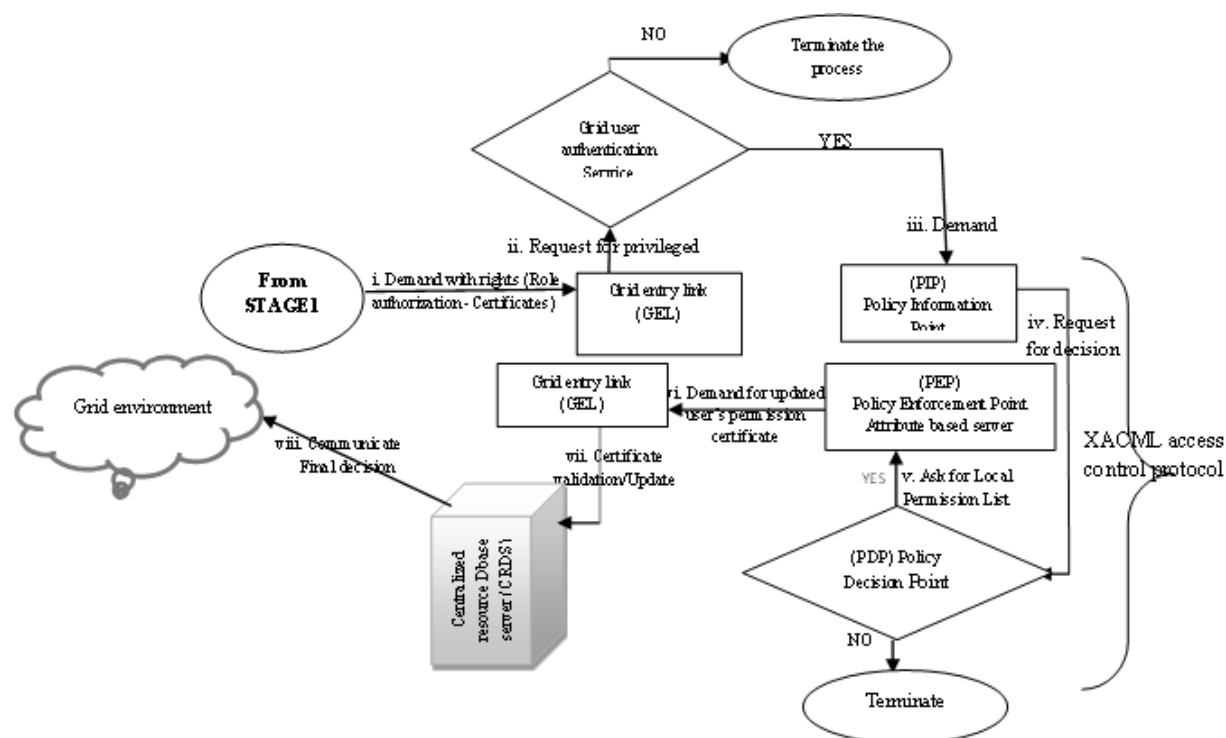


Figure 3: A 3DGBE with RBAC architectural framework of the proposed model

$Rep(C, a_i)$  yields status/repute of entity  $a_i$  to domain  $C$  in a virtual organisation considered that  $a_i$  is not an entity in domain  $C$ . It is worth mentioning that  $A$ ,  $B$ , and  $C$  represent three different domains being considered while  $a_i$ ,  $b_i$  and  $c_i$  are entities in the three domains. Hence,

$$SR(A, B, C) = \lambda_1 DSR(A, B, C) + \lambda_2 Rep(A, B, C) \quad (1)$$

Equation 1 is used to evaluate the  $SR$  in the three domains  $A$ ,  $B$  and  $C$  where the weight  $\lambda_1$  and  $\lambda_2$  are positive and  $\lambda_1 + \lambda_2 \leq 1$ .

$$DSR(A_i, A_j) = \frac{\sum_{a \in A_j} DSR(A_i, a)}{|A_j|} \quad (2)$$

Where  $a$  is an entity from the domain  $A$ . Given two different domains  $A_i$  and  $A_j$  with  $i, j \in 2[1 \dots n]$ , where  $i \neq j$ , and  $n$  is the number of domains.

Therefore,

$$DSR(A, C) = \frac{\sum_{c \in C} DSR(A, c)}{|C|} \quad (3)$$

When considering any domain,  $A$ ,  $B$  or  $C$ , Equation 2 is generic and can therefore be used to compute direct security rate ( $DSR$ ) between them. The same is applicable to Equation 3 where domains  $A$  and  $C$  were specifically considered.

## 9. REPUTE AND STATUS ACROSS DOMAINS

For domains  $A_i$  to  $A_j$  with  $i \neq j$ , the status of entities is determined as follows:

$$Rep(A_i, A_j) = \sum_{a \in A_j} \theta_a Approv(A_i, a) Rep(A_i, a) \quad (4)$$

Where  $\theta_a > 0$  is the weight given to  $Approv(A_i, a)$  for  $a \in A$  and  $\sum_{a \in A} \theta_a = 1$ . Equation 4 implies that the  $Rep$  can be determined from any desired domain and can be extended to any number of domains.

## 10. IMPLEMENTATION AND EVALUATION

Various simulation experiments were carried out using different simulators however in this example, Grid Security Services Simulator (G3S) was used [29]. To carry out an empirical evaluation of the access control architecture, the simulation was developed in Java making use of Jbuilder. In the three domains in this experimental grid based environment: domain  $A$  was made up of a cluster of seven nodes (or computers while the other two domains were LANs (Local Area Network) comprising of 13 computers each. The simulated grid environment was developed using the Globus toolkit 5.0.5. All the hardware of the test bed was embedded in Linux Ubuntu 12.04.

A computer hosted a database with the information of all users and acted as the LSMU for each domain while a computer server with a static IP address was chosen as the CSMU for the experimental grid. For an efficient and reliable evaluation, the resources and entities considered were accessible when a grid user requested their services.

Table 1: Simulation parameters with their corresponding values

Parameters	Corresponding values
$\lambda_1$	0.25
$\lambda_2$	0.36
$DSR(a_i, \dots, a_j)$	0.34

### 10.1 Evaluation of 3DGBE and MAC

In the experiment, 3DGBE access control was compared with MAC, which is a popular access control method. Table 1 provides the detail of the parameters used in the simulation experiment. Users were provided and assigned with both MAC-based and 3DGBE access control

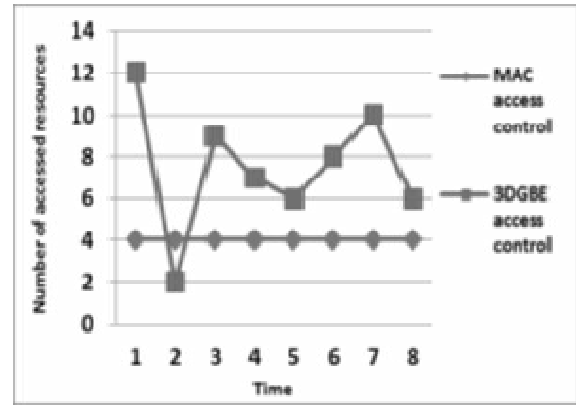


Figure 4: Number of available resources in the two access control policies 3DGBE and MAC

simultaneously.

The number of resources varied over different time periods. It was noted that the number of available resources varied over time in the 3DGBE access control architecture whereas it remained unchanged in the MAC-based access control system (see Figure 4).

It can thus be deduced that access to resources would be flexible when deploying a 3DGBE architecture.

Equation 2 was used to evaluate the security rate without considering any weights. Entities in either of the domains  $A$ ,  $B$  or  $C$  could request resources from any desired domain and the destination domain then evaluated such



requests. The result of the SR was thereafter obtained (see Figure 5).

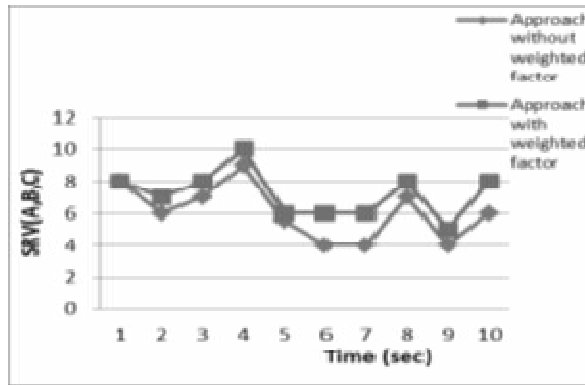


Figure 5: Secure rate comparison using two approaches

Equation 1 was used for calculating the SR between the domains. The security rate value will vary if there are no weighted values for  $\theta_j$ . Table 2 gives a summary of the required parameters. The simulation result revealed that the available number of grid nodes has a direct influence on the turnaround time as shown in Figure 6.

Table 2: Simulation parameters for  $\lambda_1$  DSR, REP of Domains A, B, C alongside the number of entities

Parameters	Corresponding values
$\lambda_1$	0.6
First (initial) value of DSR (A,B,C)	0.58
First (initial) value of Rep(A,B,C)	0.44
Entities in domain A	20
Entities in domain B	15
Entities in domain C	23

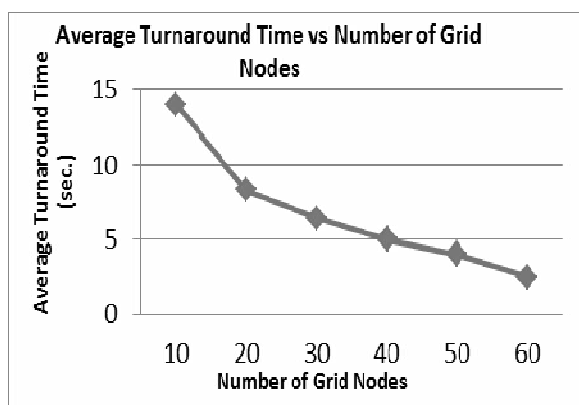


Figure 6: Average turnaround time versus number of grid nodes

This implies that as the number of grid nodes increases the average turnaround time reduces and thereby increases the number of service requesters (grid users) on the grid.

To further prove and sustain the argument that the model developed and implemented is scalable, Figure 7 shows that as the number of service requesters increases, there is little and slight effect on the turnaround time which does not impact on the users' services and request time.

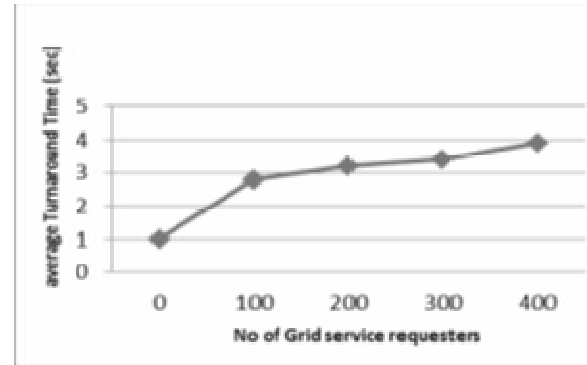


Figure 7: Average turnaround time versus number of service requesters

In order to further sustain the argument that the 3DGBE architecture is scalable therefore, the effect of increase in the number of nodes against the volume of data that is transferred within a given period of time (throughput), were observed and measured.

The result of the comparison of 3DGBE (which uses X.509 certificates) with MAC, CAS, AKENTI and PERMIS (that use own their certificate formats) is presented in Figure 8. The result shows that 3DGBE has the highest degree of interoperability when compared to the others.

In the initial setup indicates domains A, B and C contain

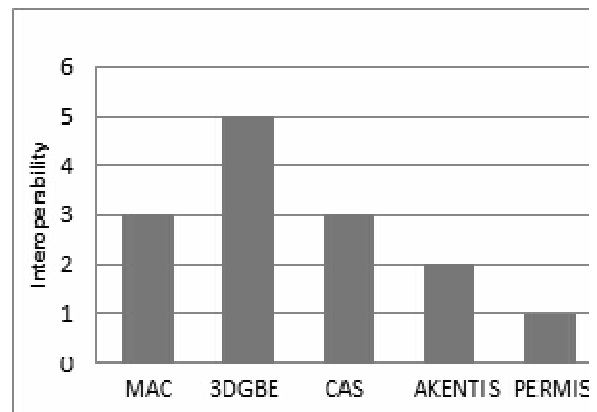


Figure 8: Comparative evaluation of interoperability of 3DGBE with the existing system

7, 13 and 13 nodes respectively. To ascertain the effect of an increase in nodes on the performance of the throughput, the number of nodes in each domain was increased as follows: domain A had 12 nodes; domain B had 20 nodes while domain C had 25 nodes.

The result obtained (see Figure 9), shows an increase in throughput as follows: when the number of grid nodes in domain A comprises 12, the throughput is 100MB/s, when the number of grid nodes in domain B is increased to 20, the throughput is 2200MB/s, while 3100MB/s is attained when the number of grid nodes in domain C is increased to 25.

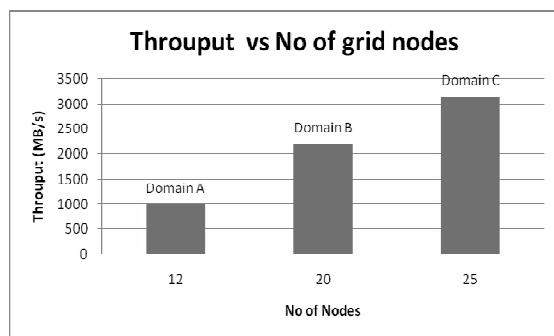


Figure 9: Throughput (MB/s) vs No of nodes

From Figure 9, it can be deduced that as the number of grid node increases, the throughput also increases thereby increasing the number of resources being accessed within a given time. This proves that the scalability of the 3DGBE architecture.

The use of grid middleware has been identified as one of the ways for solving the challenge of interoperability among multiple administrative domains. This model adopted the XACML access control protocol, which showed the highest level of interoperability when compared to others.

*Appropriate handling of authentication and authorisation through LSMU and CSMU:* The central security-monitoring unit (CSMU) maintains a high degree of interoperability between the users on the grid. For a resource request to be allowed, approval needs to have been given by the local security monitoring unit (LSMU). The CSMU serves as the central point which makes the final decision for grid resources to be accessed. There is smooth correspondence between the LSMU of each of the three domains and CSMU. The purpose of this is to ensure adequate and an efficient data sharing mechanism among the domains with the view of achieving interoperation of authorisation. CSMU forwards requests and authorisation from all the domains across the domain to access any required resources.

*Operating system interoperability:* Aside from the fact that the architecture permits various applications to run

(application interoperability), the architecture has proved to be interoperable in both the LINUX (Ubuntu 10.04) and Windows operating system with different middlewares (Globus, Glite and UNICORE). The evaluations that were carried out were done on both operating systems, LINUX performs better than Windows. The feature of operating system interoperability is noticeably weak in some of the existing models such as MAC, AKENTIS and PERMIS.

Interoperability with Grid middleware: Middleware can be regarded as:

*“A mediator layer that provides a consistent and homogeneous access to resources managed locally with different syntax and access methods” (Priol, 2005: 32)*

Aneka, Alchemi, Cosm P2P Toolkit, Globus, Gridbus, Grid Datafarm, GridSim (Toolkit for Grid Resource Modeling and Scheduling Simulation), Jxta Peer to Peer Network, Legion, NorduGrid middleware, PUNCH, Simgrid, Storage Resource Broker (SRB), ProActive, Unicore and Vishwa are prominent grid middleware [4]. With some of the listed grid middlewares, the virtual organisation interoperability issue remains a problem. This is because of the absence of upper-level semantic concepts in their grid middleware layers [30].

To address this challenge, a tri-middleware integration approach was used. The 3DGBE was enabled with three different middlewares across the three available domains on the grid, namely Globus 5.0, gLite and Alchemi for domains A, B and C respectively (see Figure 14).

With this approach of grid interoperability that is based only on the middleware integration, various middlewares were deployed on different domains and allow the same set of users to share and access resources with well-established and defined virtual origination's policies, irrespective of the grid middleware they intend to use.

The problem of middleware differences was solved effectively by using common standards. The middlewares were implemented as a subset of specifications of the different grid middlewares. With this interoperation based approach to middleware integration, different middlewares need not necessarily communicate with each other to be able to merge and share resources.

The various computing resources (installed in the three domains) were all made accessible to all the grid users regardless and independent of their middlewares they intend to adopt. The middleware integration of three or more grid resources is easier and grid users across the three domains can access resources without any hindrance.

Aside from Globus 5.0.2, two other middlewares were installed for the interoperability integration testing: gLite

and Alchemi. These three middlewares were chosen because they were available.

Local clusters and their security were considered as the basic elements of infrastructure that could be reused to aid interoperability across the three domains. Globus 5.0.2, gLite and Alchemi aid Torque/MAUI as local scheduler, which supports the ability to share local clusters and all available resources. These middlewares use X.509 certificates hence the same grid resources can be accessed, shared and distributed by these three different grid middlewares. Achieving interoperability with three different middlewares is simple with Globus 5.0.2, gLite and Alchemi respectively. Both gLite and Alchemi adopted the Grid System Infrastructure (GSI) model developed by Globus for user's authorisation.

The model (GSI) makes use of a digital certificates and proxies for the authentication and authorisation of hosts and users. Established on X.509 digital certificates and proxies, GSI was extended in both the gLite and Alchemi with the agreement of the Virtual Organisation Membership Service (VOMS), which released fully X.509 compatible signed extensions to proxies. Additional information about the users, which is required for the mapping on various levels of authorisation, is achieved through these extensions. Since VOMS proxy is compatible with the X.509 proxy, therefore the former's proxy can be taken as authentication and authorisation credential when deploying it on the three grid middlewares.

The distribution of resources across a tri-middleware based architecture is the second focus in achieving interoperability. The cluster manager in charge of the local resources was configured in such a way that jobs could be submitted despite differences in middlewares. The local scheduler in the architecture is Torque/MAUI. This scheduler is supported by Globus 5.0.2, gLite and Alchemi hence it is very easy to express new queues new and added middlewares in order to utilise the same resources.

*Declarative language (queries) for interoperability:* A big challenge for developing and implementing an interoperable 3DGBE lies in the ability to efficiently, sufficiently express "cross-queries" (inter-domain queries) that relate information from different domains. To overcome this challenge, a Meta-Query Language (MQL) [31], which is similar to the Structures Query Language (SQL) was adopted.

MQL was used for querying and restructuring tables containing information across different domains. As illustrated in the scenario presented, MQL was used to query and restructure information across Domains A, B and C within the federation. Hence, interoperability was achieved across these three domains through MQL dynamic query mappings.

Test scenario:

Three different databases were created for domains A, B and C: a University, Hospital and Banking database respectively (see Tables III, IV and V).

### 10.2 Inter-domain queries

*CASE 1:* Databases were created for domains A and B and they were aggregate and joined to enhance further operations (see Figure 10).

*CASE 2:* To combine the information of the databases of domains A and C, the query as depicted in Figure 11 was issued and the report generated shows a UNION of both

```
SELECT DB1.*, DB2.*
FROM DomainA.Database1.dbo.myTable AS
DB1
INNER JOIN
DomainB.Database2.dbo.myTable AS DB2
ON DB1.id = DB2.id
```

Figure 10: Query for aggregating data from Domains A and B

```
-- FROM Domain_A
SELECT * FROM
[MyDatabaseOnDomain_A].[dbo].[MyTable]
Table1
INNER JOIN
Domain_C].[MyDatabaseOnDomain_C].[dbo].
[MyOtherTable] Table2
ON Table1.ID = Table2.ID
```

Figure 11: Query for aggregating data from Domains A and C

```
SELECT a.ID_NO,
a.Surname,a.Nationality,
b.File_No,b.Patient_Condition,
c.Service_Code,a.Tax_ID,b.Age
```

Figure 12: Cross-domain query for joining data from Domains A, B and C

databases for both domains.

Cross-domain queries were applied to each of the newly obtained tables. For example to obtain ID No, Surname, and Nationality from Table 3; File No, Patient\_Condition and Age from Table 4; as well as Service\_Code and Tax\_ID from Table 5, from respectively domains A, B and C, the cross-domain query depicted in Figure 12, was used.

Table 3: University Database for Domain A

ID No	Student No	Surname	First name	Initial	DOB	Passport No
1000	3008814	Azeez	Nureni	N.A	1978/07/10	QOO12345
1001	3066278	Adewale	Abiola	M.D	1980/04/08	RE0047476
1200	2340078	Abidoye	Philip	P.O	1976/06/09	WE345678
1220	2357858	Scholtz	Josue	S.J	1981/04/05	VB7878784
1260	3400993	Magnuth	Henry	H.M	1975/06/09	FD8787878
1320	3476002	Andy	Liu	X.L	1984/02/09	RE7878784
1400	3455266	Achmed	Imran	I.A	1986/02/09	UD5785785
1523	2004556	Jonathan	Magnus	I.M	1974/07/10	TY8989896

Table 4: Hospital Database for Domain B

Patient ID	File No	Surname	First name	Patient Condition	Date Of Admin
1000	0031	Azeez	Nureni	Medical checkup	2012/09/04
1200	0045	Abidoye	Philip	Eye problem	2012/06/04
1320	0067	Andy	Liu	Hearing problem	2012/03/06
1400	0012	Achmed	Imran	Back pain and X-ray	2011/09/04
1523	0056	Jonathan	Magnus	Pregnancy	2012/01/02
1001	0023	Adewale	Abiola	Blood test	2012/09/09
1220	0013	Scholtz	Josue	Car accident	2011/09/01
1260	0011	Magnuth	Henry	Head injury	2012/03/06

Table 5: Banking Database for Domain C

ID No	Customer Name	Service Code	Account ID	Current Balance
1000	Azeez, N.A	FD9989	45454599XX	R 56.5XX
1200	Abidoye, P.O	YU7878	57757577XX	R 100XX
1320	Andy, X.L	HJ7880	47747744XX	R 562XX
1400	Achmed, I.A	WE4545	67677676XX	R 00.7XX
1523	Jonathan, I.M	QW567	56565655XX	R 06.7XX
1001	Adewale, M.D	NH7676	86868612XX	R 129XX
1220	Scholtz, S.J	YE85852	67676768XX	R 451XX
1260	Magnuth, H.M	BG2323	13243535XX	R 000XX

The cross-domain queries were introduced purposely to handle heterogeneity of information represented in different structures, to provide distinct aggregation capability in addition to the principal objective of multi-domain database interoperability.

### 10.3 Efficient access control

Access control remains a bottleneck when accessing resources in a multi-domain environment such as a grid. Each user participating in grid resource sharing tends to gain access to resources within its jurisdiction. Some grid users might want to access resources for which they are not authorized.

To achieve efficient access control, hierarchical role based access control was adopted for specifying role, services as well as permission for each user from any domain. To explain this, consider a specific scenario

(Health); where each of the domains has roles, services and permission defined among the users (see Algorithm 3).

Terms and Definitions as used in this context:

- Let H1, H2 and H3 denote the hierarchies and let the role hierarchy (RH) denoted as H1, H2 and H3 be assigned to domains A, B and C respectively where  $H1 > H2 > H3$

It could be recalled that a “... hierarchy is mathematically a partial order defining a seniority relation between roles, whereby the seniors’ roles acquire the permission of their juniors, and junior roles acquire the user membership of their seniors” [35]

Algorithm 3: Algorithm for efficient access control in a 3DGBE

---

Required: Domains A, B and C, LSMU, CSMU  
 Grid User (GU) identification;  
 Get the Domain's hierarchy as {H1, H2, H3};  
 Assign hierarchy to the chosen domain;  
 Obtain GU role;  
 Retrieve GU services - permission;  
 Proceed to the grid

---

Thus,

- Let Role\_Domain A denotes all roles defined in domain A;
- Let Role\_Domain B denotes all roles defined in domain B; and
- Let Role\_Domain C denotes all roles defined in domain C.

Role and services specification for DOMAIN A:

1. Role\_Domain A = {Physician, Cardiologist, Neurologist, Obstetrician, Pathologist, Pulmonologist, Surgeon, Pediatrician, Oncologist, Dermatologist}
2. Services (permission)
  - {Physician (write patient record, read patient record, write prescription, read prescription, examine patient)}
  - {Cardiologist (treat heart disease, write patient record, read patient record, write prescription, read prescription)}
  - {Neurologist (treats brain, examine nervous system, write patient record, read patient record)}

Role and services specification for DOMAIN B:

1. Role\_Domain B = {patient, nurse, pharmacist, dentist, Psychiatrist, Podiatrists}
2. Services (permission)
  - {patient (read prescription, read patient record)}
  - {nurse (write patient record, read prescription, read patient record)}
  - {pharmacist (read prescription, read patient record, select prescription)}

Role and services specification for DOMAIN C:

1. Role\_Domain C = {Ultrasound Technologist, X-Ray Technician, Clinical Technologist, Clinical Technologist, Dental Assistant, Dental Laboratory Technician}
2. Services (permission)

- {Ultrasound Technologist (read patient record, take ultrasound, analyse images)}
- {X-Ray Technician (read patient record, perform x-ray on patient, interpret and analyse x-ray result)}
- {Clinical Technologist (read patient record, perform medical test, interpret result)}

Whenever a grid user (GU) specifies his domain, the corresponding hierarchy of such a user will be instantly verified and produced. The hierarchy is divided into three layers; hierarchy 1 (H1) for domain A, hierarchy 2 (H2) for domain B and hierarchy 3 (H3) for domain C.

Any GU with H1 as hierarchy is from domain A and can access resources from any desired domain whose services are defined. The formulation is such that  $H1 > H2 > H3$  thus H1 has the highest hierarchy and can access all the resources within its domain and the domain under it, that is, domains B and C.

Similarly, H2 permits grid users to access all available information in its domain and resources below it, that is, in H3. However, H3 permits grid users to access resources within its domain alone. This initial access control framework is efficient in a 3DGBE as users whose identities are not linked to a specific hierarchy will automatically be denied access to resources.

The prototype was implemented in a Java Runtime Environment 1.7.0.5 for the workflows that define the model's task. The implementation reveals that the access control adopted is efficient within the three domains considered.

Figure 13 is a comprehensive pictorial explanation of domain roles and their services as spelt out for each user. From the foregoing, it is clear that a cardiologist who has his roles defined in domain A of H1 has the corresponding listed services allocated to it. A dentist whose domain is B with domain hierarchy H2 can only access the allotted services. Any attempt to access other information or services, will result in a "rejection or denial of service" which signifies the efficiency of the access control put in place. Finally, the same condition is applicable to the ultrasound technologist who has his/her services defined in domain C.

## 11. CONCLUSION

Evidence from the literature reviewed, showed that scalability, interoperability as well as efficient access control are three basic security challenges that need to be addressed if the full scale-benefits of grid computing are to be realized.

Based on the results obtained, the architectural framework has proved to be scalable when the average

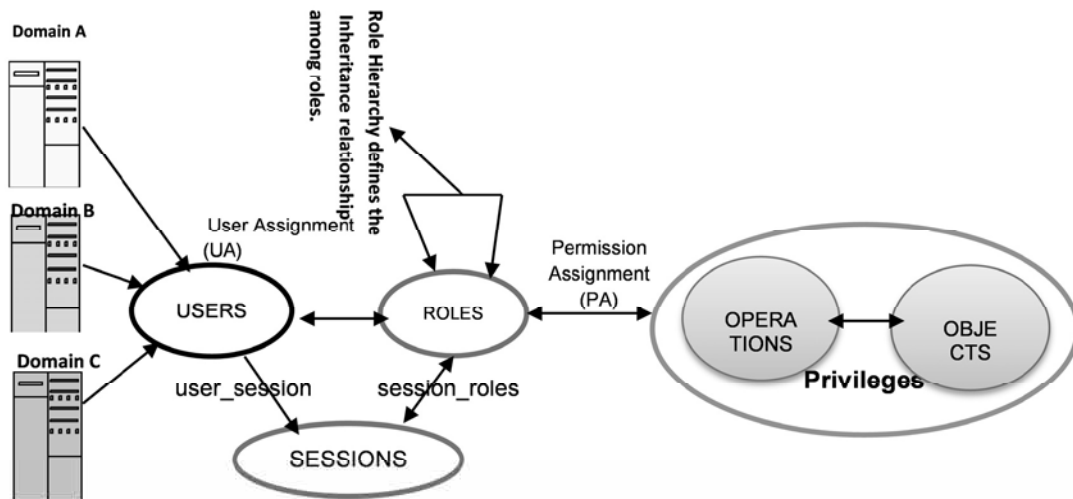


Figure 13: Implementation of hierarchical RBAC and 3DGBE

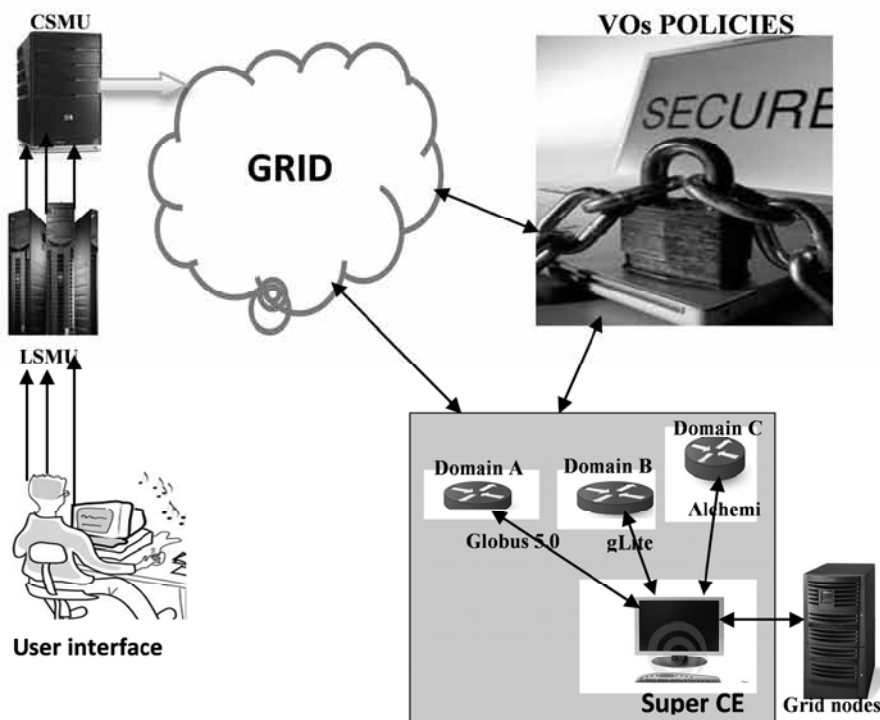


Figure 14: Tri-middleware based infrastructure for 3DGBE interoperability

turnaround time was measured against the number of grid nodes. More convincing results were achieved when the throughput and number of nodes as well as when the average turnaround was measured against the number of grid requesters.

The results obtained in terms of interoperability when the operating systems, grid middleware, LSMU and CSMU as well as database were implemented and experimented with, proved that the model's framework is interoperable.

Finally, the efficient access control was evaluated with a role based access control and implemented with a health scenario, and it yielded the expected result.

Other issues that need to be investigated in grid computing are: grid maintenance, grid coordination, pricing, grid auditing and scheduling. These pose challenges that deserve attention for future work. The objectives set out in this research work were achieved. It is therefore, believed that a full-scale implementation of

this model, on a real grid system, will ensure a secure, scalable and interoperable grid-based environment.

## 12. ACKNOWLEDGEMENT

The authors would like to thank the Research Committee of the University of the Western Cape for funding and the Center for High Performance Computer (CHPC) in Cape Town for assistance.

## 13. REFERENCES

- [1] N. A. Azeez, T. Iyamu, and I. M. Venter, "Grid security loopholes with proposed countermeasures," in *ISCIS 2011*. Springer Verlag, London, 2011, pp. 411–418.
- [2] T. Herath and H. R. Rao, "Protection motivation and deterrence: a framework for security policy compliance in organisations," *European Journal of Information Systems*, vol. 18, no. 2, pp. 106–112, 2009.
- [3] M.-S. Hwang and W.-P. Yang, "A new dynamic access control scheme based on subject-object list," *Data and Knowledge Engineering*, vol. 14, no. 1, pp. 45–56, 1994.
- [4] R. Buyya, "Economic-based distributed resource management and scheduling for grid computing," Ph.D. dissertation, Monash University, Melbourne, Australia, 2002.
- [5] H. Baktash, M. B. Karimi, M. R. Meybodi, and A. Bouyer, "2L-RBACG: A new framework for resource access control in grid environments," in *2010 Fifth Int. Conf. on Digital Information Management (ICDIM)*. Thunder Bay: IEEE Computer Society, 2010, pp. 359–366.
- [6] GStat, "Grid gstat 2.0," 2010, <http://gstat.gridops.org/gstat/sa-gr>.
- [7] Z. Mao, N. Li, H. Chen, and X. Jiang, "Trojan horse resistant discretionary access control," in *SACMAT 09: Proc. 14th ACM Symp. On Access Control Models and Technologies*. Stresa, Italy: IEEE Computer Society, 2009, pp. 237–246.
- [8] B. Bouwman, S. Mauw, M. Petkovic, and E. Philips Res.-Ordina, "Rights management for role-based access contro," in *Consumer Communications and Networking Conference, CCNC. Las Vegas, N: IEEE Computer Society*, 2008, pp. 1085 – 1090.
- [9] Z.-D. Shen, F. Yan, W.-Z. Qiang, X.-P. Wu, and H.-G. Zhang, "Grid system integrated with trusted computing platform," *Computer and Computational Sciences, International Multi-Symposiums on*, vol. 1, pp. 619–625, 2006.
- [10] R. Sandhu, E. Coyne, H. Feinstein, and C. Youman, "Role-based access control models," *IEEE Computer*, vol. 29, no. 2, pp. 38–47, 1996.
- [11] R. Lakshminish, L. Ling, and I. Arun, "Scalable delivery of dynamic content using a comprehensive edge cache grid," *IEEE Trans. on Knowl. Data Eng.*, pp. 614–63, May 2007.
- [12] O. Rahmeh and P. Johnson, "Towards scalable and reliable grid networks," in *IEEE/ACS International Conference on Computer Systems and Applications (AICCSA-2008)*. Doha, Qatar: IEEE Computer Society, 2008, pp. 253–259.
- [13] A. Detsch, L. Gaspary, M. Barcellos, and G. Cavalheiro, "Towards a flexible security framework for peer-to-peer based grid computing," in *2nd Workshop on Middleware for Grid Computing*, Toronto, Canada: ACM, 2004, pp. 52–56.
- [14] S. Basit, B. James, B. Elisa, and G. Arif, "Secure interoperation in a multidomain environment employing rbac policies," *IEEE Trans. Knowl. Data Eng.*, pp. 1557–1577, 2005.
- [15] G. Pankaj, "Application of a distributed security method to end-2-end services security in independent heterogenous cloud computing environment," in *IEEE World Congress on Services*. Washington, DC: IEEE Computer Society, 2011, pp. 379–384.
- [16] Z. Tari and A. Fry, "Controlling aggregation in distributed object systems: A graph-based approach," *IEEE Trans. Parallel Distrib. Syst.*, pp. 23–32, December 2001.
- [17] H. Yanxiang, L. Fei, and H. Wensheng, "The design and implementation of security communication model in grid networks," in *Int'l Conference on Computer Science and Information Technology*, IEEE, ICCSI, 2008, pp. 421–424.
- [18] A.-B. Ali, Z. Hussein, and S. Francois, "Access control mechanism for mobile ad hoc network of networks (MANoN)," *Software Technology Research Laboratory, De Montfort University, Leicester, Tech. Rep.*, 2009.
- [19] H. Mohteshim, "Passive and active attacks against wireless LAN," in *IASTED, 2004. Int. Assoc. of Sci. and Technology for Development*, 2005. [Online]. Available: <http://www.iasted.org/conferences/2004/Innsbruck/pdcn.htm>
- [20] J. McLean, "The algebra of security," in *IEEE Symp. on Security and Privacy*. Naval Research Laboratory, Washington, D.C.: IEEE Comput. Soc., 2008.
- [21] G. Laccetti and G. Schmid, "A framework model for grid security," *Future Generation Computer Systems*, vol. 23, no. 5, pp. 702–713, June 2007.
- [22] NHSE, National HPCC Software Exchange, pp. 4–8, 2009. [Online]. Available: <http://wotug.org/parallel/nhse/>
- [23] A. Imine, A. Cherif, and M. Rusinowitch, "An optimistic mandatory access control model for distributed collaborative editors," *INRIA, Tech. Rep.*, 2009.
- [24] I. Foster and C. Kesselman, "Globus: A metacomputing infrastructure toolkit," *The International Journal of Supercomputer Applications and High Performance Computing*, vol. 11, pp. 115–122, 1997.
- [25] D. Chadwick, "Authorisation in grid computing," *Information Security Tech.*, vol. 10, pp. 33–40, 2005.
- [26] C. Rongxing, Lu; Zhenfu, "A simpler user authentication scheme for grid computing," *International Journal of Network Security*, vol. 7, pp.202–210, 2008.
- [27] Z. Weide, W. David, D. V. and Glenn, and H. Marty, "Flexible and secure logging of grid data access," in *7th IEEE/ACM Int. Conf. on Grid Computing (Gridn 2006)*. Barcelona, Spain: IEEE/ACM Computer Society, 2006, pp. 1–8.
- [28] MSDN, "Data confidentiality," 2005. [Online]. Available: <http://msdn.microsoft.com/en-us/library/ff650720.asp>
- [29] N. Syed and R. Michel, "Grid security services simulator (G3S)—a simulation tool for the design and analysis of grid security solution," in *Proc. First Int. Conf. on e-Science and Grid Computing (e-Science '05)*. IEEE Computer Society, 2005, pp. 421–428.

- [30] D. Welch, & S. Lathrop, 2003. Wireless Security Threat Taxonomy. Proceedings of the 2003 IEEE workshop on information assurance United States Military Academy West Point. NY.
- [31] M. W. W Vermeer, & P. M. G. Apers, 1996. On the applicability of schema integration techniques to database interoperation. In International Conference on Conceptual Modeling/the Entity Relationship Approach, pages 179-194.
- [32] L. V. Lakshmanan, F. Sadri, & I. N. Subramanian, 1996. SchemaSQL - a language for interoperability in relational multi-database systems. In Proceedings of the 22nd VLDB Conference
- [33] L. Bo, F. Ian, S. Frank, A. Rachana, & F. Tim, (1990). Attribute Based Access Control for Grid Computing. pp. 1-13. International Journal of grid computing, vol.9 no.3, March 1990
- [34] T. Priol, 2005. GRID Middleware. South Korea, Advanced Grid Research Workshops through European and Asian Co-operation published by the European Research Consortium for Informatics and Mathematics, pp. 1-11.
- [35] BIBLIOGRAPHY Chuan-Lun, R., Xiao-Hui, Z., Zhong-Xian, L., Xin-Xin, N., & Yi-xian, Y. (2010). Towards Hierarchical-User RBAC model. International Conference on Machine Learning and Cybernetics (ICMLC), 2010 (pp. 2870- 2874). Qingdao, China: IEEE.



## IGNORANCE TO AWARENESS: TOWARDS AN INFORMATION SECURITY AWARENESS PROCESS

T. Gundu\* and S.V. Flowerday\*\*

\* *Department Information Systems, University of Fort Hare, 50 Church Street, East London, South Africa, 5201 E-mail: tapgun@gmail.com*

\*\* *Department Information Systems, University of Fort Hare, 50 Church Street, East London, South Africa, 5201 E-mail: sflowerday@ufh.ac.za*

**Abstract:** With most employees in small and medium enterprise (SME) engineering firms now having access to their own personal workstations, the need for information security management to safeguard against loss/alteration or theft of the firms' important information has increased. These SMEs tend to be more concerned with vulnerabilities from external threats, although industry research suggests that a substantial proportion of security incidents originate from insiders within the firm. Hence, physical preventative measures such as antivirus software and firewalls are proving to solve only part of the problem as the employees using them lack adequate information security knowledge. This tends to expose a firm to risks and costly mistakes made by naïve/uninformed employees. This paper presents an information security awareness process that seeks to cultivate positive security behaviours using a behavioural intention model based on the Theory of Reasoned Action, the Protection Motivation Theory and the Behaviourism Theory. The process and model have been refined, tested through action research at an SME engineering firm in South Africa, and the findings are presented and discussed in this paper.

**Keywords:** Information Security Awareness, Security Behaviour, Information Security Training.

### 1. INTRODUCTION

SMEs, especially those in the engineering sector, are continually investing significantly in their overall Information and Communication Technologies (ICTs) making Information Security a major concern for the safeguarding of their information assets [10]; [15].

Most of these SMEs have information security policies that present rules to be adhered to [19]. These rules provide a solid foundation for the development and implementation of secure practices within the firms. However, the existence of these formal security policies does not necessarily mean that employees will adhere to the rules [10]. Subsequently, employees need to be aware of the security practices prescribed in the firm's policy.

Information security awareness and training are frequently used for raising awareness of employees and promoting appropriate information security behaviour. This ensures their employees realise the importance of security and the adverse consequences of information security failure plus that there is the potential for people to deliberately or accidentally steal, damage, or misuse data stored within a firm's information systems and throughout the organisation [20]; [45].

Engineering firms rely heavily on digital information stored on networked servers. This information includes patented and unpatented private and confidential designs, plus drawings and client information that are prone to security threats. Engineering SMEs tend to ignore the risk of the uninformed employee and are more concerned with vulnerabilities from external threats; however, industry research suggests that the uninformed employee, by not

behaving securely, may expose the firm to serious security risks, for example: data corruption, deletion, and even commercial espionage [1]; [5]; [6]; [22]; [33].

Insider risk can result from two sources: intentional and unintentional behaviour [45]. This paper focuses on unintentional naïve mistakes although intentional dangerous tinkering by disgruntled employees is also a significant threat. Unintentionally uninformed employees (insiders) may expose a firm's information assets to risk by making naïve mistakes, visiting malware infested websites, responding to phishing emails, using weak passwords, storing their login information in unsecured locations, or giving out sensitive information over the phone when exposed to social engineering techniques. Unintentional mistakes by the employee is not an attempt to discredit the firm or make a profit by selling confidential data, but rather as a result of inadequate employee training about information security, that is their lack of security awareness and the consequences of their actions. This weakness can never be totally eliminated, but a well-structured security awareness campaign helps to reduce the risk to acceptable levels [19]; [22].

SME Engineering firms have high levels of trust in their employees not to compromise security; hence, they believe information security awareness is not an issue for them [42]. Ironically, it is more important for SMEs compared to larger firms as employees often have multiple roles and thus have access to a variety of financial, organizational, customer and employee information. Furthermore, there is less segregation of duties in SME engineering firms, thus less control over access to information. Whilst exposed to many of the same threats and vulnerabilities as large organisations,

SMEs do not have access to the same level of resources [42]; this makes their risk even higher.

The purpose of this paper is to present, refine and validate a process that can be followed by SMEs to ensure that their employees are information security aware. This process is mainly based on a behavioural intention model to be presented in section 3.2 and Kruger and Kearney's [21] information security measuring concepts.

The behavioural intention model bases its argument on three principal theories: the Theory of Reasoned Action (TRA) [3], the Protection Motivation Theory (PMT) [28] and the Behaviourism Theory (BT) [47]. Previous works have used research frameworks that integrated TRA, PMT and BT with other theories (even if unconsciously) [10]; [13]; [30]. According to Anderson and Agarwal's [27] review of literature in this area, no prior information security research has used all three theories in a single information security study. Although research has been carried out in the area of information security awareness, there is a lack of literature on the effectiveness of information security awareness methods on the basis of psychological theories as well as a lack of description of the underlying theory of these methods. Psychology is the science of the mind and behaviour. Social psychology has been used for many years for research in the area of education, learning and human behaviour [29].

Action Research was conducted at a civil engineering firm to refine and validate the process. Elden and Chisholm [44] note that action research is change oriented, seeking to introduce changes with positive social values, the key focus of the practice being on a problem and its solution.

The remainder of the paper is organised as follows: first, the information security awareness process is presented, then follows the behavioural intentional model; thirdly, the method for measuring information security is discussed; followed by the analysis and results; finally, the paper concludes by discussing its findings.

## 2. THE INFORMATION SECURITY AWARENESS PROCESS

Information security theories posit that in order for security efforts to be effective, firms must ensure that employees are part of the security effort [4]; [32]; [34]; [38]; [45].

This section discusses the proposed information security awareness process in the form of a flowchart. Figure 1 shows the proposed information security awareness process for SME engineering firms. The flowchart has four processes (P1, P2, P3 and P4) and three checks (C1, C2 and C3). When planning an information security awareness program, the first step should be to check the existence of an up-to-date Information Security Policy (C1 and C2); however, the firm where the action research

was conducted had a sound and up-to-date policy that accurately reflected its overall posture towards information security. The step of drafting or updating an Information Security Policy (P1 and P2) was not carried out and is beyond the scope of this study.

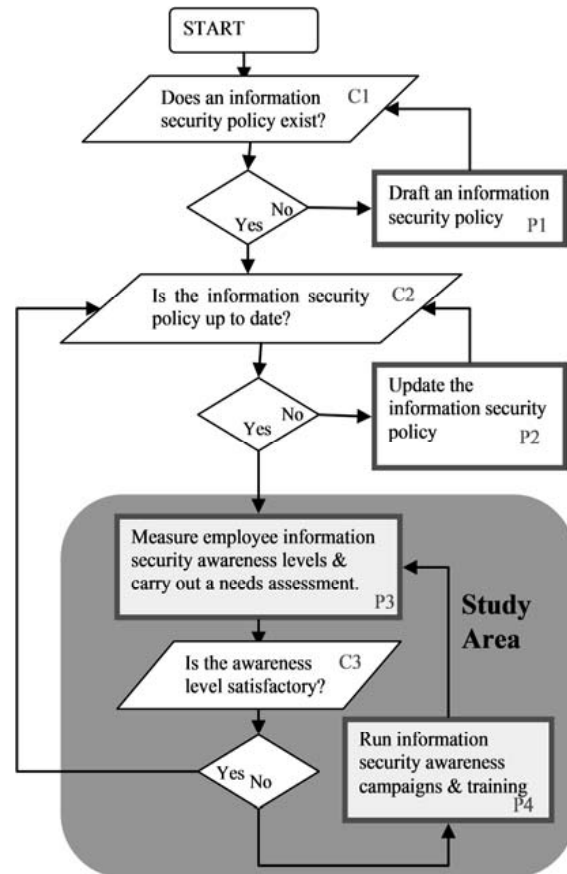


Figure 1: Information security awareness process

The next step is to measure employees' current level of information security understanding (P3) so as to identify any knowledge gaps. During the action research, this needs assessment process highlighted the firm's awareness and training requirements. For example, in the first iteration of the action research, the measurement revealed that employees had an inadequate understanding of password creation, safe Internet usage, virus and firewall understanding, thus highlighting some topics for awareness training. These results also justified to the firm's management the need to allocate resources towards information security awareness and training. The method for measuring employee awareness levels was adapted from Kruger and Kearney's [21] previous research; the details of this method will follow in section 4.

The next step would then be to verify if the current level of information security awareness is at an acceptable level (C3). When conducting the action research, it was found that the level of information security awareness during the first iteration was unsatisfactory and exposed the need for information security awareness campaigns and training. If the levels are unsatisfactory, awareness

campaigns and training sessions should be conducted. During the action research, an e-learning based awareness campaign/training was conducted (P4). Its implementation and maintenance is discussed in detail in section 4. The awareness level was measured again after the awareness campaign and results showed that the knowledge gap was closing, but the results were not yet satisfactory according to the scales used (these will be discussed in the data analysis section). The process was then run again for a second and third iteration. The results of the third iteration were satisfactory and the process was stopped.

### 3. INFORMATION SECURITY AWARENESS CAMPAIGN AND TRAINING (P4)

Awareness from a different perspective: "It is believed that about 200 years ago people did not know about the germ theory; they did not know that they should wash their hands and boil surgical tools to limit the spread of disease and infection. Even though people know these things today, do they always wash their hands before eating, or even after doing something icky?" [39].

Unfortunately, not everyone does so even when they know better. This highlights that the real challenge is not just to teach people, but also to help them change their behaviour. Security knowledge cannot help much if employees do not act on it; hence, this section provides guidelines for implementing and maintaining comprehensive e-learning information security awareness and training campaigns.

Security awareness and training assists in tempering the attitude that security policy is restrictive and interferes with an employee's ability to do his/her work. The better the employee's understanding of information security issues, the more they understand the importance of security and the ways in which security protects them and enables them to do their work in a safer and more effective environment [19].

Information security campaigns are divided into awareness and training. Awareness aims to raise the collective knowledge of information security and its controls, while training aims at facilitating a more in-depth level of employee information security understanding. An effective information security awareness and training programme seeks to explain proper rules of behaviour when using the firm's computer/information systems. The programme communicates information security policies and procedures that need to be followed. Additionally, the campaign imposes sanctions when noncompliance occurs [10].

The BERR 2008 survey [2] suggests that the majority of firms rely upon written materials for training in one form or another. However, simply developing and circulating a policy will not be sufficient to foster appropriate

understanding and behaviour. Most companies use the traditional classroom style for awareness and training. However, this study seeks to apply the now widely used tried and tested e-learning concept to information security awareness and training. Jenkins et al [16] and Ricer et al [26] report that there is no significant difference between people who learn using a computer or the traditional classroom style in the short or long-term retention of knowledge.

Additionally this section introduces the behavioural intention model. This model attempts to explain how employee information security awareness knowledge can affect behavioural intentions (towards policy compliance and positive security culture). Behaviourists believe that employees are born with limited innate reflexes (stimulus-response units that do not need to be learnt) and that all of an employee's complex behaviours are as a result of learning through interaction with the environment [47]. Thus, belief in information security awareness and training should help mould information security behaviours. The information security awareness campaigns and training in P4 on the Information Awareness Process (Figure 1) are based on a behavioural intention model to be explained next.

#### 3.1 Theoretical background of the behavioural intention model

Based on the problems presented in the preceding sections, this section serves to propose, explain and relate the Theory of Reasoned Action (TRA), the Protection Motivation Theory (PMT) and the Behaviourism Theory (BT) to the behavioural intention model.

##### 3.1.1 Theory of Reasoned Action

TRA framework specifically evaluates the relative importance of two incentive components: (1) attitude (2) subjective norm. It suggests that a person's Behavioural Intention (BI) depends on the person's Attitude (A) about the behaviour and Subjective Norms (SN) i.e.  $BI = A + SN$ . Attitude towards behaviour is defined as the individual's positive or negative feelings about performing certain actions. Subjective norm is defined as an individual's perception of whether people important to the individual think the behaviour should be performed. As a general rule, the more favourable the attitude and the subjective norm, the greater the perceived control and therefore the stronger the employee's intention to perform the behaviour in question [7]; [17]; [23]; [29].

The Theory of Reasoned Action helps to explain how the employee's attitude towards security and perceived corporate expectation affects the employee's behaviour towards information security. Consequently, the employee's attitude and perceived expectations influence the employee's behavioural intention.

The employee's attitude is affected by cultural, dispositional and knowledge influences. Cultural influences are associated with the employee's background. Dispositional influences are associated with the employee's usual way of doing things. Knowledge influences are associated with the level of knowledge of the subject in question. The employee's attitude can therefore be moulded by information security awareness campaigns and training. The subjective norm is what the employee perceives the firm requires of him/her and perception of how peers would behave in similar scenarios [9]; [13]; [30]. Corporate expectations can therefore be communicated to employees via information security and training sessions. In summary, information security awareness campaigns will help change employee attitudes towards information security and will aid in communicating the firm's expectations to its employees.

### 3.1.2 Protection Motivation Theory

Protection Motivation Theory (PMT) was developed by Rogers (1983). It was developed from the expectancy value theories and the cognitive processing theories, its aim being to assist and clarify fear appeals. PMT has been noted as one of the most powerful explanatory theories for predicting an individual's intention to engage in protective actions [27]. Information security awareness and training instil knowledge in the employees and assists in motivating protection. In essence, protection motivation emanates from both the threat appraisal and the coping appraisal. Threat appraisal describes an individual's assessment of the level of danger posed by a threatening event [28]; [40]. It is composed of perceived vulnerability and perceived severity.

#### *Threat appraisal:*

1. Perceived vulnerability i.e. an employee's assessment of the probability of threatening events. In this study it refers to threats resulting from noncompliance with the firm's information security policy (ISP).
2. Perceived severity i.e. the severity of the consequences of the event. In this instance, imminent threats to the firm's information security may arise from noncompliance with the firm's ISP.

The coping appraisal aspect of PMT refers to the employee's assessment of his or her ability to cope with and avoid the potential loss or damage arising from the threat [40]. Coping appraisals are made up of self-efficacy, response efficacy and response cost.

#### *Coping appraisal:*

1. Self-efficacy: this factor emphasizes the employee's ability or judgment regarding his or her capabilities to cope with or perform the recommended behaviour. In the context of this paper, it refers to the sorts of skills and measures needed to protect the firm's information assets [11]; [30]; [40].

2. Response efficacy: this factor relates to the belief about the perceived benefits of the action taken by the individual [28]. Here, it refers to compliance with the information security policy as being an effective mechanism for detecting a threat to the firm's information assets.
3. Response cost: this factor emphasizes the perceived opportunity costs in terms of monetary, time and effort expended in adopting the recommended behaviour, in this instance the cost of complying with the ISP. Previous research has used PMT and found it useful in predicting behaviours related to an individual's computer security behaviour both at home and in the work situation [9]; [27], as well as Information Security Policy (ISP) compliance [10]; [30].

### 3.1.3 The Behaviourism Theory (BT)

Watson coined the term "*behaviourism* [47]." Critical of Wundt's emphasis on internal states, Watson urged psychology to focus on obvious measureable behaviours [47]. Watson believed that theorising thoughts, intentions or other subjective experiences was unscientific [47]. Behaviourism as a theory was primarily developed by Skinner [47]. According to Skinner [47] it loosely encompasses the work of other behavioural researchers like Thorndike, Tolman, Guthrie and Hull.

These investigators had similar underlying assumptions on the processes of learning. These basic assumptions are summarised as follows: First, learning is manifested by a change in behaviour. Second, the environment shapes behaviour. And third, the principles of contiguity (how close in time two events must be for a bond to be formed) and reinforcement (any means of increasing the likelihood that an event will be repeated) are central to explaining the learning process. For Behaviourism, learning is the acquisition of new behaviour through conditioning.

#### *There are two types of possible conditioning:*

1. Classical conditioning: where the behaviour becomes a reflex response to stimulus as in the case of Pavlov's Dogs. Pavlov was interested in studying reflexes when he saw that the dogs drooled without the proper stimulus. Although no food was in sight, the dogs still salivated. It turned out that the dogs were reacting to lab coats. Every time the dogs were served food, the person who served the food was wearing a lab coat [49]. Therefore, the dogs reacted as if food was on its way whenever they saw a lab coat. In a series of experiments, Pavlov then tried to figure out how these phenomena were linked. For example, he struck a bell when the dogs were fed. If the bell was sounded in close association with their meal, the dogs learned to

associate the sound of the bell with food. After a while, at the mere sound of the bell, they responded by salivating. Pavlov's work laid the foundation for many other psychologists including Watson's ideas. Watson and Pavlov shared both a disdain for "mentalistic" concepts (such as consciousness) and a belief that the basic laws of learning were the same for all animals whether dogs or humans [49].

2. Operant conditioning highlights reinforcement of behaviour by a reward or punishment. The theory of operant conditioning was developed by Skinner [47] and is known as Radical Behaviourism. According to Reynold [48] the word 'operant' refers to the way in which behaviour 'operates on the environment'. Briefly, a behaviour may result either in reinforcement, which increases the likelihood of the behaviour recurring, or punishment, which decreases the likelihood of the behaviour recurring. It is important to note that, punishment is not considered to be applicable if it does not result in the reduction of the behaviour, and so the terms punishment and reinforcement are determined as a result of the actions. Within this framework, behaviourists are particularly interested in measurable changes in behaviour [48]. In operant conditioning we learn to associate a response (our behaviour) and its consequence and thus to repeat acts followed by good results and avoid acts followed by bad results [48].

#### 3.1.4 The Behavioural Intention Model

Following the preceding discussion, it can be observed that the TRA, PMT or the BT can effect desirable behavioural intention. However, the behavioural intention model in Figure 2 attempts to encourage better behavioural intentions by combining the three theories into one model. Discussions on the behavioural intention model are explained in this section.

Subjective norms have a positive effect on information security policy (ISP) compliance behavioural intention. TRA indicates that individuals' attitudes impact on behavioural intentions [24]. To that end, a positive attitude toward ISP compliance bodes well for good behavioural intention. Conversely, negative attitudes will diminish an individual's ISP compliance and good behavioural intention. Thus, individuals with positive beliefs and values about their firm's ISP might display favourable tendencies towards complying with such rules, requirements and guidelines [10]; [13].

Attitude toward Information Security Policy (ISP) compliance will have a positive effect on ISP compliance behavioural intention. With respect to ISP, it is to be expected that individuals with high information security capabilities and competence will appreciate the need to

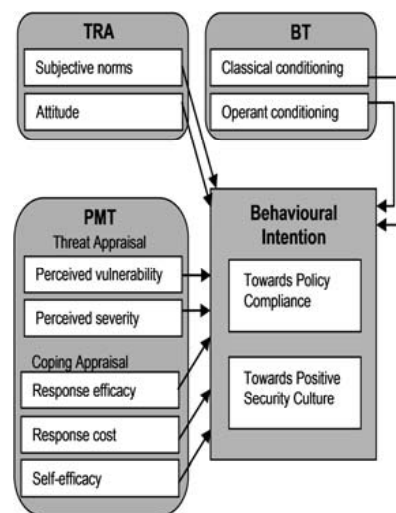


Figure 2: Behavioural intention model

follow organisational ISPs, and such individuals may be better placed to realise the threats of noncompliance [43]. Self-efficacy will have a positive effect on ISP compliance behavioural intention. According to Pahnla et al [30], response costs may include monetary expense, timing inconveniences, embarrassment or other negative consequences, which result from an individual's behaviour. Employees are reluctant to follow or adopt recommended responses if they perceive that a considerable amount of resources i.e. time, effort, and money will be used in pursuit of a low rewarding goal [8]; [9]. Conversely, if small amounts of resources are required in implementing a measure, it may be adopted [36]; [41]. Reducing the Response Cost tends to increase the likelihood of an individual performing a recommended behaviour [40]. Past studies have confirmed that Response Costs are negatively related to intention to use security measures [9]; [41].

Response Cost will have a negative effect on ISP compliance behavioural intention because usually employees believe information security measures are difficult and lengthy.

When an individual possesses requisite knowledge about the effectiveness of a recommended coping mechanism in providing protection from a threat or danger, the individual is more likely to adopt an adaptive behaviour [9]; [28]; [40]. If an individual has doubts regarding the effectiveness of a measure, he or she may not readily accept it [18]. Accordingly, individuals who believe that their organization's ISP has guidelines and coping mechanisms to avert threats and dangers in their context, they are more likely to develop an intention to adopt it [10].

Response efficacy will have a positive effect on ISP compliance behavioural intention. In general, when employees perceive a threat, they often adjust their behaviour in response to the level of risk and determine if they are willing to accept the risk or not [8]; [41]. Thus,

an individual's perceived severity tends to be positively linked to their intentions to follow protective actions [36]. If an individual perceives a threat to his or her firm's Information Systems (IS) assets, such an individual will more than likely follow guidelines and requirements laid out in their ISP [13]; [30].

Perceived severity will have a positive effect on ISP compliance behavioural intention with respect to safe computing in the firm; however, individuals who consider themselves immune to security threats are more likely to ignore security measures at work [10]; [13]; [30]. It is reasonable to expect that an individual who perceives high risk to their firm's information system resource will be more likely to adopt protective behaviours.

Therefore, perceived vulnerability will have a positive effect on Information Security Policy (ISP) compliance behavioural intention because employees will be made aware of the vulnerability of the firms' information assets.

### 3.1.5 *Information Dissemination Method (E-Learning)*

When information security campaign material based on the needs assessment has been compiled, there is a need to choose a method for communicating the information to the employees. During the action research in this study, an e-learning method was used instead of the conventional classroom style because it provided a configurable infrastructure that integrated learning material, policies, and services into a single solution which quickly, effectively and economically created and delivered awareness and training content. E-Learning allows employees to train at their own convenience and learn at their own pace. It has also proved to be cheaper than bringing everyone together, in terms of time and money. This section therefore seeks to explain how e-learning can be used as a tool for communicating and testing information security awareness training.

E-learning has grown considerably over the past several years as technology has been integrated into education and training. E-learning may be defined as instruction delivered electronically via the Internet, Intranets, or multimedia platforms such as CD-ROM or DVD [35]. The literature review highlighted that research work on e-learning as a tool for information security awareness and training is still in its infancy and that no such tool has been used to date in SMEs.

The e-learning awareness and training program for this study was designed and developed by the researcher with assistance from a multimedia designer and a Web page developer using Macromedia Flash, Macromedia Dream Weaver, PDF, PowerPoint, Access, Gold Wave, and Photoshop software in order to present the program material in a visual and auditory format. This was presented in the form of a website containing information identified by the needs assessment and most relevant

recent information security topics. Since information security is a diverse area with many topics, the importance of each topic varies from one firm to another depending on the nature of the risks faced so there is no universal information security awareness training. The training/awareness and testing could be completed in 1-3 hours depending on the speed at which the employee worked. The website for training and awareness was constructed as follows:

**Home Page:** provides an introduction to information security and the motivation behind the training/ awareness campaign. Employees need to be motivated as to why information security is important. The home page then links to the awareness pages.

**The Awareness/Training Pages:** supply information on topical issues and examples of breaches. These pages contain all the information about information security required by employees.

**The Test Page:** was used as the data collection tool for acquiring data from the employees; this was used to measure their information security awareness levels.

All the pages had attractive information security pictures/video clips/jokes in an effort to create a more relaxed e-learning environment.

The employees participating in the study received an email with instructions on how to use the awareness and training material including a link to the awareness and training website.

E-Learning is a broad term and this paper wishes to stimulate the development of E-Awareness initiatives.

## 4. MEASURING INFORMATION SECURITY AWARENESS LEVELS (P3)

After the security awareness campaign was launched, it was important to measure its success and draw conclusions from the measured results. Measurement provides evidence of the campaign's effectiveness and reveals where knowledge gaps still exist. Measurements were not limited to a verification of whether the message was received by the target audience, but detected the effectiveness of the message, method and behavioural change.

According to a survey by Richardson [31], 32% do not measure information security awareness in their firms, because there are no commonly agreed and understood standard measurements for the effectiveness of information security awareness campaigns and training. Two distinctive challenges are identified when

developing a measuring tool and performing the actual measurements. These challenges are “what to measure” and “how to measure it” [12]; [21].

#### *What to measure:*

Kruger and Kearney [21] identified three components to be measured, namely what the employee knows (Knowledge), how they feel about the topic (Attitude), and what they do (Behaviour).

The attitude of employees towards information security is important because unless they believe that information security is important, they are unlikely to work securely, irrespective of how much they know about security requirements. Knowledge is important because even if an employee believes security is important, he or she cannot convert that intention into action without the necessary knowledge and understanding. Finally, no matter what employees believe or know about information security, they will not have a positive impact on security unless they behave in a secure fashion. Figure 3 below shows how enhanced security is achieved by correlating attitude, knowledge and behaviour.



Figure 3: Enhanced Security

#### *How to measure:*

Measuring such intangibles as Attitudes, Knowledge and Behaviour is difficult. The action research made use of multiple data collection techniques such as assessment tests, online surveys, participant observation, informal interviews and document surveys for gathering data. However, only the results from online assessment tests were used to calculate security awareness levels; information gathered using the other techniques was only used for needs assessments.

Online Survey and Assessment Tests enable identification of broad trends [14]. An agreement scale was used to allow employees to indicate degrees of agreement with statements about information security.

The assessment test contained questions that seek to test for knowledge, attitude and behaviour. The following are examples of the questions asked:

#### *Example statement for test of knowledge:*

Internet access to the firm's systems is a corporate resource and should be used for business purposes only.

1. True 2. False 3. Do not know

#### *Example statement to test attitude:*

Laptops are usually covered with existing insurance cover so there is no special need to include them in security policies.

1. True 2. False 3. Do not know

#### *Example statement to test behaviour:*

I am aware that one should never give one's password to somebody else; however, my work is of such a nature that I do give my password from time to time to a colleague (only to those I trust!).

1. True 2. False 3. Do not know

## 5. DATA ANALYSIS AND RESULTS

The engineering firm where the action research was conducted was established in 1997. It develops designs, plans, models and geotechnical surveys for the clients it consults. It has thirty two employees, four of whom have no access to the firm's computer resources. This left a sample size of twenty eight employees. The action research was conducted over a ten-month time period from February, 2011 to November, 2011.

In this action research, the researcher was not regarded as an objective, passive outsider. The firm's management expected him to be an active participator, helping to plan and deliver the training program and evaluate its results. When the information security awareness of the employees was measured for the first time during the needs assessment, only 21% (6 employees) had sufficient levels of information security. Table 1 summarises the information security understanding of the employees per iteration.

Table 1: Employees information security awareness understanding levels

	Needs assessment	Iteration 1	Iteration 2	Iteration 3
Employees understanding level	6 (21%)	18 (64%)	24 (86%)	27 (96%)

The number of employees with sufficient levels of information security understanding increased on the second iteration due to an increase in knowledge. The majority of employees had sufficient information security understanding after iteration 2 and 3.

All the employees were shown their test results and the overall group results during each iteration in order to motivate those who had not performed well. However, the number of employees showing sufficient levels of

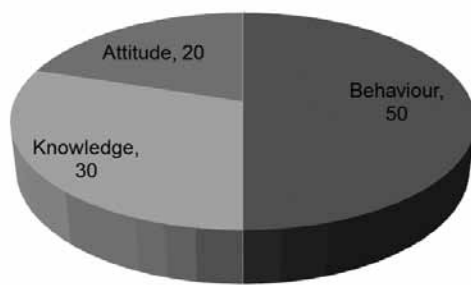


Figure 4: Awareness importance scale [21]

information security understanding is not a true reflection of a firm's overall information security awareness levels; hence Kruger and Kearney's [21] method of analysing data acquired through the measuring methods discussed in the preceding sections was used. This method involved weighting the three aspects being measured in Figure 4.

This weighting was verified with the Managing Director and the Human Resources Manager of the firm who agreed that behaviour was the most important measure followed by knowledge then lastly attitude. The results and importance weightings were processed in a spreadsheet application and the output was finally presented in the form of graphs and awareness maps as comparable to Kruger and Kearney's study [21]. Table 2 below shows the scale used to interpret the level of awareness. Kruger and Kearney's scale was slightly modified to take into consideration recommendations by the firm's Managing Director. Figure 5 summarises the results categorised by Knowledge, Attitude and Behaviour.

Table 2: Awareness level measurements [21]

Awareness	Measurement (%)
Good	75
Average	60
Poor	30

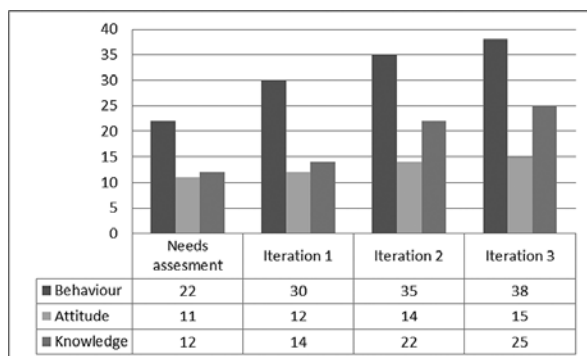


Figure 5: Results summary

The 78% awareness level in the 3<sup>rd</sup> iteration was satisfactory and there was no need for a fourth although it is advisable to run the process at least once a year as the skills and knowledge of the employees may become outdated.

It was possible to measure the effectiveness of the information security awareness training by using tools and methods outlined by Kruger and Kearney [21]. These enabled the firm to evaluate the extent to which awareness activities had impacted on behaviour, attitude, and knowledge and therefore, whether or not the initial training objectives had been met.

## 6. FINDINGS

This study confirmed that having and implementing an information security policy does not automatically guarantee that all employees will understand their role in ensuring the security and safeguarding of information assets. It is therefore critical to design and align an information security awareness campaign to the information security policy's high-level goals, objectives and requirements.

The findings of the study support the Theory of Reasoned Action (TRA), the Protection Motivation Theory (PMT) and the Behaviourism Theory (BT). Awareness campaigns were aimed at communicating the firm's stance (subjective norm) on information security, threat appraisal, coping appraisal and in an effort to mould the employees' attitude towards positive behavioural intention. The results showed that an increase in knowledge made a positive change in attitude and behaviour.

However it was discovered that even though initially their security knowledge levels were very low, the employees had a positive attitude towards securing the firm's information assets; however, they did not have the skills and knowledge to behave in a secure manner confirming that the risk to which employees expose a firm is indeed due to unintentional naïve mistakes as was revealed by literature.

What is disappointing is that although knowledge increased dramatically during the iterations, the increase in attitude was marginal. This could be because employees have a certain attitude towards the firm and this attitude cannot be altered by information security awareness alone.

This study revealed that information security awareness programs require the largest portion of the information security budget which should be channelled to the design and implementation of an information security awareness campaign. This supports the findings of Voss [46]. It was revealed that the general costs of running information security awareness campaigns and training can be divided into direct and indirect costs.



### Direct costs

- Salary/incentives for the security awareness coordinator or team;
- Training, including instructor fees and room rentals (in the case of classroom style training); and
- Materials, such as slides, web designing, videos, posters, hand-outs and gadgets.

### Indirect costs

- Time spent by other employees or departments involved in promoting security awareness; and
- Time spent by the target audience on courses and training.

Making use of e-learning campaign methods significantly reduced the costs of running the awareness campaign. Direct costs involved only the website designing cost, and the firm's in-house technician who was trained on updating and maintaining the website thereafter. Indirect costs reduced as employees took the courses during times they were not busy reducing the chance of productive time being lost.

While carrying out the action research the objectives were to refine and validate the process and change the behaviour of the employees at the particular SME. However, good information security behaviour cultivates an unpredicted information security culture. Hence it can be concluded that good information security awareness campaigns will ultimately result in a positive information security culture.

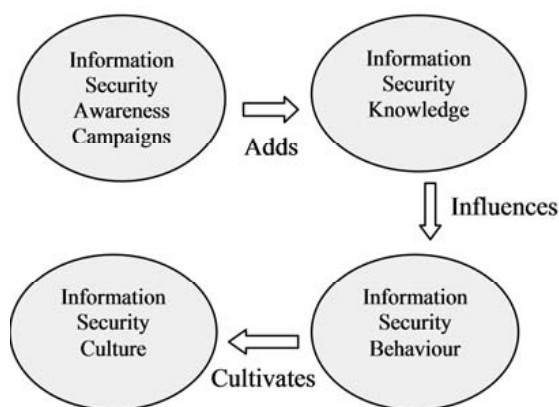


Figure 4: From information security awareness to information security culture

## 7. CONCLUSION

This paper was conceived against the backdrop of efforts made by SME firms to protect their information assets. This paper introduced an information security awareness process, which included behavioural intention models based on three persuasive theories i.e. Theory of Reasoned Action, Protection Motivation Theory and the

Behaviourism Theory. The research findings showed that information security awareness levels greatly influence behavioural intentions.

The information security awareness process and behavioural intention was verified through expert review by initially nine information security experts. Additionally, it was refined and validated through action research. After the action research, three more experts reviewed the process and model against the results from the empirical work to further validate them. The information security process yielded positive information security behaviour from employees at the action research host firm during all iterations. The researcher is therefore almost certain that similar results would be achieved if the process and model were put into effect at SMEs with similar characteristics to the one where the study was conducted.

The authors recognise that although e-learning is not a novel idea, it is a relatively new aspect in the field of information security and has great potential to increase e-security awareness initiatives. This study area will become more apparent as e-learning within information security expands. Relating to that, this study has been able to promote e-learning as an effective type of learning compared to the traditional classroom style of learning.

This research study explored the risks exposed by the uninformed naïve employee to SME firms' information assets. However, the risks exposed by the malicious insider as well as the outsider still require further exploration.

## 8. REFERENCES

- [1] R. Willson and M. Siponen. "Overcoming the insider: reducing employee computer crime through situational crime prevention", *Communications of the ACM*. Vol 52(9), September 2009. NY, USA.
- [2] BERR. "Information Security Breaches Survey" – *Technical Report*. Department for Business Enterprise & Regulatory Reform. April 2008. URN 08/788.
- [3] M. Fishbein, and I. Ajzen. *Belief, attitude, intention, and behaviour: An introduction to theory and research*, Massachusetts: Addison-Wesley, 1975.
- [4] A. Da Veiga & J.H.P. Eloff. "A Framework and assessment instrument for Information Security Culture," *Computers & Security*, Vol 29(2), pp 196-207, March 2010.
- [5] S. Furnell. "Malicious or misinformed? Exploring a contributor to the insider threat," *Computer Fraud & Security*. Vol 2006(9), pp 8-12, September 2006.
- [6] S. Furnell and K. Thompson. "From culture to disobedience: Recognising the varying user acceptance of IT security" *Computer Fraud & Security*. Vol 2009 (2), pp 5-10, February 2009.

- [7] J.L. Hale, B.J. Householder and K.L. Greene. The theory of reasoned action. In J. P. Dillard, and M. Pfau, *The persuasion handbook: Developments in theory and practice* (pp. 259 – 286). California: Thousand Oaks, 2003.
- [8] S. Milne, P. Sheeran and S. Orbell. "Prediction and intervention in health-related behaviour: a meta-analytic of protection motivation theory," *Journal of Applied Social Psychology*, Vol 30(1), pp 106-43, 2000.
- [9] Y. Lee and K.R. Larsen. "Threat or coping appraisal: determinants of SMB executives' decision to adopt anti-malware software," *European Journal of Information Systems*, Vol 18(2), pp 177-87, 2009.
- [10] T. Herath and H.R. Rao. "Encouraging information security behaviours in organizations: Role of penalties, pressures and perceived effectiveness," *Decision Support System*, Vol 47, pp 154 – 165, 2009.
- [11] A. Bandura. "Social cognitive theory of self-regulation," *Organizational Behaviour and Human Decision Processes*, Vol 50, pp 248-87, 1991.
- [12] G. Hinson, "Seven myths about information security metrics," originally published in *ISSA Journal*, July 2006, Available at: <http://www.noticebored.com/html/metrics.html> (Accessed Feb. 2010.)
- [13] B. Bulgurcu, H. Cavusoglu and I. Benbasat. "Information security policy compliance: an empirical study of rationality-based beliefs and information security awareness," *MIS Quarterly*, Vol 34(3), pp 523-48, 2010.
- [14] E. Hofstee. Literature Review. *In constructing a good dissertation*. Johannesburg: EPE, 2006.
- [15] ISACA. (2009). An Introduction to the Business Model for Information Security. California. Available from: <http://www.isaca.org/AMTemplate.cfm?Section=Deliverables&Template=/ContentManagement/ContentDisplay.cfm&ContentID=48017> (Accessed 3 February 2010).
- [16] S. Jenkins, R. Goal and D. Morrele. "Computer-assisted instruction versus traditional lecture for medical student teaching of dermatology morphology: A randomized control trial," *Journal of the American Academy of Dermatology*. Vol 59(2), pp 255–259, 2008.
- [17] K. Miller. *Communications theories: perspectives, processes, and contexts*, New York: McGraw-Hill, 2005.
- [18] P.A. Rippetoe and R.W. Rogers. "Effects of components of protection motivation theory on adaptive and maladaptive coping with a health threat," *Journal of Personnel Social Psychology*. Vol 52, pp 596-604, 1987.
- [19] E. Johnson. "Security Awareness: Switch to a better program," *Network Security*. Vol 6, pp 15-18, 2006.
- [20] M.E. Kabay. "Improving Information Assurance Education Key to Improving Secure(ity) Management," *Journal of Network and Systems Management*. Vol 13, pp 247-251, 2005.
- [21] H.A. Kruger and W.D. Kearney. "A Prototype for assessing information security awareness," *Computers & Security*. Vol 25(4), pp 289 – 296, 2006.
- [22] R.L. Krutz and D.V. Russell. *The CISP Prep Guide*. New York: John Wiley & Sons, 2001.
- [23] K. Miller. *Communications theories: perspectives, processes, and contexts*. New York: McGraw-Hill, 2005.
- [24] I. Ajzen. "The theory of planned behaviour". *Organizational Behaviour and Human Decision Processes*. Vol 50(2), pp 179-211, 1991.
- [25] R. Power. "CSI/FBI Computer Crime and Security," *Computer Security Journal*, Vol 17, pp 7-30, 2002.
- [26] R.E. Ricer, A.T. Filak, and J. Short. "Does a high tech (computerized, animated, PowerPoint) presentation increase retention of material compared to a low tech (black on clear overheads) presentation?" *Journal of Teaching and Learning in Medicine*. Vol 17(2), pp107–111, 2005.
- [27] C.L. Anderson and R. Agarwal. "Practicing safe computing: a multimethod empirical examination of home computer user security behavioural intentions," *MIS Quarterly*. Vol 34(3), pp 613-43, 2010.
- [28] R. Rogers. Cognitive and physiological processes in fear-based attitude change: a revised theory of protection motivation. In: J. Cacioppo, R. Petty, editors. *Social psychophysiology: a sourcebook*. New York: Guilford Press, pp 153-76, 1983.
- [29] M.A. Hogg and D. Abrahams, *Social identifications: A social psychology of intergroup relations and group processes*. Routledge London and New York, 1988.
- [30] S. Pahnla, M. Siponen and A. Mahomood. "Employees' behaviour towards IS security policy compliance," *Proceedings of the 40th Hawaii International Conference on System Sciences*, January, pp 3-6, Los Alamitos, CA; 2007.
- [31] R. Richardson. CSI Computer Crime & Security Survey. CSI, 2008. Available from: <http://www.cse.msstate.edu/~cse6243/readings/CSIsurvey2008.pdf> (Accessed 14 December 2009).
- [32] C. Russell. "Security Awareness - Implementing an Effective Strategy," SANS Institute, *InfoSec Reading Room*, 2002.
- [33] R.K. Sarkar. "Assessing insider threats to information security using technical, behavioural and organisational measures," *Information Security Technical Report*. Vol 15(15), pp 112-133, August 2010.
- [34] B. Schneier. *Schneier on Security*. New Jersey: John Wiley & Sons, 2008.
- [35] K.L. Smart and J.J. Cappel. "Students' perceptions of online learning: A comparative study," *Journal of Information Technology Education*. Vol 5, pp 201–202, 2006.
- [36] C. Pechmann, G. Zhao, M. Goldberg and E.T. Reibling E.T. "What to convey in antismoking advertisements of adolescents: the use of protection motivation theory to identify effective message themes," *Journal of Marketing*. Vol 6, pp 1-18, 2003.

- [37] J. Van Niekerk and R. von Solms. "Organisational Learning Models for Information Security," *Peer reviewed Proceedings of the ISSA 2004* enabling tomorrow conference 30 June – 2 July 2004, Gallagher Estate, Midrand.
- [38] J. Van Niekerk and R. von Solms. "Information Security Culture: a management perspective." *Computers & Security*. Vol 29(4), pp 476-86, 2010.
- [39] H. William. "Methods and techniques of implementing a security awareness program". *SANS Institute, InfoSec Reading Room*, 2002.
- [40] I.M.Y. Woon, G.W. Tan and R.T. Low. "A protection motivation theory approach to home wireless security". In: D. Avison, D. Galletta and J.I. DeGross, editors. *Proceedings of the 26th International Conference on Information Systems*, In Las Vegas, December 11-14, pp 367-380; USA; 2005.
- [41] M. Workman, H.H. Bommer and D. Straub. "Security lapses and the omission of information security measures: a threat control model and empirical test," *Computers in Human Behaviour*. Vol 24, pp 816, 2008.
- [42] P.A.H. William. "In a 'trusting' environment, everyone is responsible for information security." *Information Security Technical report*. Vol 13, pp 207 – 215, 2008.
- [43] P. Ifinedo. "Understanding information systems security policy compliance: An integration of the theory of planned behaviour and the protection motivation theory," *Computers & Security*. Vol 31(1), pp 83-85, 2012.
- [44] M. Elden and R.F. Chisholm. "Emerging Varieties of Action Research: Introduction to the Special Issue," *Human Relations*. Vol 46(2), pp. 121-142, 1993.
- [45] J. Cox. "Information systems user security: A structured model of the knowing-doing gap." *Computers in Human Behaviour*. Vol 28, pp 1849–1858, 2012
- [46] B.D. Voss. "The Ultimate Defense of Depth: Security Awareness in Your Company". *SANS Institute, InfoSec Reading Room*, 2001.
- [47] B.F. Skinner. *Science and human behaviour*. New York: Free Press, 1965.
- [48] G.S. Reynold. *A primer of operant conditioning*. (Rev Ed) Michigan: Scott, Foresman, 1975.
- [49] A.W. Staats, and C.K. Staats. "Attitudes established by classical conditioning." *The Journal of Abnormal and Social Psychology*. Vol 57(1), pp 37-48, 1958.

# MULTI-AGENT AUGMENTED COMPUTER VISION TECHNOLOGIES TO SUPPORT HUMAN MONITORING OF SECURE COMPUTING FACILITIES

Marius Potgieter\* and Johan Van Niekerk\*\*

\* School of Information and Communication Technology, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa 041 504 1111 E-mail: s208108589@live.nmmu.ac.za

\*\* School of Information and Communication Technology, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa 041 504 3048 E-mail: johanvn@nmmu.ac.za

**Abstract:** Humans are poorly equipped to perform repetitive tasks without adversely affecting the efficiency with which they are performing the task. One such task is the monitoring of CCTV footage to prevent the theft of, or tampering with, computing equipment. This paper introduces an approach towards security monitoring that uses a Computer Vision augmented with Speeded-Up Robust Features (SURF) as the catalyst to provide event-driven object detection to assist in securing an environment. A multi-agent artificial intelligence is used to improve the processing of event detection during the execution of these computer vision algorithms. The scenario of a secure computer environment is used to demonstrate the problems with current approaches and present an alternative to human monitoring using Computer Vision. The paper demonstrates that some of the physical aspects of information security can be improved through the use of SURF algorithms.

**Keywords:** Computer Vision, Features, SURF, SIFT, Pattern Recognition, Multi-Agent Theory.

## 1. INTRODUCTION

The modern world is a fast paced environment where access to information and the ability to respond to constant change is vital to organizational success. Nowadays, access to information is so crucial that some authors no longer see it as a competitive advantage, but rather as a must have commodity, similar to electricity [1]. It is thus vital to protect organizational information against harm or loss. The process of protecting information is known as information security.

Information security is commonly implemented in the form of various information security controls. These controls are described in International Standards such as ISO/IEC 27002 [2] or ISO/IEC 13335 [3]. These standards provide three broad categories of Information Security controls; namely physical, technical and operational controls.

Physical and technical controls are often used in combination with each other. For example the protection of mission critical computing facilities often take the form of a computer room with various physical and technical controls to provide security. The controls typically include access control of a technical nature as well as physical monitoring performed by humans monitoring video cameras through a CCTV system.

Unfortunately humans are poorly equipped to perform repetitive tasks, like the monitoring of CCTV footage, without adversely affecting the efficiency with which they are performing the task [4]. This paper demonstrates how computer vision technologies, supported by artificial intelligence constructs, can be used to augment such

human monitoring of CCTV footage in order to reduce the risk of a human not detecting a security incident. Such computer augmented monitoring can reduce the risk posed by possible human complacency.

The remainder of the paper will provide a brief introduction to current computer vision and will then introduce a scenario where the physical monitoring of a secure computer environment is important. The paper will then introduce Speeded-Up Robust Features (SURF) algorithms and demonstrate how such techniques can be used to augment current camera-based monitoring of such secure environments in order to reduce reliance on human operators. The role(s) of artificial intelligence constructs within the computer vision process is highlighted.

## 2. COMPUTER AUGMENTED CCTV MONITORING

The use of computer vision technologies to support the human monitoring of CCTV footage has become well established. According to [5] the UK had more than 4 million such cameras in 2007. Many of these were used in public spaces to track suspicious persons or events [5].

For example, the real time monitoring of CCTV to improve security at mass transport facilities, like airports, have become commonplace. Such monitoring is often supported by facial recognition systems. Currently such systems rely on computer vision technologies such as Active Appearance Model (AAM), Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Support Vector Machines (SVM) [6] [7] [8].

In general facial recognition systems use either feature based algorithms or appearance based algorithms [8].

For facial recognition appearance based algorithms are considered the better choice because feature based algorithms still have many limitations. However, for event detection, where the recognition and tracking of specific objects are important, feature based algorithms would be preferable since these approaches provides more accurate identification of separate entities.

Speeded-Up Robust Features (SURF), which this paper focusses on, is one such feature based algorithm that could potentially be used to improve the automated detection of security incidents from CCTV footage. The next section will briefly introduce a secure computer room scenario that will be used by the remainder of the paper to demonstrate how SURF can be used as part of a multi-agent approach to identify security incidents.

### 3. SCENARIO

#### 3.1 Introduction

For many organizations the physical security of computing infrastructure is still considered the first line of defence towards information security. In many such secure computing facilities physical access to computing re- sources by unauthorized persons is prevented through the use of access control mechanisms. The facilities are further secured through the use of CCTV cameras to monitor events within the facility. This monitoring is typically performed by humans who pro-actively try to identify and prevent events which could lead to security incidents.

Figure 1 shows such an example computer facility containing equipment that needs to be monitored. Both access to and possession of equipment is generally controlled.



Figure 1: Computer Facility

In this facility the access control systems provide the technical control of the physical locking system that allows entry to the secure environment. The supporting operational control would be to allow only certain people access to this area.

The video cameras provide physical monitoring of the area using video cameras usually on a CCTV system. The video streams can also be recorded and stored for archival purposes. To enforce operational controls based on the footage a person needs to monitor it.

#### 3.2 Problems

One problem with many security controls are that they do not provide fool proof solutions to the outcomes required by operational controls. In the case of access control systems you would have people granting access to the secure environment to someone that would not normally have access. Granting access to unauthorized people exposes the secure environment to a variety of threats.

Similarly video monitoring needs to be performed by a human to act on threats to the environment. Depending on the amount of footage to monitor this can create unrealistic assignment of human resources to such a task. Even this does not provide fool proof monitoring since threats to the environment is the exception and not the norm. Research has shown that the person monitoring can get complacent and not be alert for possible threats due to a limited capacity to maintain attention span during repetitive tasks [4]. According to Harris humans are poorly equipped to detect low-sign-to-noise-ratio signals embedded in the context of varying background configurations.

These problems only outline some vulnerability that controls need to deal with. Global surveys found that insider threats pose the greatest danger to controls placed to secure assets [9]. This provides a greater risk to security from an insider than that from external sources. Shifting the responsibility of monitoring from a human to that of a computer based monitoring system can firstly reduce human error induced by factors such as fatigue and complacency. Computer based monitoring will also reduce the risk posed by insider threats. Such computer based solutions does not necessarily have to replace humans, but can augment human monitoring in order to reduce fatigue and improve monitoring efficiency.

The use of algorithms to assist with computer monitoring of video data is often hampered due to imprecise data. This includes scale and perspective variations that are caused by viewing an object from difference distances and angles; it also includes factors like a change in illumination. [10]. This paper will suggest using Speeded-Up Robust Features (SURF) as a tool to augment computer vision to allow for more robust object tracking and monitoring. SURF algorithms have been shown to be more robust than other approaches when faced with such varying or imprecise data [11]. The SURF algorithms are further refined through the use of Genetic Algorithms and Artificial Neural Networks during the process.

## 4. IMPLEMENTATION

### 4.1 Introduction

Object recognition has been a focus point in machine vision research for the past decade for the advancement of robotics, security, defence technology and other related fields. The processes mentioned in the different research materials will be examined to provide insight into different aspects of object recognition that will ultimately form the basis on which the holistic approach to a generic object recognition engine that will facilitate a secure monitoring system.

The process of scene modelling and recognition has been thoroughly explored and defined where interest points are detected, descriptors are generated and features generated with those descriptors. These features describe an image and thus the objects within this image. The algorithm that has been used with great success in computer vision is called Scale-invariant feature transform (or SIFT) [12]. It should be noted that most algorithms within computer vision including SIFT are performed on grey-scale images.

The following steps are followed to apply a SIFT algorithm on an image to support Computer Vision:

1. Image Processing – Collect image from source and convert to grey-scale
2. Image Analysis – Interest Point generated on source image and features generated around them
3. Pattern Recognition – Compare an match interest points/features from source with objects within security database being monitored
4. Event Detection – Attach events to objects that have been interacted with in source



Figure 2: Image with a few interest points with their scale and rotation angle drawn

This variant of feature descriptors also describes an image at set interest points. The interest points are selected using algorithms that define areas within an image that looks most distinct and can be well defined as shown in Figure 2. Around these interest points the feature algorithm calculates the maxima and minima in

the difference of Gaussians function. A difference of Gaussians function discards details of an image as the image gets blurred at different levels (also referred to as scale-space) and compare with the original image. The change that occurs between images will be the Gaussian Distance [13].

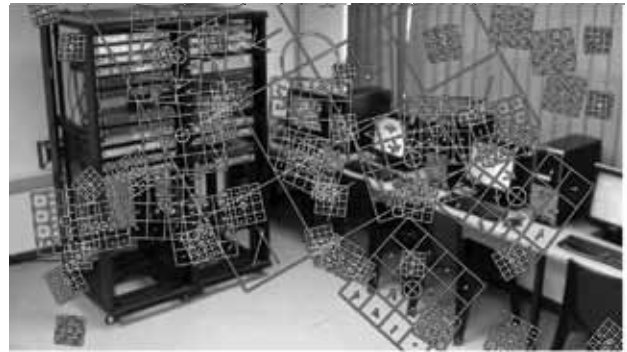


Figure 3: Interest points with their corresponding histograms

By obtaining the maxima and minima of this function a feature that consists of vectors that define the gradient intensity and general direction around those points is obtained. Each individual vector can then form part of a subsection of the feature that can be defined as a histogram as shown in Figure 3. Each subsection can be summed into a general direction and scale which provides a summary of the feature.



Figure 4: Single interest point and histograms defining the summed area around it defined by vectors

This histogram, which contains the distribution of vectors, defines the feature as a list of integers that corresponds to the histogram values. It is this list that is used to compare to other images, matching features to obtain a positive match (Figure 8). The matching process

uses nearest neighbour and Best-Bin-First search techniques to increase the speed of indexing and matching.

The variant of features this paper will suggest to be used with the proposed computer vision technology is called Speeded-Up Robust Features [14]. They are basically calculated the same way as SIFT features but have been improved to increase performance when performing matches. The way it achieves this is by using integral images, areas that have been summed within a grid of values. The comparison can be seen with SIFT with the various histogram values that form the feature but in the case of SURF these subsections are summed to form so-called Haar-like features. Haar-like features are similar to Haar wavelets that are square shaped functions that form predefined variations of the feature. As an example, a simple rectangular Haar-like feature can be defined as the difference of the sum of pixels of areas inside the rectangle [15]. Such a feature can be at any position and scale within the original image. The simplicity of the features makes it easy to evaluate and thus gives it a speed advantage compared to the more sophisticated SIFT variation. The speed at which the computer vision algorithm performs the matching is important since our goal is to perform real-time detection of events within an environment containing multiple objects that needs to be detected and matched with the database of objects.

The use of categorization before classification in object recognition is suggested to decrease the pool size of images that is used to compare descriptors. The use of matching kernel functions to compare descriptor vectors significantly reduces processing required to describe an image by using this kernel in a support vector machine (SVM) [16]. A SVM is a concept in statistics and computer science for a set of related supervised learning methods that analyse data and recognize patterns.

The settings for determining the threshold of these images is traditionally manually set depending on its use. This does not favour a generic system that could cater for a variety of image conditions. Investigation into using genetic algorithms with image processing has determined that an adaptive learning system can be used to adjust these parameters using genetic algorithms with a neural network [17].

Image segmentation is an important part of object recognition which defines the process of separating individual objects from an image containing multiple objects. This allows for feature recognition to occur on an object-to-object basis without having to account for the background or surrounding objects [18].

The above mentioned related research will be incorporated in one form or another to create the Computer Vision Engine. The resulting engine will incorporate features that provide the performance enhancements required to process high quality images as

well as create a generic engine capable of processing any range of image variations. Additional functionality such as learning algorithms (genetic algorithms and neural networks) as well as incorporating image segmentation and categorization that focuses on an adaptive approach to object recognition was mentioned. The use of multiple learning algorithms to approach a complex problem like this is called a multi-agent approach.

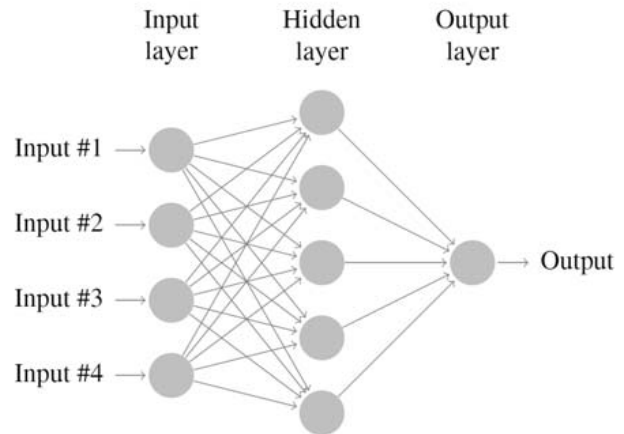


Figure 5: Conceptual Structure of an Artificial Neural Network

#### 4.2 Multi-Agents

The concept of a multi-agent system lies with perceiving an environment in a specific way and then acting on that perception. This is important in trying to evaluate a complex environment that cannot be perceived using a single view-point. These views can be gathered from multiple agents that supply the data on the environment and then processes this data either individually or as a group by means of a specific applied process with an end goal in mind. There typically exist three types of agents [19]:

1. Passive: plays a simple role within a system (does not supply additional input to system)
2. Active: provides an output within the system with a specific goal in mind e.g. Camera provides visual feedback
3. Complex: involves calculations/algorithms applied to input from environment to supply required goal (e.g. feature generation within the computer vision system)

The result of the agents can be either is used to provide actions to automate a process in the system or to supply another agent with data. This process can create a sense of intelligence within the process model where the agents act according to what information they are given. The use of AI techniques within a multi-agent system has been suggested to validate the perceived intelligence by providing intelligence with active-learning concepts within itself e.g. Genetic Algorithms It will not be in the

scope of this paper to identify the characteristics of different multi-agent systems but to implement them as a holistic approach to computer vision. This approach to Computer Vision will rely on different agents to perceive our secure environment and provide actions within our system to improve performance or supply additional functionality. The three concept agents we will be looking at is GAs, Neural Networks as an AI agents (complex agents) and depth perception provided by sensors e.g. Microsoft Kinect (active agent). It will be the authors' suggestion that a multi-agent approach of distributing provides robustness to the system.

*Genetic Algorithm:* Genetic Algorithms (GA) describe a biological process where a solution to a problem is generated through techniques common to natural evolution like inheritance and mutation. In a genetic algorithm a semi-randomly generated population called chromosomes is supplied. These will go through a "breeding" process where features of two chromosomes join through inheritance, selection, mutation or crossover. The newly formed chromosomes will be selected using a fitness function that generally describes how close to an ideal solution they came and the process repeats with the new fitter chromosomes to start the mating process. Each time this happens we call it a generation, where a typical genetic algorithm can go through thousands of generations towards a satisfactory fitness level have been reached. A genetic algorithm typically contains a termination condition which typically is that it has reached a solution or as close to a solution as it can establish or a fixed amount of generations have been reached [20].

*Implementation:* Computer Vision contains various parameters which supply filters as criteria for adjusting an image (frame) to extract features like edges, blobs (areas that differ in colour or brightness compared to surroundings) and interest points, to name a few. These filters depend highly on these parameters to output accurate features. Unfortunately because of environmental changes such as light-level, capture device quality and other factors we cannot simply supply static parameters to these filters. The proposed system thus uses genetic algorithms to evaluate a scene or environment being monitored specifically during the processor's idle times when few or no events are being recorded. This provides a way to calculate a genetically evolved parameter values that will supply the computer vision system with greater flexibility to adapt to changes in our environment. The use of a GA in this case is highly effective since the evolutionary process is ideally suited to a constantly changing environment.

*Artificial Neural Network:* Artificial Neural Networks (ANN) is an artificial intelligence technology based on a biological paradigm. Initially ANN described a biological architecture between individual neurons which can be extensive and extremely complex. The connections between these neurons are called synapses that provide

interconnectivity between neurons. A signal that moves between neurons along synapses activates neurons. The state of neurons activating is referred as neurons firing. This signal can either be increased or decreased depending on the effect of the neuron on the signal. We can observe a clear input and output within this neural network with the signals changing depending on the effect of the neurons.

Within a computer (software) model this can be represented as different nodes that apply functions to an input (signal) and provides a clear output based on the effect of the nodes on the input values. Each node can be interconnected with multiple other nodes (network) where the signal moves between the nodes changing the output based on what the nodes function applied to the input (referred to as the hidden layer). An example of the interconnected nature of a typical ANN is given in Figure 5. The actual "intelligence" of such a network lies in the interconnected nature of the nodes and the weights assigned to each incoming connection to a specific node. Once trained, an ANN is an extremely efficient decision structure since it can accept multiple inputs and return a single output of a specified format [21].

*Implementation:* The role of the ANN within our Computer Vision system is to provide analysis of features that were generated within our Image Processing Phase and needs to be analysed (Image Analysis Phase). Each feature (input) moves through our neural network and gets weighted depending on different algorithms to determine their weighting in how distinct they describe our image in comparison to other features. This provides the system with a method of eliminating features that would otherwise increase processing time needed in identifying objects and increasing importance of features that can be used as indexers to increase the speed of the system.

#### 4.3 Computer Vision-Technology

Feature descriptor algorithms generally consist of interest point detection followed by descriptor calculation. The proposed algorithm is to include steps between the previously defined steps to increase its performance and accuracy. To simplify the process these steps have been classified into phases.

The following phases represent the changes to the process:

1. Image processing, deals with the preparation of the image for feature extraction which will include creating separate threads for processing each image in smaller parts;
2. Image analysis, where the interest points will be detected which will provide logical objects to focus the detection on;



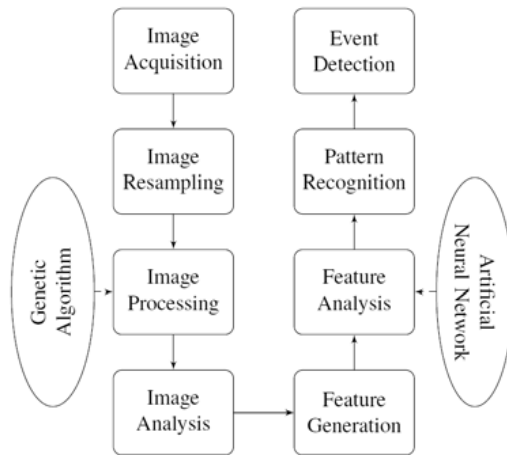


Figure 6: Computer Vision Process Flow

3. Pattern recognition, matches the detected feature descriptors to the predefined library of objects through classification then recognition; and
4. Event Detection will provide information about the detection object either as a still image or video in relation to its position, relative distance to other objects and other parameters which enable an event driven object engine.

**Image processing:** This phase of the process involves acquiring the image from the camera. The ability to locate objects in a two-dimensional space image is one of the main obstacles of object recognition. The use of canny edge detection is used to distinguish between various objects within a static image. The Canny Edge Detection algorithm takes the derivative of an image to find the gradients, this determines the directional gradients. The amplitude of the gradients is used to either include or exclude the gradient as part of the contours forming the edge. Once edges of objects are found pass these objects through the image analysis phase that will implement the SURF algorithms.

**Acquire Video Feed as Images:** This step simply acquires sets of images from the video source. Video cameras possess different levels of frame rates rated in Frames per Second (fps). Since most cameras work between 20-30 fps it was assumed that within one second there will not be sufficient difference in the frames caught in that second to justify processing each frame. Thus only one single frame is taken from each acquisition period that will be set manually by the discretion of the user.

**Image Resampling and Combination:** Since the system aims to apply a generic approach to video monitoring, the image is resampled into standard resolutions and quality to facilitate the analysis phase. Image transformation that increases the efficiency of key point matching will be applied using different image filters [12]. The use of genetic algorithms to adjust the parameters for the image filters has been suggested when dealing with image segmentation and could supply a viable alternative to

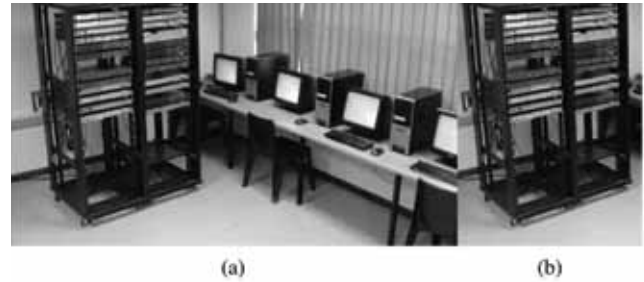


Figure 7: (a) Contains frame as seen from video source. (b) Image stored in database and categorised as the object it represents and its ownership. This image will have its features compared to that of the frame in the video.

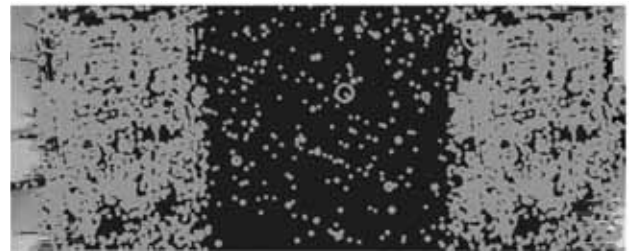


Figure 8: Interest points are matched between the video frame and this one object in the database containing other objects. As can be seen here there is clearly a high match for the object.

manually setting parameters of other types of filters [18]. The approach followed in this research used this suggestion from [18] as part of the multi-agent approach towards refining the SURF algorithms. Figure 6 demonstrates the overall computer vision algorithm process flow. The specific place in this process flow where the above-mentioned GA fits in is also indicated on the figure.

Each captured frame will be packaged with all its associated images that have been resampled (in this case to grey-scale versions of the frames).

**Image analysis:** The SURF algorithm is used to form the integral image (summed area table) which uses subsets of the image in grids and sums them [22]. This increases the efficiency of Hessian Blob Detectors that detects points or regions that differs from their surroundings. Using the box filters the system locate interest points that would best fit that given box filter in a specific Hessian determinant Figure 4. Calculating these box filter responses is extremely expensive on processing power, requiring 126 million lookups for a 1280 x 1024 image. It has been suggested that one can decrease the bottleneck in this multi-pass approach by sharing the results across each pass [23]. This can be improved by creating a parallel computing environment where the power of current processors are used to run all passes at once and select the best result from the output. Once the image have been reduced to an integral image, consisting of interest points, the resulting output can be converted into

Haar Filters that will form part of each individual SURF-descriptor. In the multi-agent approach of this research the resulting outputs from the Haar Filters were fed into an ANN which performed the final discrimination to determine the usefulness of descriptors.

These SURF-descriptors collectively will form the images pattern recognition signature that will be used to compare with trained objects in the database.

*Pattern recognition:* Pattern recognition is performed by comparing the set of SURF-descriptors with those of the trained images. When a match occurs that is at an acceptable tolerance level the match gets accepted as the object being observed. This method however gets exponentially more demanding on processing power depending on the size of the trained image database. The use of a SVM (Support Vector Machine) to classify local descriptors in order to perform object categorization can significantly reduce the amount of objects the process needs to compare [16]. The kernel function that categorizes the input descriptors creates a structure to identify an object without referring to the individual descriptors. This allows the use of an index system to narrow the parameters of the set used for comparison. If the observed object is a close match to one of the indexed categories, the collection assigned to it will be loaded into memory and each individual set object will be compared to make a more accurate recognition. The result of the comparison should have more matches to an object within the loaded set, if this is not the case the object should be trained as an object that has not yet been recorded. Additional information about trained objects can be added at this stage that could be used to assist additional classification. This could also assist in creating a vocabulary tree that would provide scalability of the database [24]. Potential automatic classification of an object can be implemented through the use of online image galleries such as Google Images [25].

*Event Detection:* The matched object contains a series of matching features which can be plotted on a homography matrix. The homography matrix provides detail about the transformation between the observed image and the object in the database that has been trained. By examining the homography matrix the mapped transformations can be used to determine the object's general rotation, location and alignment to its trained version as well as other objects that have been matched. The state of these conditions can be used to provide event based triggers that can be used to create event handlers. These handlers assist in creating interactive object recognition systems as well as reducing the continuous recognition of the same object by tracking its current position and excluding it from the pattern recognition phase.

It is at this point where operational controls can be applied to objects that have been observed in order to provide security of those objects. In the case of the

scenario of a secure computer facility one will have an object database of computer equipment and some generic human models to track. Objects and the humans that interact with them will provide events that can be used to alert users of possible security threats. This automated process provides security monitoring an alternative to current approaches where a user was needed to be on the lookout for all these events on multiple video sources. The result of this suggested system is to only alert users of the events detected by the system that was flagged by the system as a possible security threat. The flagging process is not discussed within this paper but forms part of algorithms matching events detected in the computer vision system to security operational controls.

## 5. DISCUSSION

This research was based on the premise that SURF requires less processing power than SIFT which means that SURF should be better at processing real-time footage. The previous section demonstrated how SURF algorithms could be used to perform such image processing and why this would be ideal for implementing a security monitoring system. During the course of the research SURF technology was used in a program to analyse physical events in a simplistic secure computer facility. This prototype implementation demonstrated that the use of SURF technologies for this type of system is viable. The system was used to successfully detect various events in the environment, such as the one shown in Figure 9. After an initial testing it was found that the SURF algorithms would need additional refinement as the amount of "clutter" in the environment increases. Genetic algorithms and Artificial Neural Networks have been used to make the SURF algorithms more adaptable. The GA was used primarily to "evolve" more refined parameters for the SURF filters during idle processor cycles. The ANN, on the other hand, was used to provide an extra layer of discriminatory logic to improve the outputs of Haar-like filters.

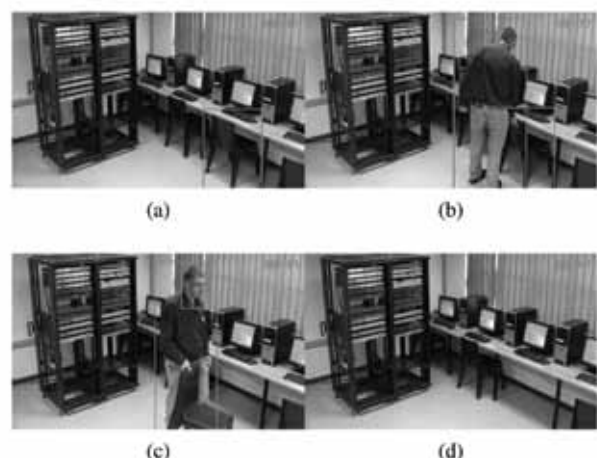


Figure 9: Four frames from a video taken at certain time sequences illustrating event tracking.

The prototype demonstrated that it was viable to detect actual events using real-time footage. The events still need to be interpreted to determine whether or not they are real security incidents or normal behaviour. Such interpretation has not been automated yet. However, it should be possible to do such interpretation using adaptive AI technologies similar to the Genetic Algorithms and Neural Networks that has already been used in the current implementation. This will form the basis of future research.

## 6. CONCLUSION

The paper presented an implementation of a computer vision technology to augment human monitoring of secure computing facilities through the use of SURF. The computer vision algorithms utilized during this implementation incorporated artificial intelligence constructs to further refine the computer vision algorithms. It was established that current systems that rely on users to classify possible threats over an extended period of time is not recommended. The information provided by the object recognition process provided by SURF supplied the required event detection needed to be used within security monitoring scenarios. These events can be clearly linked to operational controls used to secure objects within a secure environment to supply alerts to a user to enforce those controls. The work in this paper demonstrated that computer vision technologies, specifically SURF algorithms, can play a role in improving the enforcement of operational controls in secure computing environments. As far as could be determined no previous work has demonstrated the use of SURF algorithms for the detection of events from CCTV footage. This step towards an augmented security solution that can be used in diverse situations is clearly needed in an age where security of many assets, including human lives, is of utmost importance.

## REFERENCES

- [1] N. G. Carr, "IT Doesn't Matter," *Harvard Business Review*, vol. 81, no. 5, pp. 41–49, 2003.
- [2] "ISO/IEC 27001:2005, information technology - security techniques - information security management systems-requirements," 2005.
- [3] "ISO/IEC 13335:2004, Information technology - Security techniques - Management of information and communications technology security," 2004.
- [4] D. H. Harris, "How to Really Improve Airport Security," *Ergonomics in Design: The Quarterly of Human Factors Applications*, vol. 10, no. 1, pp. 17–22, Jan. 2002.
- [5] N. Dadashi, A. Stedmon, and T. Pridmore, "Semi-automated cctv surveillance: The effects of system confidence, system accuracy and task complexity on operator vigilance, reliance and workload," *Applied Ergonomics*, no. 0, pp. –, 2012.
- [6] W. Zong and G.-B. Huang, "Face recognition based on extreme learning machine," *Neurocomputing*, vol. 74, no. 16, pp. 2541 – 2551, 2011, Advances in Extreme Learning Machine: Theory and Applications - Biological Inspired Systems. Computational and Ambient Intelligence. Selected papers of the 10th International Work-Conference on Artificial Neural Networks (IWANN2009).
- [7] E. Gumus, N. Kilic, A. Sertbas, and O. N. Ucan, "Evaluation of face recognition techniques using pca, wavelets and svm," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6404 – 6408, 2010.
- [8] U. I. Bajwa, I. A. Taj, and M. W. Anwar, "A unified classifier for robust face recognition based on combining multiple subspace algorithms," *Optics Communications*, vol. 285, no. 2122, pp. 4324 – 4332, 2012.
- [9] C. Colwill, "Human factors in information security: The insider threat Who can you trust these days?" *Information Security Technical Report*, vol. 14, no. 4, pp. 186–196, 2009.
- [10] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," *Computer Vision ECCV 2006*, vol. 3951, no. 3, pp. 404–417, 2006.
- [11] L. Juan and O. Gwun, "A Comparison of SIFT , PCA-SIFT and SURF," *Image Processing*, vol. 3, no. 4, pp. 143–152, 2009. [Online]. Available: <http://www.cscjournals.org/csc/manuscriptinfo.php?ManuscriptCod>
- [12] D. Lowe, "Object recognition from local scale-invariant features," *Proceedings of the Seventh IEEE International Conference on Computer Vision*, pp. 1150–1157 vol.2, 1999.
- [13] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 207, no. 1167, pp. 187–217, 1980.
- [14] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [15] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, 2001, pp. I–511 – I–518 vol.1.

- [16] J. Eichhorn and O. Chappelle, "Object categorization with SVM: kernels for local features," *Biological Cybernetics*, vol. 190, no. 137, pp. 365–382, 2004.
- [17] B. Bhanu and S. Lee, "Adaptive image segmentation using a genetic algorithm," *Systems, Man and Cybernetics*, 1995.
- [18] K. Hammouche, M. Diaf, and P. Siarry, "A multilevel automatic thresholding method based on a genetic algorithm for a fast image segmentation," *Computer Vision and Image Understanding*, vol. 109, no. 2, pp. 163–175, 2008.
- [19] Y. Kubera, P. Mathieu, and S. Picault, "Everything can be agent," in *Autonomous Agents Multiagent Systems/Agent Theories, Architectures, and Languages*, 2010, pp. 1547–1548.
- [20] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009.
- [21] A. P. Engelbrecht, *Computational Intelligence: An Introduction*, 2nd ed. Wiley Publishing, 2007.
- [22] P. Viola and M. Jones, "Robust Real-time Object Detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2001.
- [23] T. B. Terriberry, L. M. French, and J. Helmsen, "GPU Accelerating Speeded-Up Robust Features," *Proc of 3DPVT*, no. 2, pp. 355–362, 2008.
- [24] D. Nister and H. Stewenius, "Scalable Recognition with a Vocabulary Tree," *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Volume 2 CVPR06*, vol. 2, no. c, pp. 2161–2168, 2006.
- [25] R. Fergus, P. Perona, and A. Zisserman, "A Visual Category Filter for Google Images," *Proc 8th European Conf on Computer Vision*, vol. 3021, pp. 1–14, 2004.