# Towards Antiphishing with Intrusion Detection (September 2009)

V. Vishe, HS Venter and JHP Eloff

***Abstract*— Phishing is a technique where an attacker gains an identity of a web user by misleading a web user to fraudulent websites. Phishing techniques combine social engineering with technology. Financial loss due to phishing has been increasing every year. There are many tools available for antiphishing; however attackers find innovative ways to bypass existing tools. In this paper we propose a unique architecture containing a basic model which addresses most of the existing phishing problems. An Intrusion Detection System (IDS) is set in place to prevent any damage from phishing sites. This architecture learns about new attacks from the IDS and modifies itself accordingly.**

***Index Terms*— Phishing, WHOIS, APDS**

## I. INTRODUCTION AND BACKGROUND

Phishing is a technique wherein the attacker sends an email containing links to fraudulent websites to victims These websites are hosted by attackers either on legitimate workstations or on a compromised network of PCs known as a Botnet. [7] The attacker who sends the Phishing mail is called a phisher. The fake website usually looks very similar to a legitimate one to appear convincing to the victim. Once the victim enters his details on the fake website, they are sent to the phisher who uses them for financial gain.

Figure 1: 827% Rise in Phishing attacks from Jan 2008 [2]

Manuscript submitted on June 7, 2009.

V. Vishe is with the Department of Computer Science at the University of Pretoria, Pretoria, South Arica, phone: +27-842301150; fax: 012 311 7839; (e-mail: vishev@telkom.co.za).

Prof. Hein S. Venter is an Associate Professor with the Department of Computer Science at the University of Pretoria, Pretoria, South Africa (e-mail: hventer@cs.up.ac.za).

Prof. Jan Eloff is an Extraordinary Professor with the Department of Computer Science at the University of Pretoria, Pretoria, South Africa. Also he is currently appointed as the Research Director at SAP Meraka UTD (e-mail: jan.eloff@sap.com).

As shown in Figure 1 [2], Phishing attacks are exponentially growing every year causing increasing financial loss. The tools available currently are not very effective when it comes to preventing new attacks. Also, many of them have a high false positive ratio in identifying legitimate websites as phishing sites. [4]

Most of the anti-phishing tools available currently are based on an automation process which applies certain rules in certain order to determine that the site is a probable phishing site. As the process is automated it becomes easy for the phisher to find loopholes in the rules and thus ways to bypass them.

This paper tries to address this particular issue. [3] The present anti-phishing tools fail to prevent damage from new attacks because they use static rules that fall short of identifying phishing sites or take time to add new rules to avert new attacks.

The aim of this research is to create an intelligent and scalable architecture through which existing and new Phishing attacks can be identified and avoided.

The rest of the paper is organized as follows. Section II explains the proposed architecture. Section III summarizes the challenges and future work.

## II. ARCHITECTURE

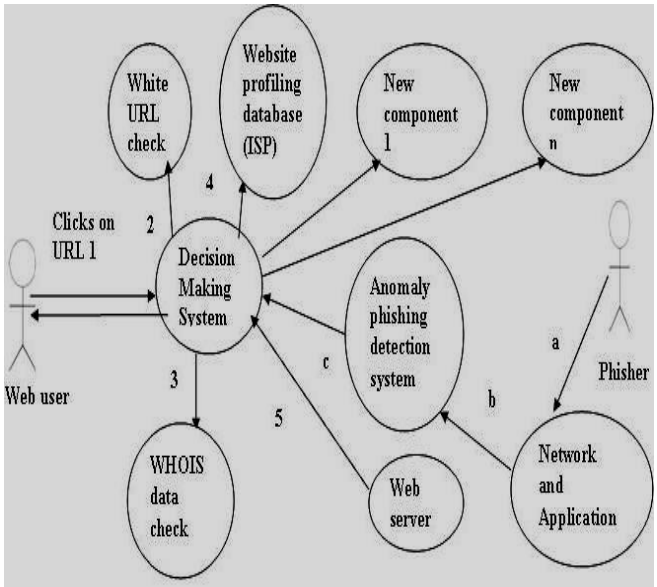


Figure 2: Architecture

As shown in Figure 2, the architecture consists of five basic components in order to detect most of the existing phishing attacks. They are as follows,

1. White URL Check [6]: checks a URL against a list of URLs trusted by web user.
2. WHOIS [5] data check: a technique that queries its

database in order to find registration information for any URL. WHOIS Database stores details such as phone number and email of the person who registered the site. The age of a URL can also be found using a WHOIS database query.
3. Web site profiling database: stores profiles of websites. Whenever a new URL reaches the Internet Service Provider (ISP), this database is updated with the new profile created for that URL.
4. Web Server [8]: checks the response from web server for any script injection. E.g. checks hidden malicious javascript in response.
5. Decision making System (DMS) is the core engine of the architecture which coordinates between different components while calculating the probability of a website to determine if it is a phishing site. Each component analyses its output for a URL, calculates a certain weight based on risk that it is a phishing attempt and sends it back to the DMS. After gathering the weights from various components, the DMS calculates the probability of a URL being a phishing attack or not.
As shown in Figure.2 when a web user enters a URL, the DMS takes control. It first performs white URL check (at step 2) to determine if the URL is from a trusted list, then a WHOIS db check (at step 3) to know the age and registration details, next it creates a profile of the URL clicked by the user and compares it with the web profiling database (at step 4), One of the parameters for creating profile is, the time required to get the response to ISP from Web server. The time may differ in case of phishing URL so will the profiles. Finally it checks the response returned from the web server for any scripting injection.(at step 5) As mentioned above the DMS then calculates the probability of the URL belonging to the Phishing category based on weights returned by the four components. This result is sent back to the web user to indicate a possible phishing attack. (at step 6)

The tricky part comes in when a phisher tries innovative phishing attacks. As the basic model is static and based on certain rules, it will fail to detect any new kind of phishing attack. To counter this problem an additional component called Anomaly Phishing Detection System (APDS) is added. APDS is based on similar concept as intrusion detection. APDS works on a behavioral profile of a phisher which works on application layer and checks for any anomaly. As soon as any new attack comes in, it is identified as an anomaly and will be captured by the APDS. APDS alerts DMS at runtime which then verifies the possibility of a phishing attack. If it is a phishing attack the DMS is dynamic and intelligent enough to create a new component that will encounter the new phishing attack. This new component will then be transferred to all web users to protect them from the new attack. Thus the more a phisher tries to attack a system the stronger the system becomes.

This Architecture is flexible as the web user can control the number of components that should be in use. By specifying fewer components the privacy of the user will be better protected as the data will not be spread across multiple components and the performance will be higher though by trading off security. This architecture is scalable as the new components are added dynamically.

## III CHALLENGES AND FUTURE OF THE ARCHITECTURE:

The major challenge faced by this architecture is that, it becomes heavier with time as more components get added to the architecture. This will downgrade the performance and speed of the system. However this particular issue can be addressed by caching and combining components with overlapping logic.

Future work on this architecture will involve developing a prototype and measuring its effectiveness. Also this architecture can be extended to avoid not only phishing attacks but different kinds of network attacks.

## REFERENCES


[1] PhishTank http://www.phishtank.com

[2] Anti-Phishing Working Group (APWG) Phishing activity trends report , 2nd Half 2008 www.antiphishing.org/reports/apwg_report_H2_2008.pdf

[3] Y. Zhang, J.Hong, and L. Cranor. Cantina: A content based approach to detecting phishing web sites. In WWW, 2007

[4] Y. Zhang, S. Egelman, L. Cranor, and J. Hong Phinding Phish: Evaluating Anti-Phishing Tools. In Proceedings of the 14th Annual Network & Distributed System Security Symposium (NDSS 2007), San Diego, CA, 28th February - 2nd March, 2007.

[5] "WHOIS Protocol Specification," Aug. 2005; http://www.rfc-editor.org/rfc/rfc3912.txt

[6] C. M. McRae and R.B.Vaughn, "Phighting the phisher: Using Web Bugs and Honeytokens to Investigate the Source of Phishing Attacks (Published Conference Proceedings style)," in *Proc. 40th Hawaii International Conference on System Sciences(HICSS 2007) IEEE*

[7] Y. Xie, F. Yu, K. Achan, R. Panigrahy, G. Hulten, I. Osipkov, "Spamming Botnets: Signatures and Characteristics" Microsoft Research, Silicon Valley - 2008

[8] D. Harley, A. Lee, "A Pretty Kettle of Phish" http://www.eset.com/download/whitepapers/Pretty_Kettle_of_Phish.pdf


## BIOGRAPHY

Vivek Vishe obtained his Bachelor of Engineering degree in Information Technology from University of Mumbai in August 2004. He is a part time Computer Science student at The University of Pretoria pursuing research in antiphishing. He has been working in the telecommunication industry for last 5 years; currently he is employed at Telkom South Africa in Integration. (EAI)