Towards an Automated Security Awareness System in a Virtualized Environment

William Aubrey Labuschagne[1] and Mariki Eloff[2]
[1]Defence, Peace, Safety and Security, Council for Scientific and Industrial Research, Pretoria, South Africa
[2]School of Computing, University of South Africa, Pretoria, South Africa
wlabuschagne@csir.co.za
Eloffmm@unisa.ac.za

Abstract: A majority of African Internet users do not have access to the Internet. The lack of infrastructure in rural areas affects Internet usage. Since costs are high and the bandwidth low, these factors encourage users to access the Internet using shared resources. This is an efficient solution to access the Internet. However users might not be aware of the security threats that exist on using shared resources. Many companies provide security solutions to automatically protect resources on the network and security awareness training to users. This ensures that users are aware of the security threats and provide methods to mitigate them. These measures are useful in a corporate environment where funds exist to enable these security solutions. Public platforms, for example Internet Cafes and schools, allows multiple users to access the Internet using shared resources. This implies that multiple people will use the same computer to perform required tasks. Numerous security threats exist within the Internet sphere that could affect users utilizing shared resources these include but are not limited to viruses, keyloggers and phishing attacks. This shared environment could provide a platform that promotes the spread of virus infections. Users using these platforms should be made aware of these threats and monitor the effectiveness of the security awareness campaign. This paper proposes a system used to address these issues from a single platform. The Shared Public Security Awareness (SPSA) system is an automated virtualized system used to determine the current security awareness levels of users on a shared platform accessing the Internet. The system uses virtual machines to provide users with access to the Internet, assess the security awareness levels of the users, determines if any web browser components were infected by web based malware during browsing sessions, provides users with access to security related material affecting the users and provide reports on online behaviour. This paper evaluates the proposed SPSA system as a mechanism to conduct a security awareness campaign in a shared resource environment while providing a capability to analyze the online behaviour of users that affects the security of this environment.

Keyword: internet cafes, security awareness, security training, virtualized environments, cyber literacy, internet

1. Introduction

The Internet provides a vast range of information resources and services which form part of everyday life. Usages include but are not limited to searching for information, conducting business, paying bills and the purchase of goods. Moreover the development of human capital has been identified as an important economical performance indicator in rural areas (Agarwal, Rahman & Errington 2009).This can be attained with access to knowledge available on the Internet. However the high adoption and use of the Internet by citizens introduced an opportunity for cyber criminals to utilize this platform to coordinate cyber attacks with the intention to cause damage. Kim identified a comprehensive list which includes loss of money, defamation, invasion of privacy, physical harm, loss of time and psychological damage (Kim et al. 2011). Most companies provide security measures against these attacks for their employees. In most instances employees require to attend security awareness training programs to equip them with strategies on how to mitigate these cyber attacks when encountered. Furthermore the network infrastructure is secured with expensive security solutions. Therefore, people working at companies are best equipped against cyber attacks. Users in rural areas are in a disadvantages position. In most instanced these users do not have ownership of resources to access the Internet. The cost of access to the Internet and equipment inhibits ownership of resources like computers. The need to access the Internet was addressed with entrepreneurial initiatives which provide access to the Internet with the use of shared resources. This implies multiple users using the same computer to access the Internet. An example of this implementation are schools and Internet cafes. However, users sharing the same resources could assist in the spread of malware infections. In the event of discovering a malware infection at these establishments, the services provided need to be suspended which has an effect on revenue for the owners. Another issue which could be encountered at these establishments is security literacy.

Most of these users are not aware of the cyber threats that are devised and deployed by criminals. Security awareness programs are used to educate the users and provide them with measures to

identify and mitigate the threats encountered. Grobler studied the cyber awareness initiatives in South Africa (Grobler et al. 2011). She reported on initiatives by the Council for Scientific and Industrial Research (CSIR), the University of Pretoria (UP), the University of Fort Hare (UFH) and Nelson Mandela Metropolitan University (NMMU). The CSIR collaborated with the University of Venda to raise security awareness in the rural areas by developing content which addresses cyber security topics and training community members which then in turn will train the community. The UP project, PumaScope, equips students with the required security knowledge to educate scholars at identified schools. UFH tested the proficiency levels of the user in a particular area. NMMU addresses educating users through the use of games and eLearning platforms to provide access to security awareness content for a wider audience. A need has been identified to provide an automated platform which incorporates the core ideas of the mentioned initiatives a platform that could be used to determine the proficiency levels of the users and provide access to resources to improve security awareness in rural areas.

This paper looks at the design of an automated tool, Shared Public Security Awareness (SPSA) system, which promotes security awareness in rural areas where the community uses shared computer resources to access the Internet. These resources can be located at schools or Internet café where access to the Internet is provided through the use of shared computers. Establishments would be used throughout the paper that references the communal area where the shared computer resources are located. The deployment of the SPSA system addresses three primary functions: The first function is to provide the capability to conduct a security awareness program which consist of assessing the literacy of the users and deliver the security awareness topics to the users. The second function analyzes the online behaviour of users and the collection of malware which would assist in developing strategies which addresses the security threats encountered at these establishments. The third function provides a turnkey solution which automates the functionality of the SPSA system with limited intervention from personal to administrate the system.

The rest of the paper is organized as follows: section 2 summarizes research related to the component identification of the SPSA system. The main contribution: the design of the SPSA system is outlined in section 3. Conclusions and future work are discussed in section 4.

2. Related research and underlying concepts

The SPSA virtualized and collection system requirements are discussed in this section. These establishments provide resources which enables user's access to the Internet through the use of web browsers. Cyber attackers have adopted attacking strategies which include automated exploitation of computer systems without the intervention of the user. The resources used by these establishments must be protected against possible attacks originating from the Internet. Also a mechanism is required to identify the threat and evaluate the actions performed by users which initiated these attacks. The system should exhibit the following capabilities:

- A robust and automated architecture which ensures availability and configurability of the system. This is achieved with the implementation of virtualization and customization of existing systems (See Section 2.1).
- The identification of threats originated from users visiting malicious web sites, accomplished with the collection and analysis of data generated during browsing sessions (Section 2.2).

2.1 Virtualization, automation and customization

The SPSA system underlying architecture consists of virtual machines. Bell defines a virtual machine as software that functions as a computer without physically being a computer (Bell, Lintumaa 2011). The use of virtual machines provides numerous of advantages.

The implementation of virtualized environments is cost effective. England proposed a model for deploying virtual machines as a securing mechanism for the enterprise desktop (England, Manferdelli 2006). Some organizations require users to conduct classified work. In these organizations the users will be provided with two physical computers: one to conduct normal duties and the other for classified duties. This is not cost effective. The use of virtual machines would allow both functionalities to be conducted within a virtualized environment and provide the required security measures.

Virtual machines can be controlled programmatically with the use of scripting language which automates the process of operations which include start-up and shutdown. Light proposes the use of scripts to control virtual machines within an automated sandbox (Ligh et al. 2010). He also described the malware analysis cycle with the use of virtualization which is supported by Harlan (Harlan 2005). The cycle described by Light is adapted for the SPSA system. A baseline virtual machine is created. A copy is made of the baseline virtual machine and then loaded daily for usage at these establishments. This will ensure that uninfected virtual machines are deployed for use every day. It also provides the opportunity to examine the virtual machines for possible infections; this is achieved by storing the virtual machine used during the day.

The added benefit of virtual machines is the efficiency of restoring to a state which users can use to access the Internet after malware infections. An environment which uses physical machines requires reinstalling the operating system after a malware infection. During this period the establishment cannot conduct business. The use of virtual machines minimizes the period of inactivity. Gold reported in 2007 of cyber attackers targeting virtualization (Steve 2007). Some malware is virtual machine aware which implies that the malware would not execute in the virtual machine environment (Zhu, Chin 2007). The malware writers added this feature to protect the malware against virtualised environments used by malware analysts. This could be beneficial to the establishments and reduce the infection rate due to the inactivity of the malware.

Users at establishments require access to the Internet. A customizable user management system would be required to control the sequence users follow to access the Internet and expose features of the SPSA system to the users. These features include the completion of a questionnaire and coverage and comprehension of the security awareness topics. The continuous exposure to security related content contributes the success of a security awareness program (Kruger, Kearney 2006). The SPSA system is designed to present security awareness content to the user before accessing the Internet thus reminding the user of safe practises against cyber attacks. Easyhotspot is an alternative solution for hotspot billing system released under the GNU general public license which implies that the software could be modified with the needed requirements of the SPSA system (The EasyHotspot team 2007). Easyhotspot consists of a user management system which allows users to access the Internet through the portal (See **Figure 1**). Modifications to the portal would presents users with access to the security awareness content or the questionnaire.

Figure 1: EasyHotspot management system

2.2 Threat collection and analysis

Abraham summarised an overview of social engineering malware which entices users to perform detrimental actions which could infect the computer system (Abraham, Chengalur-Smith 2010). The malware utilizes numerous avenues which include websites, social software and email for infection. Web browsers are used to access these avenues on the Internet. The inspection of the web sites visited is crucial in the identification of threats and determining the effectiveness of the security awareness program. Polychronakis proposed the design of a URL collection system used in exploring the life cycle of web based malware (Polychronakis, Mavrommatis & Provos 2008). The system analyzed the web pages for malicious content; this was achieved by visiting the URL and monitoring the system for new processes, file system changes and registry modifications. Provos also proposed

a similar approach which consisted of identification of URL's, in-depth verification of maliciousness and aggregation of malicious URL's into site level ratings (Provos et al. 2007). These approaches are risky; a controlled approach is required by collecting the content from the URL and testing the content for maliciousness. Collection of the content from the web sites could be achieved with a web crawler. Mohhr discussed Heritrix which is an open source extensible, web scale, archival-quality web crawler (Mohr et al. 2004). Ikinci demonstrated the effectiveness of Heritrix as part of the MonkeySpider system used in the detection of malicious websites (Ikinci, Holz & Freiling 2008). The SPSA system follows a similar approach as demonstrated in the MonkeySpider system which includes the use of antivirus software in the identification of malicious content. These components discussed provide an automated and virtualized platform for the SPSA system.

The following section discusses the technical implementation of the components.

3. Shared public security awareness (SPSA) system architecture

The SPSA system consists of subsystems which as whole provide a virtualized automated platform to access the Internet, collect Internet behavioural data and delivery of a security awareness program at these establishments. These subsystems can operate independently of each other and thus are discussed separately. The automated virtualized environment is discussed in Section 3.1 and 3.2, followed by Section 3.3. and 3,4 which addresses the collection of data generated during browsing sessions and concluding with the elaboration of the security awareness program delivery mechanism in Section 3.5 and 3.6.

3.1 Internet access system

The Internet Access System is a modified user management system which based on configuration will direct users first to complete the security awareness questionnaire or direct users to the security awareness content before allowing access to the Internet (See **Figure 2**). The selection policy determines which functionality the user will interact with. The questionnaire functionality is used to assess the security knowledge of the user while the content functionality provides the user with an opportunity to learn about security related topics.

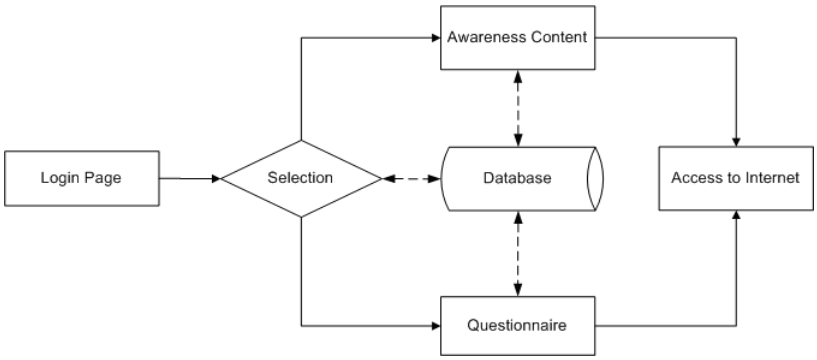

Figure 2: Internet access system

3.2 Virtual machine manager

The Virtual Machine (VM) Manager automates the operations of the SPSA system (See **Figure 3**). At the start of each day the VM manager loads a "clean" virtual machine for usage. A "clean" virtual machine represents a baseline installation of the operating system which has not been used by the users of these establishments. All components required to access the Internet are installed and configured. During the setup phase all software is tested for viruses and only reputable websites are visited to download software or update software. The task scheduler will initiate predefined scripts which will active the URL collection system to capture HTTP packet information into a file. Users will arrive at the workstations and start browsing websites. At the end of the day the task scheduler will initiate a script which will extract the data out of the file created and store the data in a database. The VM manager will shutdown the virtual machine which was used during the day, creates a backup of the virtual machine and assigns a date label to the virtual machine should forensics or malware analysis be required on the virtual machine.

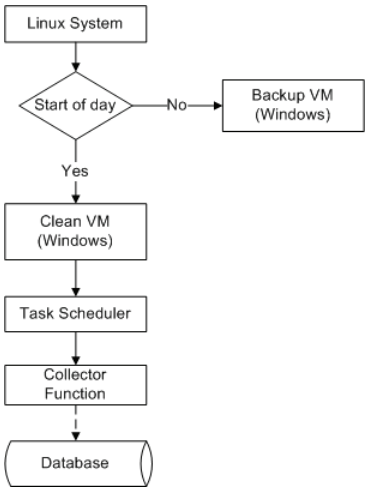

Figure 3: Daily virtual machine operations

3.3 URL collection system

The URL collection system is used in the collection of the web page address visited by the user and these include the web pages that are visited without the prior knowledge of the user. The URL collection is initialized during the start sequence of the user's virtual machine. TShark is a network protocol analyzer which provides the capability to capture packet data from a live network. Studies conducted by (Nascimento, Correia 2011) and (En-Najjary, Urvoy-Keller 2010) used TShark for the collection of specified network traffic. During the operation of the SPSA system a filter will be used to specify the required data to capture. Only outgoing HTTP traffic data is required which saves disk usage. The request line in the HTTP data packet contains the required data. The URL information is important to the work described here. According to (Forouzan 2003), "The URL is a standard for specifying any kind of information on the Internet. The URL defines four things: method, host computer, port, and the path." He states that host and path provide information on where the information is located. The URL provides a route to the content that was accessed by the user. TShark filter is configured only to collect the request line information encapsulated in the Hypertext Transfer Protocol (HTTP) header. An output file containing the captured data will be created when the time expires. This will contain the address of the webpage the user visited. Storage of the data is required and this is achieved by the URL transporter system which will analyze and extract the data from output file created by TShark. The URL transporter system is an application which will be executed at predetermined times during the day to poll a specified directory and extract the data from all the files within the directory and transport it to the external storage components for example a database server.

3.4 URL inspector

The URL Inspector component is designed to examine the URL's visited by the user. It consists of two components namely the URL Analyzer and the Malware Collection and Classification (MCC) system (See **Figure 4**). The URL Analyzer will examine each collected URL in the database against the Google Safe Browsing database, a service provided by Google, which enables applications to examine the location of the website against known phishing and malware websites (Google Code Lab 2008). This information is captured in a report. The MCC system also uses the URL captured in the database. The system consists of an Internet crawler called Heritrix which will be used to download the content of the URL and then use an anti-virus (AV) application called ClamAV to determine if the content is malicious. The list of malware found will be captured in a report. The report could assist in the identification of threats specific to these establishments and be used as a measure to determine the effectiveness of security awareness programs.

The data gathered about the browsing behaviour which include the destination address and the content of the web pages visited will be useful to determine the effectiveness of security awareness campaign by investigating the behavioural changes of the Internet users at these establishments

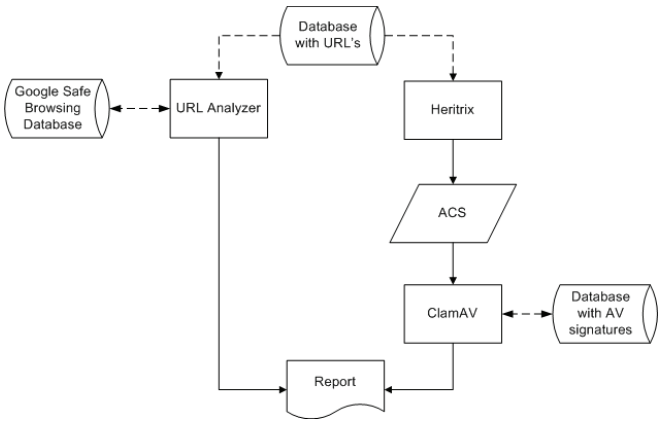

Figure 4: URL inspector

3.5 Awareness collection system

The security awareness levels of the users will be determined by completing a questionnaire. The users visiting these establishments are required to login. Thereafter the users will be presented with a set of questions which assesses the knowledge in security awareness related topics. Wilson reported on the best practises in the development of a security awareness program (Wilson, Hash 2003). One of the sections in the report discussed a comprehensive list of awareness topics some of these include but is not limited to:

- Password usage and management
- Spam
- Social Engineering
- Web usage
- Shoulder surfing
- Desktop security
- Unknown e-mail/attachments
- Incident response – contact whom? "What do I do?"

The Awareness Collection System was developed with requirements identified for the design of a security awareness game (See **Figure 5**). Game play encourages learning and with the use of game play components users are enticed to return to continue with the game. Using these principles would extend the contact time between the SPSA system and the user. Labuschagne recommended the use of Appointment, Influence and Status, and Progression dynamics (Labuschagne et al. 2011a). These dynamics are demonstrated visually with the use of badges. A badge is a visual indicator of an achievement. The appointment dynamic is represented with an image and is calculated with the consecutive logins over a period of three days. The user has to ensure that they continuing using the system after the badge have been obtained. The badge would be revoked should the user miss one day from using the system. The badge will be assigned again to the user after three consecutive day usage of the system. The status badge is provided when a user answers five questions correctly. The badge will be revoked in the event of an incorrect answer. Therefore the user is encouraged to provide the correct answers. The progression dynamic is represented with the progress bar which provides the user with a visual indicator on progress. The user is presented with randomized multiple choice questions. Labuschagne also identified security awareness topics which are applicable to establishments which allow resources to be shared amongst users accessing the Internet (Labuschagne et al. 2011b). These topics are more specific to the environment and include social media security awareness topics which is lacking in the work conducted by Wilson (Wilson, Hash 2003). The questions categories include but are not limited to the following:

- Spam
- Cyber bullying
- Malware
- Social Engineering

- Social Networking Sites
- Phishing

Figure 5: Screenshot of security awareness questionnaire

A report will be generated upon the completion of the questionnaire. The report indicates areas of weakness for the user and provides the user access to resources which addresses the areas of concern. A comprehensive report could assist in the identification of security awareness topics specific to the establishment. These results could also be incorporated in to the E-Awareness Model (E-AM) proposed by Kritzinger and Von Solms. This model would not allow home users to access the Internet if their security awareness levels are not satisfactory. Also the users are required to complete remedial work to address the shortcomings before access to the Internet is granted (Kritzinger, von Solms 2010). The SPSA system is designed to determine the security awareness levels and provide users to opportunity to improve their security knowledge with topics specific to users at these establishments.

3.6 Awareness content system

The Awareness Content System makes use of a content management system (CMS) to deliver the material to the user. The CMS used for the study purpose is called Moodle. It is a software package for producing Internet-based courses and web sites (Dougiamas 1999). Some typical features of Moodle are assignment submission, discussion forum, files download, grading, instant messages, online calendar, online news and announcement, online quiz and a wiki. These features provide a platform that integrates into the requirements of the SPSA system in the delivery of security awareness content to the users and provide a mechanism for assessment. The CMS stores that material of the identified security awareness topics which the user can easily access. One of the topics addresses the dangers of short URL's which could be encountered on social media platforms (See **Figure 6**). The user is provided with background information on the threat and suggests actions to perform once the threat is encountered. The CMS also provides functionality to assess the user's knowledge on the topic that was accessed by the user. The material content is collected from different sources which include vendor specific security best practises provided to the community. For instance, McCarthy composed a guide to Facebook security which addresses safety topics relating to the social networking platform (McCarthy, Watson & Weldon-Siviy 2011). One of the topics in this guide provides readers the necessary steps required to protect their Facebook accounts. The material for the SPSA system is updated once new information has become available. The material on the SPSA needs to current to address the latest threats identified by security vendors. This is possible by following information security threat trends that affects the categories identified for the establishments.

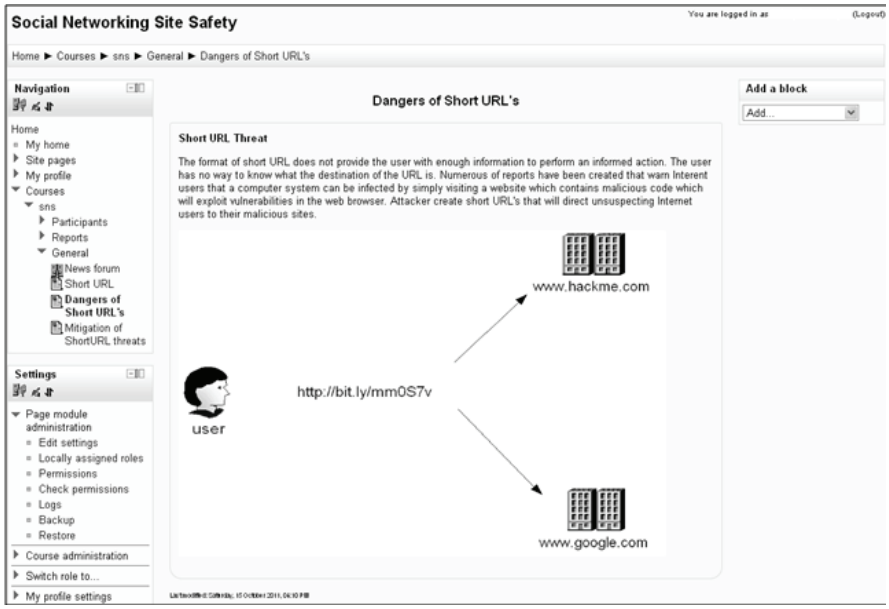

Figure 6: Awareness content system

4. Conclusion

This paper describes the design of an automated and virtualized platform used to promote security awareness in rural areas where the community access the Internet through shared resources. The SPSA system is a collection of components identified in the body of knowledge which provides a singular tool to measure the proficiency of the community and promotes security awareness. The SPSA system resolves the problem of associated with conducting security awareness programs in rural areas; these include but are not limited to travelling to the destination, establishing trust with the community and the frequency of exposing the users at these establishments to security related content. It provides an automated and virtualized infrastructure which improves the availability of resources to access the Internet, collects data about the browsing behaviour of the users, the identification and classification of threats encountered by the users, and conducts a security awareness program. The SPSA system does however have limitations. Currently the SPSA system consists of two subsystems: The automated virtualized platform which delivers the security awareness program and a separate platform which is designed for the evaluation of content visited by the users during the browsing session. The process to transfer the data collected by the automated virtualized platform is not automated. The majority of these establishments do not have the infrastructure to provide enough bandwidth to harvest all the content from the web pages as this process requires the research team to collect the data from the establishments and complete the process at another location which provides high bandwidth infrastructure. Furthermore the identification of malicious sites and software is limited to the signatures identified by security vendors. The SPSA system does not provide a component to automatically update the security awareness content.

Future research will include an additional component to determine if the virtual machine used by the user resembles malware infection behaviour. This would improve the accuracy of malware infection identification. In addition, the SPSA system requires a mechanism to assess the factors affecting the behavioural change of the users at these establishments. This is required to evaluate the effectiveness of the SPSA system. The evaluation of the effectiveness of the SPSA system would be determined with the deployment of the system in identified rural areas.

References

Abraham, S. & Chengalur-Smith, I. 2010, "An overview of social engineering malware: Trends, tactics, and implications", *Technology in Society,* vol. 32, no. 3, pp. 183-196.

Agarwal, S., Rahman, S. & Errington, A. 2009, "Measuring the determinants of relative economic performance of rural areas", *Journal of Rural Studies,* vol. 25, no. 3, pp. 309-321.

Bell, M. & Lintumaa, K. 2011, *Virtual Machines: Added planning to the forensic acquisition process.*, InSecure.

Dougiamas, M. 1999, *Modular Object-Oriented Dynamic Learning Environment*, 2.1.2 edn, Moodle Pty Ltd.

England, P. & Manferdelli, J. 2006, "Virtual machines for enterprise desktop security", *Information Security Technical Report,* vol. 11, no. 4, pp. 193-202.

En-Najjary, T. & Urvoy-Keller, G. 2010, "A first look at traffic classification in enterprise networks", *Proceedings of the 6th International Wireless Communications and Mobile Computing Conference*ACM, , pp. 764.

Forouzan, B.A. 2003, "Hypertext Tansfer Protocol" in *TCP/IP Protocol Suite*, 2nd edn, McGrawHill, , pp. 649-663.

Google Code Lab 2008, *Google Safe Browsing*. [online], http://code.google.com/apis/safebrowsing/

Grobler, M., Flowerday, S., Von Solms, R. & Venter, V. 2011, "Cyber Awareness Initiatives in South Africa: A National Perspective", *Southern African Cyber Security Awareness Workshop*Defence, Peace, Safety and Security (CSIR), South Africa, 12 May 2011, pp. 32.

Harlan, C. 2005, "Malware analysis for windows administrators", *Digital Investigation,* vol. 2, no. 1, pp. 19-22.

Ikinci, A., Holz, T. & Freiling, F. 2008, "Monkey-spider: Detecting malicious websites with low-interaction honeyclients", *Proceedings of Sicherheit, Schutz und Zuverlässigkeit,* .

Kim, W., Jeong, O., Kim, C. & So, J. 2011, "The dark side of the Internet: Attacks, costs and responses", *Information Systems,* vol. 36, no. 3, pp. 675-705.

Kritzinger, E. & von Solms, S.H. 2010, "Cyber security for home users: A new way of protection through awareness enforcement", *Computers & Security,* vol. 29, no. 8, pp. 840-847.

Kruger, H.A. & Kearney, W.D. 2006, "A prototype for assessing information security awareness", *Computers & Security,* vol. 25, no. 4, pp. 289-296.

Labuschagne, W.A., Burke, I., Veerasmay, N. & Eloff, M.M. 2011a, "Design of cyber security awareness game utilizing a social media framework.", *Information Security South Africa*South Africa, 15 May 2011.

Labuschagne, W.A., Eloff, M.M., Veerasmay, N., Leenen, L. & Mujinga, M. 2011b, "Design of a Cyber Security Awareness Campaign for Internet Cafe Users in Rural Areas", *Southern African Cyber Security Awareness Workshop*Defence, Peace, Safety and Security (CSIR), South Africa, 12 May 2011, pp. 42.

Ligh, M.H., Adair, S., Hartstein, B. & Richard, M. 2010, *Malware Analyst's Cookbook and DVD: Tools and Techniques for Fighting Malicious Code,* 1st edn, Wiley.

McCarthy, L., Watson, K. & Weldon-Siviy, D. 2011, *A Guide to Facebook Security For Young Adults, Parents, and Educators*, Facebook.

Mohr, G., Kimpton, M., Stack, M. & Ranitovic, I. 2004, "Introduction to Heritrix an archival quality web crawler", *Proceedings of the 4th International Web Archiving Workshop (IWAW'04)*, sep.

Nascimento, G. & Correia, M. 2011, "Anomaly-based Intrusion Detection in Software as a Service", .

Polychronakis, M., Mavrommatis, P. & Provos, N. 2008, "Ghost turns zombie: exploring the life cycle of web-based malware", *Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*USENIX Association, Berkeley, CA, USA, pp. 11:1.

Provos, N., McNamee, D., Mavrommatis, P., Wang, K. & Modadugu, N. 2007, "The ghost in the browser analysis of web-based malware", *Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*USENIX Association, Berkeley, CA, USA, pp. 4.

Steve, G. 2007, "Time to face virtualized realities", *Infosecurity,* vol. 4, no. 4, pp. 35-38.

The EasyHotspot team 2007, *EasyHotspot*., [online], http://easyhotspot.inov.asia/

Wilson, M. & Hash, J. 2003, *Building an Information Technology Security Awareness and Training Program*, National Institute of Standards and Technology, Gaithersburg.

Zhu, D. & Chin, E. 2007, "Detection of VM-Aware Malware", University of Berkeley, [online: http://radlab.cs.berkeley.edu/w/upload/3/3d/Detecting_VM_Aware_Malware.pdf