An Automated Analyzer for Users' Anti-Phishing Behaviour within a LAN

Abdullah M. Alnajim

Abstract*— Phishing is a security attack that seeks to trick people into revealing sensitive information about themselves and their Internet accounts. This paper proposes a novel anti-phishing approach that is deployed within a Local Area Network (LAN). The approach is a model that automatically perform ongoing analysis for users behaviours against phishing attacks and then based on the results it decides whether to train them or not against phishing. The aim is to enhance the phishing countermeasures applied on a LAN by making users aware of phishing attacks. A prototype proof of concept implementation is presented in this paper in order to test the approach's applicability. The prototype of the new model shows that the approach model runs and performs the concept.

Index Terms*— Modeling, Analyzer, Blacklists, LAN, e-Commerce Security, Network, Proxy, Online Banking Security, Phishing, Pharming.

I. INTRODUCTION

Recently, the Internet has become one of the important ways of communication. Security critical applications (e.g. online banking login page) that are accessed using the Internet are at the risk of Internet fraud. In the field of computer security, phishing is a criminally fraudulent process of capturing confidential information such as usernames, passwords and credit card details by impersonating a trustworthy entity in an electronic communication [1,2]. Violations of the security of confidential information would result in severe consequences, such as financial loss for e-commerce and online banking organizations and for individuals [3].

Phishing is aimed at people to take advantage of the way humans interact with computers or interpret messages rather than taking advantage of technical system vulnerabilities [4].

Phishing can be performed in different ways. They are as follows [3]:

1. email-to-email: this occurs when someone receives an email asking for sensitive information to be replied to the sender email or sent to another email.
2. email-to-website: this occurs when someone receives an email with embedded web address that leads to a phishing website.
3. website-to-website: this occurs when a phishing website is reached by clicking on an online advert or through a search engine.
4. browser-to-website: this occurs when someone misspelled a web address of a legitimate website on a browser and then goes to a phishing website that has a similar address.

The Anti-Phishing Working Group (APWG) has reported that during the last three months of the 2014 (Quarter 4 (Q4)) only, the number of unique phishing reports submitted to APWG was 197,252 [5]. The report shows that this was an increase of 18 percent from the 163,333 received in Q3 of 2014. APWG also stated that the total number of phishing attacks observed in Q4 was 46,824 which targeted a total of 437 brands. APWG assured that the United States continued to be the top country hosting phishing sites [5].

There are technical advances that mitigate the problem of phishing. For instance, security toolbars, such as SpoofStick, TrustBar and SpoofGuard, can prevent phishing attacks.

Anti-phishing training for end-users is indispensable to any proposed technical solution. It is suggested that while technical improvements may continue to stop the attacks, end-user training is a key component in phishing attacks mitigation [6]. In preventing online fraud, Symantec [7] believes that users' awareness is central to helping to change their behaviours and thus reduce their mistakes with phishing emails and websites.

Anti-phishing training will make the end-user aware and it will erect an effective barrier against phishing attempts. Anti-phishing awareness was shown to have a great positive effect in mitigating the risk of phishing [8].

This paper proposes a novel anti-phishing approach that is deployed within a Local Area Network (LAN). The approach is a model that automatically and continuously analyzes users behaviours against phishing attacks and then based on the results it decides whether to train them or not against phishing. The aim is to enhance the phishing countermeasures applied on a LAN by making users aware of phishing attacks and how to prevent them.

In this research, there is an assumption that phishing attacks do not use either software to change the host files in users' operating systems or any malicious software, such as a virus, worm or Trojan horse, that runs in users' operating systems. These are called 'Pharming' and 'Malware' and are different from phishing.

The remainder of the paper is organized as follows. Section two reviews the literature regarding phishing detection methods. The third section presents the proposed model that is applied on a LAN to detect phishing attacks. The fourth section discusses and analyzes the pros and cons of the proposed model. The final section concludes the paper and discusses the possible way of future work.

II. RELATED WORK

There are technical (e.g. toolbars) and training (e.g. tips) approaches to mitigate phishing. The anti-phishing toolbars are web browser plug-ins that warn users when they reach a suspected phishing site [4]. Anti-phishing tools use two major methods for mitigating phishing sites. The first method is to use a blacklist that lists phishing URLs. Blacklists have a high level of accuracy because they are constructed by paid experts who verify a reported URL and add it to the blacklists if it is

Revised Version Manuscript Received on June 30, 2015.

Dr. Abdullah M. Alnajim, Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

considered as a phishing website. The second method is to use heuristics to check the host name and the URL for common spoofing techniques. The heuristics approach is not 100 percent accurate since it produces low false negatives (FN), i.e. a phishing site is mistakenly judged as legitimate, which implies they do not catch all phishing sites. The heuristics often produce high false positives (FP), i.e. incorrectly identifying a legitimate site as fraudulent [9].

To increase the accuracy FP and FN rates, Xiang et. al. [10] proposed CANTINA+ which is a comprehensive feature-based approach including eight novel features, which exploits the HTML Document Object Model (DOM), search engines and third party services with machine learning techniques to detect phishing. Xiang et. al. [10] designed two filters to help reduce FP. The first is phishing detector that uses hashing to catch highly similar phishing attacks. The second is a login form filter, which directly classifies Web pages with no identified login form as legitimate. CANTINA+ eventually is evaluated and achieved good accuracy rates but yet did not reach a 100 percent accurate FP and TP rates.

The anti-phishing tools always works in a way that receives users' submission of phishing URLs. Usually, they are not fast and efficient enough to find and take down phishing attacks [11]. Bo et. al. [11] propose a hybrid method to discover phishing attacks in an active way based on DNS query logs and known phishing URLs. They analyzed phishing reports from Anti-phishing Alliance of China (APAC) and developed their system to report living phishing URLs automatically to APAC every day. They evaluated the system and reported that it is good complement to traditional anti-phishing tools.

In addition to the anti-phishing tools, there are different anti-phishing training approaches to make users aware of phishing emails and websites and to learn how to avoid them. The most basic approach is publishing guidelines for the Internet users to follow when they go online. These guidelines are referred as tips for users [12]. All the information used in the training approaches is based on tips for users.

Many financial and commercial, private and government institutions (e.g. eBay and HSBC) have provided anti-phishing training tips for detecting phishing emails and websites. The aim of the tips is to train users to look for phishing clues located in emails and websites to enable them to make better decisions in distinguishing phishing emails and websites. People in general do not read anti-phishing online training materials although some of them are found effective when used [12].

Many commercial institutions, such as Microsoft, periodically send email security information to help their customers in protecting their online security [13]. This email provides practical security tips, useful resources and links, and a forum to ask security-related questions.

Microsoft states that the email is suitable for customers to stay up to date on the latest issues and events with:

- Security tips including anti-phishing tips.
- Security critical updates.
- Answers to frequently asked questions (FAQs) on security topics.
- Information about security trials and downloads.
- Tips from security team for home users.

Theses emails are usually sent in text and HTML formats. The limitation of this approach is that customers who are interested in receiving these emails need to subscribe with the commercial institutions (i.e. anti-phishing emails providers) in order to be included in receiving these emails.

An online game was proposed in order to teach users good habits to help them avoid phishing attacks [14]. Kumaraguru et al [15] considered training people about phishing email during their normal use of email. Their aim was to teach people what phishing clues to look for located in emails. They found that email training approach works better than the current practice of publishing or sending anti-phishing tips.

Alnajim and Munro [16] proposed a novel anti-phishing approach that uses training intervention (APTIPWD). The approach helps users to make correct decisions in distinguishing phishing and legitimate websites. It brings information to end-users and helps them immediately after they have made a mistake in order to detect phishing websites by themselves. The new approach also keeps anti-phishing training ongoing process. This means, in all time, once users tries to submit information to phishing website, they will be trained (see Figure 1).

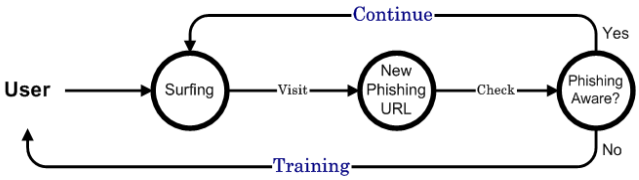

Fig 1. The broad idea of APTIPWD

There are many anti-phishing tips that can be used in the intervention message. The effectiveness of most common users' tips for detecting phishing websites using novel effectiveness criteria was examined [12]. The aim of the tips' effectiveness examination was to find fewer anti-phishing tips that users can focus on to detect phishing attacks by themselves. Therefore, the most effective anti-phishing tip was used [16]. The tip was as follows: "*a fake website's address is different from what you are used to, perhaps there are extra characters or words in it or it uses a completely different name or no name at all, just numbers. Check the True URL (Web Address). The true URL of the site can be seen in the page 'Properties' or 'Page Info': While you are on the website and using the mouse Go Right Click then Go 'Properties' or 'Page Info'. If you don't know the real web address for the legitimate organization, you can find it by using a search engine such as Google*".

Countries around the world follow high level procedures (a country-based) in order to mitigate phishing attacks. Alnajim [17] introduced and analyzed a high level anti-phishing countermeasure implemented in Saudi Arabia. The Saudi Arabian countermeasure is obviously applied on the Internet traffic within Saudi Arabia. Therefore, it was very important to analyze the countermeasure model against all possible Phishing attacks scenarios initiated by or designed to attack users inside and outside Saudi Arabia. The location of the source or the destination of Phishing attacks is vital in the

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

analysis of the model because the model works only in Saudi Arabia. Hence, based on the location, the scenarios used in the analysis came across all possible sources of Phishing attacks as well as all possible destinations of Phishing attacks. Based on the analysis methodology mentioned, Alnajim [17] examined the model and found that the model is effective when phishing websites are reached by users who surf the Internet inside Saudi Arabia whereas it is ineffective in protecting users from falling in Phishing when the websites are reached by users who surf the Internet from outside Saudi Arabia.

Alnajim [18] then proposed a novel country based model to detect phishing attacks. The aim is to enhance the phishing countermeasures applied on a country's Internet infrastructure. This is because of that the anti-phishing framework in Saudi Arabia is exposed to users when they fall to phishing attacks and thus enhancing anti-phishing behaviors by training them to detect phishing instead of only blocking phishing websites is proposed. The idea presented by Alnajim and Munro [16] is applied on the current anti-phishing framework implemented in Saudi Arabia [17].

Alnajim [18] new model has advantages and limitations. The advantage is that the model is exposed to phishing victims who are inside the country deployed it (e.g. Saudi Arabia). This enhances the anti-phishing countermeasures deployed nowadays in Saudi Arabia. Whereas a potential drawback could be that it makes the Internet traffic slower. This is because of extra component (i.e. Intervention Server) added to the anti-phishing detection framework in Saudi Arabia.

III. THE NEW MODEL

A. Objective

Organizations, such as universities and companies, have many users to their internal LANs. They use their LANs to do their tasks, access the network resources, use the Internet or communicate with others. They may be exposed to phishing attacks since they are connected to the Internet. Therefore, making users aware of phishing attacks and how to prevent them would enhance the phishing countermeasures. Due to this, this research proposes a model that checks continuously the LAN users' phishing awareness status by automatically analyzing their behaviours against phishing attacks in order to know whether they are phishing unaware users. Based on the results a decision is taken to get them trained against phishing (in case they are unaware) by using the training intervention idea proposed previously [16].

B. Assumptions

In this research, in addition to the assumption mentioned in the introduction section, there are few technical assumptions that should be stated before presenting the new model. They are as follows:

1. The LAN is connected to the Internet.
2. The LAN resources are controlled. This means that every user should be registered and authorized to use the LAN by an authentication system. Once a user would like to use the LAN, they should authenticate themselves by providing their access credentials (e.g. id and password)
3. Every user has an email address that is linked to his network ID.

C. Design and Scenarios

A client-server model is a common design for distributed computing. The client and the server are two components that interact between each other [19].

A client-proxy-server model extends the client-server model [20]. It introduces an additional component which is a proxy. The proxy is located between the client and the server [20]. In a proxy based computer network, any URL request to the web made by a client is directed to the URL domain server. The server component is represented by the "Internet". Proxies have been widely used in many applications to perform various tasks such as clients' connections control, URLs' request control, caching and filtering data [21].

The interaction between client and server is as follows [19]:

- Client requests a service from Server.
- Server processes the requests and replies to Client.

However, in the client-proxy-server, the interaction becomes as follows:

- Client sends request for Server to Proxy.
- Proxy passes request to Server.
- Server processes the request and sends reply for the Client to Proxy.
- Proxy passes reply to Client.

In this research, the new model has additional components added in order to perform as expected. The new model main component is referred as "Automated Trainer". The Automated Trainer is a framework that includes few subcomponents to perform the task of the model. These subcomponents are as follows:

- An agent named as the User Behaviour Analyser' (UBA).
- A database named as the User Awareness Status (UAS).
- A Web Mail Server.
- A server named as the Local Fixed List of Anti-phishing Training Websites Sever (LFLAPTW).

The job of each sub-components mentioned and their interaction with each other (scenario) are described below.

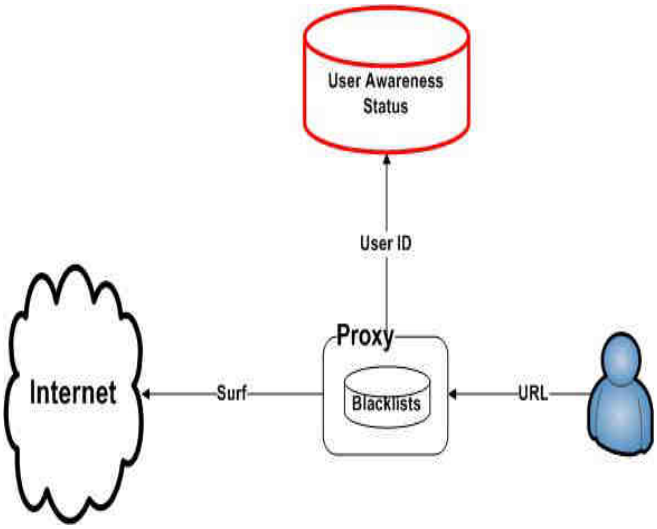

Fig 2. The Proposed User Awareness Status Feed

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

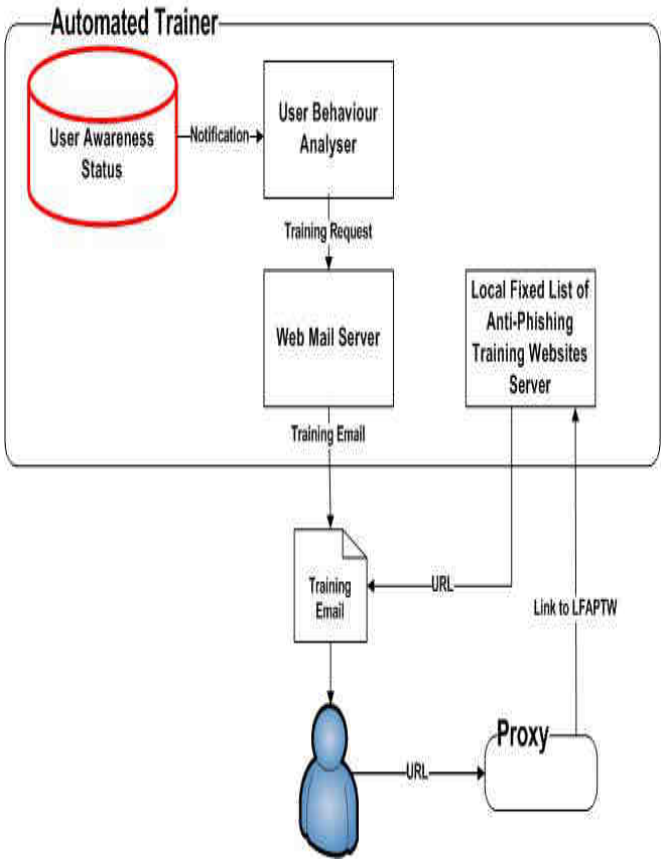

Fig 3. The Proposed Anti-phishing LAN Approach

When a user would like to surf the Internet working from a terminal within a LAN (Local Area Network), they request a URL. The URL is sent to the network proxy. The proxy checks weather the URL is blacklisted or not. Accordingly the proxy either accepts –based on a defined policy- the URL and pass it to the Internet zone or rejects it. In case of that a user requests a blacklisted phishing URL the proxy will reject it and then sends the user ID to an Automated Trainer system (described later). The system then changes the user status from 'phishing aware' to 'phishing unaware' in the User Awareness Status (UAS) database (Please see Figure 2). This database is created in order to record users anti-phishing awareness status. All users IDs recorded in the database are 'phishing aware' by default unless a notification comes from the network proxy. This is to ensure that the system is convenient.

Figure 3 shows the proposed anti-phishing LAN approach. The approach is a system called 'Automated Trainer'. This system has a primary component which is referred as 'User Behaviour Analyser' (UBA). The UBA is an agent that performs an ongoing process of analyzing user anti-phishing behaviours within a local network and decides whether users anti-phishing awareness needs to be enhanced or not (see Figure 4). The UBA takes the user status from the UAS database mentioned previously.

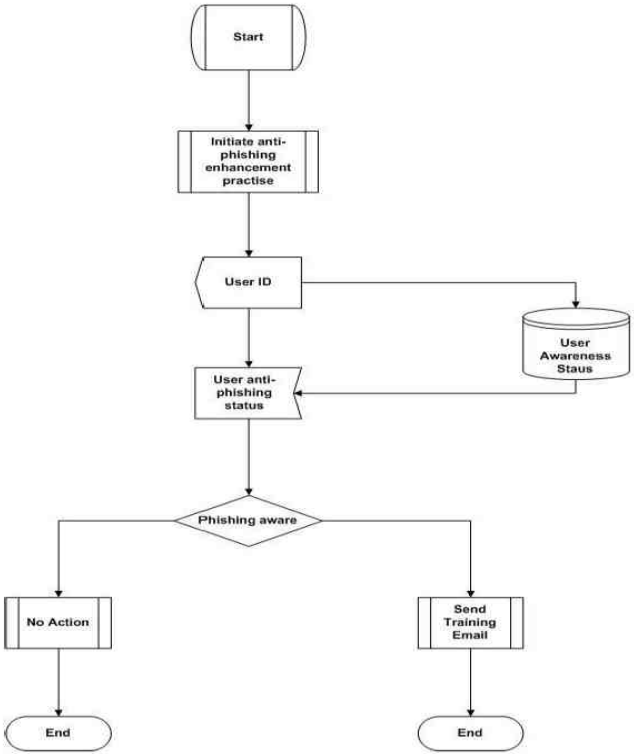

Fig 4. Flowchart of User Behaviour Analyser (UBA)

The UBA's task is to frequently checks the UAS database for users flagged 'phishing unaware'. If it finds phishing unaware users, it initiates a need-for-training request 'training request' and sends it to a Web Mail Server working within the LAN. The UBA sends a packet includes user ID, the textual email content and the fake targeted brand email address.

The web mail server is configured in a way that it receives the packet from the UBA and sends an anti-phishing Training Email to the user. The email has the sent textual email content and includes a fake URL pretended to be a URL for a genuine website. The fake URL leads to a fake website that is run locally among many websites located in a local server. This server is referred as Local Fixed List of Anti-phishing Training Websites Sever (LFLAPTW). Running these websites locally ensures users confidentiality for their data.

IV. Simulation and Discussion

A prototype proof of concept implementation is presented in this section. In order to test the approach's applicability, simulating the possible scenarios is needed.

Therefore, a demo of the proposed model is implemented. A database and functioning UBA were implemented. The websites used in the simulation were identical copies of the real ones. The legitimate and phishing websites were stored on the local machine and run by Apache server. All the phishing emails and websites used for this simulation were based on real ones collected from various phishing examples resources. The DNS (Domain Name System) host files in Windows operating system was modified so that web browsers displayed the URL of the actual phishing websites. All phishing websites were functional so their users were able to submit information. A piece of Jave code was written and Mysql database was used.

In a proxy based computer network, the proxy settings in the LAN settings of every single machine (client) that is connected to it should be configured so that the address of the

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

proxy is provided with its port. For example, clients in Qassim University network applied the university proxy in their LAN settings. Therefore, the clients used were connected to the Proxy to request any URL. This was carried out by putting the server machine as its LAN proxy on the default port 80.

Having implemented the model as a prototype proof of concept, the new model shown in figure 3 runs and perform the concept. This model would enhances the anti-phishing countermeasures deployed within a LAN by making unaware users aware of phishing attacks and how to prevent them.

In the other hand, the limitation could be that it makes the network traffic a little bit slower. This is expected because of the extra component (i.e. Automated Trainer) added to a LAN.

V. CONCLUSION

In this paper, a novel anti-phishing approach that is deployed within a Local Area Network (LAN) is proposed. The approach is a model that automatically performs ongoing analysis for users behaviours against phishing attacks and then based on the results it decides whether to train them or not against phishing. The aim is to enhance the phishing countermeasures applied on a LAN by making users aware of phishing attacks and how to prevent them.

The new approach model main component is referred as "Automated Trainer". The Automated Trainer is a framework that includes few subcomponents to perform the task of the model. These subcomponents are an agent named as the User Behaviour Analyser' (UBA), a database named as the User Awareness Status (UAS), a Web Mail Server and a server named as the Local Fixed List of Anti-phishing Training Websites Sever (LFLAPTW).

A prototype proof of concept implementation is presented in this paper in order to test the approach's applicability. The prototype of the new model shown in figure 3 runs and performs the concept.

This approach would enhance the anti-phishing countermeasures deployed within a LAN by making unaware users aware of phishing attacks and how to prevent them. In the other hand, the limitation could be that it makes the network traffic slower. This is expected because of the extra component (i.e. Automated Trainer framework) added to a LAN.

REFERENCES

1. The National Consumers League Projects (2015). Phishing. Available: http://www.fraud.org/scams/internet-fraud/phishing, last access on 15/5/2015.
2. G. K. Tak, N. Badge, P. Manwatkar, A. Ranganathan, S. Tapaswi, "Asynchronous Anti Phishing Image Captcha approach towards Phishing". Proc. the 2nd International Future Computer and Communication (ICFCC), Wuhan, IEEE Press, pp. V3-694 - V3-698.
3. A. Alnajim, and M. Munro "An Approach to the Implementation of the Anti-Phishing Tool for Phishing Websites Detection". Proc. International Conference on Intelligent Networking and Collaborative Systems (INCoS 2009). Barcelona, Spain, IEEE Press, 2009, pp. 105 - 112.
4. J. S. Downs, M. B. Holbrook and L. F. Cranor, "Decision strategies and susceptibility to phishing". Proc. the 2nd symposium on usable privacy and security. New York, USA, ACM Press, 2006, pp. 79 – 90.
5. Anti-Phishing Working Group APWG. (2015). Phishing Activity Trends Report, 4th Quarter 2014. Available: http://docs.apwg.org/reports/apwg_trends_report_q4_2014.pdf, last access on 26 June 2015.
6. S. A. Robila and J. W. Ragucci, "Don't be a Phish: Steps in User Education". Proc. 11th annual SIGCSE conference on innovation and technology in computer science education. New York, ACM Press, 2006, pp. 237 – 241.
7. Symantec. (2004). Mitigating Online Fraud: Customer Confidence, Brand Protection, and Loss Minimization. Available: http://www.antiphishing.org/sponsors_technical_papers/symantec_online_fraud.pdf, last access on 21/3/2007.
8. A. Alnajim and M. Munro, "Effects of Technical Abilities and Phishing Knowledge on Phishing Websites Detection". Proc. the IASTED International Conference on Software Engineering (SE 2009), Innsbruck, Austria, ACTA Press, 2009, pp. 120-125.
9. Y. Zhang, J. I. Hong and L. F. Cranor, "Cantina: a content-based approach to detecting phishing web sites". Proc. 16th international conference on WWW. New York, ACM Press, 2007, pp. 639 – 648.
10. G. Xiang, J. Hong, C. P. Rose, L. Cranor, "CANTINA+: A Feature-Rich Machine Learning Framework for Detecting Phishing Web Sites". ACM Transactions on Information and System Security (TISSEC), 2011 , Volume 14 Issue 2, New York, ACM Press, Article No. 21 .
11. H. Bo, W. Wei, W. Liming, G. Guanggang, X. Yali, L. Xiaodong, M. Wei, "A Hybrid System to Find&Fight Phishing Attacks Actively". IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology. Lyon, IEEE Computer Society, 2011, pp. 506-509
12. A. Alnajim and M. Munro, "An Evaluation of Users' Tips Effectiveness for Phishing Websites Detection". Proc. 3rd IEEE International Conference on Digital Information Management ICDIM, London, IEEE Press, 2008, pp. 63-68.
13. Microsoft Corporation. (2007). *Microsoft Security for Home Computer Users Newsletter.* Available: http://www.microsoft.com/protect/secnews/default.mspx, last access on 16 March 2007.
14. S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. F. Cranor, J. Hong and E. Nunge, "Anti-Phishing Phil: the design and evaluation of a game that teaches people not to fall for phish". Proc. 3rd symposium on usable privacy and security SOUPS. New York, ACM Press, 2007, pp. 88 – 99.
15. P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong and E. Nunge,"Protecting people from phishing: the design and evaluation of an embedded training email system". Proc. the SIGCHI conference on Human factors in computing systems. New York, USA, ACM Press, 2007, 905 – 914.
16. A. Alnajim and M. Munro, "An Anti-Phishing Approach that Uses Training Intervention for Phishing Websites Detection". Proc. 6th IEEE International Conference on Information Technology - New Generations (ITNG). Las Vegas, IEEE Computer Society, 2009, pp. 405-410.
17. A. Alnajim, "High Level Anti-Phishing Countermeasure: A Case Study". Proc. The The World Congress on Internet Security (WorldCIS-2011), London, UK, IEEE Press, 2011, pp. 139 – 144.
18. A. Alnajim, 2015. "A Country Based Model Towards Phishing Detection Enhancement". International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2015, Volume 5 Issue 1, pp. 52 - 57.
19. W. Jia, and W. Zhou, "Distributed Network Systems: From Concepts to Implementations". New York: Springer, 2004.
20. M. P. Singh, "the Practical Handbook of Internet Computing". USA: Chapman & Hall/CRC Publisher, 2005.
21. Y. Xiao and H. Chen, "Mobile Telemedicine: A Computing and Networking Perspective". USA: Auerbach Publications, 2008.

Dr. Abdullah M. Alnajim is an information security and academic consultant. He is also a faculty in the Information Technology Department, college of Computer at Qassim University, Saudi Arabia. Dr. Abdullah Alnajim had BSc in Computer Science from King Saud University in Saudi Arabia in 2002. Dr. Alnajim had MSc in Internet and Distributed Systems from Durham University in the United Kingdom in 2005. Dr. Alnajim had a Ph.D from the Department of Computer Science at Durham University in 2009. His Ph.D thesis was entitled as 'Fighting Internet Fraud: Anti-Phishing Effectiveness for Phishing Websites Detection'. Dr. Alnajim's research interests involve Internet security and frauds that encounter web applications especially online banking and e-commerce applications.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.