A rule based approach for detecting Phishing attacks

Vivek Shukla¹, Mahendra Kumar Rai²

¹,² M.E. (CSE), HOD (IT)
(S.R.I.T.) Jabalpur (M.P)

Abstract: Phishing attacks are one of the emerging serious threats against personal data security. These attacks are often performed by sending out emails that seem to originate from a trusted party. The objective is to deceive the recipient to release sensitive information such as usernames, passwords, banking details, or credentials. The aim of phishing is to steal a user’s identity in order to make fraudulent transactions as if the Phisher were the user. Though a large number of methods have been proposed and implemented for detecting Phishing attacks, a complete solution is missing. A great amount of research is being carried out to solve this problem using rule based method, browser based method and machine learning approaches but still there are insufficient methods that can be used against Phisher’s novel attacks which they are able change time to time. In this Dissertation entitled “A RULE BASED APPROACH FOR DETECTING PHISHING ATTACKS”, a solution is proposed for organization wide solution; rule set has been proposed for this system to filter out phishing mails at the perimeter of the organization. A balanced rule set has been used to keep false positive and false negative low.

Keywords- Phishing attack, phishing website, rule-based, machine learning, phishing detection, decision tree

I. INTRODUCTION

Phishing is the term used to describe massive e-mails that trick recipients into revealing their personal or company confidential information, such as social security and financial account numbers, account passwords and other identity or security information. These e-mails request the user’s personal information as a client of a legitimate entity with a link to a website that looks like the legitimate entity’s website or with a form contained in the body of the e-mail. The aim of phishing is to steal a user’s identity in order to make fraudulent transactions as if the phisher were the user. In a typical phishing attempt, you will receive an authentic-looking email message that appears to come from a legitimate business; e.g., bank, online shopping site. It will ask you to divulge or verify personal data such as an account number, password, credit card number or Social Security number. Often the language will be intimidating, e.g. "Your account will be closed or suspended if you don't follow these directions." Although legitimate online banking and e-commerce are very safe, one should always be careful
about giving personal financial information over the Internet. It is also possible for one to be phished by mail, telephone or even in person. Security organizations and companies have done research and development on anti-phishing techniques and tools, which include basic changes in the E-mail infrastructure to help lessen Spam, more widespread deployment of anti-Spam, anti-Malware, personal firewall products, privacy protection software, and stronger authentication for electronic transactions, etc. Some of them have good effects on decreasing the number of phishing.

Unfortunately, phishing attacks are growing both in numbers and in complexity. Phishers are always refining their techniques such as using automated tools and Botnets to increase their catch. Phishing Emails are becoming increasingly sophisticated.

II. MOTIVATION

The APWG website reports that number of crimeware-spreading sites infecting PCs with password-stealing crimeware reached an all time high of 31,173 in December, and 827% increase from January 2013[]. These attacks are ever increasing and do not show any sign of slowing. These attacks come in the form of emails asking you to reveal your personal data such as login credentials, credit card number and ATM card number etc.

Figure 1.1 Number of Crime ware Websites in 2013

Some other information APWG reports

1. Unique phishing reports submitted to APWG recorded a yearly high of 34,758 in December.
2. The number of unique keyloggers and crimeware-oriented malicious applications reached an all-time high in July reaching 1,519 in July.
3. Rogue anti-malware began to rise in July, skyrocketing in December to 9,287.

Though current solutions use strong spam filters to filter out any malware mail still Phishers are quite successful in their business. A lot of research is going on in pursuit of betterment of the technology available; browser based plug-in, phishing detection using IDS are some of the available solution. But they are not perfect and there is need of improvement.

III. PROBLEM STATEMENT

a rule set has been proposed to classify phishing emails. The chief objective of the dissertation is to propose a minimal rule set so that classifier can filter out Phishing mail with better accuracy.

We have made an attempt to design a system to keep it free from problem such as

1. Dependence on individual user’s settings
2. Dependence on web browser policies
3. Non evaluation of mails
4. Bypasing of Client-browser.
IV. BACKGROUND

One of the emerging serious threats against personal data security is phishing. Phishing attacks are often performed by sending out emails that seem to originate from a trusted party. The objective is to deceive the recipient to release sensitive information such as usernames, passwords, banking details, or credentials [15]. Phishing is the term used to describe massive e-mails that trick recipients into revealing their personal or company confidential information, such as social security and financial account numbers, account passwords and other identity or security information. These e-mails request the user’s personal information as a client of a legitimate entity with a link to a website that looks like the legitimate entity’s website or with a form contained in the body of the e-mail. The aim of phishing is to steal a user’s identity in order to make fraudulent transactions as if the Phisher were the user. Although there are clear advantages to filtering phishing attacks at the email level, there are at present not many methods specifically designed to target phishing emails, as opposed to spam emails in general. The most closely related prior attempt is in which the authors use structural features of emails to determine whether or not they represent phishing attacks. The features are mostly linguistic and include things such as the number of words in the email, the “richness” of the vocabulary, the structure of the subject line, and the presence of 18 keywords. Other examples include the filter built into Thunderbird 1.5 [21]. However, this filter is extremely simple, looking for only the presence of any one of three features, namely the presence of IP-based URLs, nonmatching URLs (discussed in Section 3.2.3), and the presence of an HTML “form” element. The Thunderbird built-in filter still only presents a warning to the user, and does not avoid the costs of storage and the user’s time. In our implementation and evaluation, we seek to fill this gap in email-based phishing filters. Our approach is generalizable beyond email filtering, however, and we do note how it could be used and what changes would be required in the context of filtering web pages as opposed to emails.

Many people have proposed ways in which to eliminate spam emails in general, which would include phishing emails.

V. RELATED WORK

Current solutions use strong spam filters to isolate phishing solicitations or capture phishing sites at the browser. To name a few are:

- Spam Assasin
- SpoofGuard
- Pilfer
- HoneyTank
- Phishwish
- IDS based phishing attack detection

**Spam Assasin**

Spam Assasin is a tool that recognizes spam containing phishing email. [1] state that SpamAssassin has a false negative of 15% for spam e-mails, and performs worst when tested with 10 fold cross validation. SpamAssassin uses a wide range of heuristic tests on mail headers in order to identify spam, and can be
Intrusion Detection System for Detecting Phishing Attacks is customized. For algorithms, it uses text analysis, Bayesian filtering, DNS blocklists, a collaborative filtering database, and a Stochastic Gradient Descent method in training a neural network. This is used for its scoring based on perception that uses a single perception with a log sig activation function that maps the weights to SpamAssassin's score space. SpamAssassin does not delete email from mailboxes, but it can route classified email to mailboxes or folders.

**Spoof Guard**

Chou et al. [1] proposes a browser-based plug-in, SpoofGuard, that monitors users' Internet activities and warns if the tool classifies a visiting website as a phishing page. SpoofGuard uses the observation that a page is loaded from an e-mail message and whether the URL was visited before. The authors propose the use of the following properties: (1) Logos – use of images. (2) Suspicious URLs – urls that contain IP address or higher length urls. (3) User Input – pages that has form input. (4) Short lived – the spoof sites are shut down with in 2 – 3 days. (5) Copies – similar contents. (6) Sloppiness or lack of familiarity with English – misspellings and grammar errors.

SpoofGuard uses 3 methods to determine impersonation: (1) a stateless method that determines whether a downloaded page is suspicious. (2) a stateful method that evaluates a downloaded page in light of previous user activity, and (3) a method that evaluates outgoing post data. SpoofGuard uses a standard aggregate function to calculate the total spoof score (TSS) computed as: TSS(page)= Σ1 n wiPi + Σ1,1 n,n wi,jPiPj + Σ1,1,1 n,n,n wi,j,kPiPjPpk ....

For a given downloaded web page and a browser state TSS produce a number Pi within [0,1] where 1 indicates a page more likely to be a spoof page. The wi’s are preset weight to minimize false positives. SpoofGuard has a configuration pop-up screen that requires a user defined spoof rating threshold. This allows setting independent weights and security levels for the domain name, url, link, password and image checks. The user interface alerts suspicious sites with a traffic light symbol lighting for the degree of the probable spoof activity. The information which was based for classifying is available for the user. Even though a link from an e-mail is a good method for phishing detection, a user clearing the browser history could result in many false positives. Sensitivity decreasing on this system would result in false negatives while increasing would result in false positives.

**PILFER**

It is a machine-learning based approach to classification [20]. PILFER decides whether some communication is deceptive, i.e. whether it is designed to trick the user into believing they are communicating with a trusted source, when in reality the communication is from an attacker. It makes this decision based on information from within the email or attack vector itself (an internal source), combined with information from external sources. This combination of information is then used as the input to a classifier, the result of which is a decision on whether the input contained data designed to deceive the user. With respect to email classification, it has two classes, namely the
class of phishing emails, and the class of good (“ham”) emails. It identifies some of the email as Phishing email based on some features. These features are

(i) IP based URLs
(ii) Age of linked-to domains
(iii) Non matching URLs
(iv) “Here” links to nonmodal domain

HoneyTank

Honey Tank collects Spam using a honeynet and automatically generates a pattern. The pattern is to be used by a network based intrusion detection systems. A HoneyTank is a workstation receiving TCP segments sent to unallocated IP addresses and replying to those segments to emulate real end systems that supports TCP services. They use Advanced Sequential Analyzer on Unix (ASAX) as the intrusion detection system. ASAX is a generic system that analyzes sequential files like security audits trails. It is composed with three parts which are analyzer, rule declarations, and format adaptor. The analyzer receives the input from the format adaptor and analyzes according to the declared rules.

Phishwish

Its primary goal is to minimize the complexity of the rule-base and configuration, and maximize the number of phishing emails detected while minimizing the number of false positives. Phishwish is applicable to emails that instruct the recipient to log into a web site. It processes text based and HTML formatted emails, although some rules are only applicable to HTML. Each rule is assigned a configurable weight, $W_i$ and a flag $X_i$. Phishwish sets $X_i$ to 1 if the rule is applicable to the email and to 0 otherwise. Each rule produces a value, $P_i$, ranging from 0.0 to 1.0. If the rule is not applicable, $P_i = 0$. The final score is $S = \sum W_i P_i / \sum W_i X_i$, with higher values of $S$ indicating a greater probability of phishing. When describing the rules, a positive result is indicative of phishing, in which case $P_i$ is set to 1 except for rules 8 and 10 where it is set to a fraction. A negative result is indicative of a valid email, in which case $P_i$ is set to 0. Business refers to the business from which the email supposedly has been sent. LoginURL refers to the URL within the email that the recipient should use to access the business’ login page. The rules fall into the following general categories:

1) Identification and analysis of the login URL in the email
2) Analysis of the email headers
3) Analysis across URLs and images in the email
4) Determining if the URL is accessible

These rules are:

Rule 1: If the email appears (based on search engine results) to not be directing the recipient to the actual login page for the business, the result is positive.

Rule 2: In HTML formatted emails, if a URL displayed to the recipient uses TLS, it is compared to the URL in the HREF tag. If the URL in the tag does not use TLS, the result is positive.

Rule 3: If the login URL is referenced as a raw IP address instead of a domain name, the result is positive.

Rule 4: If the business name appears in the login URL, but not in the domain portion, the result is positive.
Rule 5: In HTML formatted emails, if a URL is displayed to the recipient, it is compared to the URL in the HREF tag. If their domains do not match, the result is positive.

Rule 6: The chain of "Received" SMTP headers is checked to determine if the path includes a server or a mail user agent in the same DNS domain as the business. The rule is positive if such a Received header is not present.

Rule 7: Rules 7 and 9 perform a case-insensitive byte-wise comparison of the domain of all URLs in the email message with the domain of the login URL. Rule 7 analyzes non-image URLs for such inconsistencies in their domains. If inconsistencies are detected, the rule is positive. Rule 8: Rules 8 and 10 match the DNS registrant for the domain of each URL in the email with the DNS registrant for the domain in the login URL. Rule 8 analyzes non-image URLs for inconsistencies in their whois registrant information. P8 is set to the percentage of URLs whose information differs from that of the login URL.

Rule 9: This rule analyzes image URLs for inconsistencies in their domains. If inconsistencies are detected, the rule is positive.

Rule 10: This rule analyzes image URLs for inconsistencies in their whois registrant information. P10 is set in the same manner as P8 in Rule 8.

Rule 11: The rule is positive if the web page is inaccessible. The rule is considered not applicable otherwise.

IDS Based Phishing Attack Detection

This solution proposed by Hasika Pamunuwa uses IDS to detect phishing attacks[]. This System architecture has two parts. First part of the architecture seeks emails from the outside of the world and forwards it to the IDS which act as a filter. Once filtering is done; identified phishing emails are saved in database. Second part of the system is validation system it crawls backs the addresses of the suspected emails and validates whether the mail is a genuine phishing mail or not.

If we talk about its filtering system, it uses open source IDS Snort as a filter which on the following rules identifies emails as a phishing email:

1. HTML encoded in e-mail.
2. Any URL including IP addresses.
3. URLs that has been masked with HTML to a different address

Research Gaps

As discussed above most of these solutions are either client side solution a browser based plugin or spam filter.

First solution Spam Assasin[] is a tool for detecting email at the server side which has weakness that it has false negative of 15%

Second solution Spoof Guard[] is a browser based plug-in and can be bypassed by the attackers, it keeps track of browser history to be used in phishing detection, a user clearing the browser history could result in many false positives. As it takes user defined setting, sensitivity decreasing on this system would result in false negatives while increasing would result in false positives.

Fette et al. [] proposes PILFER classify phishing email with a true positive rate of 92% and a false positive rate of 0.1%.
The solution proposed by Hasika Pamunuwa uses IDS to detect phishing attacks[]. Though it has the advantage that it is an organization wide solution but its filtering algorithm is very primitive. This whole observation can be put into the following points

1. There should not be user dependence
2. All the traffic should be evaluated
3. False positive and false negative should be low
4. Filtering rule set should be accurate

VI. CONCLUDING REMARKS

In this paper, we have shown that it is possible to detect phishing emails with high accuracy by using a specialized filter, using features that are more directly applicable to phishing emails than those employed by general purpose spam filters. Although phishing is a subset of spam all, who asks to receive emails from a person pretending to be their bank for the purpose of fraud and identity theft?), it is characterized by certain unique properties that we have identified. One might be inclined to think that phishing emails should be harder to detect than general spam emails. After all, phishing emails are designed to sound like an email from a legitimate company, often a company with which the attacker hopes the user has a pre-existing relationship. Models based on “naïve” assumptions, such as certain words like "Viagra" being indicative of a class of un-desirable emails, no longer hold when the attackers are using the same words and the same overall “feel” to lure the user into a false sense of security. At the same time, phishing emails present unique opportunities for detection that are not present in general spam emails. In genera spam emails, the sender does not need to misrepresent their identity. A company offering to sell "Viagra" over the Internet does not need to convince potential buyers that they are a pharmacy that the user already has a relationship with such as CVS or RiteAid. Instead, a spammer can actually set up a (quasi-)legitimate company called Pharmacy1283, and identify themselves as such, with no need to try to convince users that they are receiving a communication from their bank, or some other entity with which they have an established relationship. It is this misrepresentation of sender identity that is key to the identification of phishing emails, and further work in the area should concentrate on features to identify this deceptive behavior. As the phishing attacks evolve over time to employ alternate deceptive behaviors, so does the information available to combat these attacks. The approach used is flexible, and new external information sources can be added as they become available. These sources could take the form of web services, or other tagged resources, to provide additional information to the decision making process. Many phishing attacks include copies of corporate logos, and if one could map a logo back to its legitimate owner’s website, that would be valuable information in determining the authenticity of a website or email displaying that logo. As image sharing and tagging services such as Flickr [29] are increasing in use, it is not unreasonable to think that some day in the near future, one might actually be able to search with an image and get back a description as a result. There are a number of emerging technologies that could greatly assist phishing classification that we have not considered. For instance, Sender ID Framework (SIDF) [19] and DomainKeys [28], along with other such sender authentication technologies,
should help to both reduce false positives and make detection of spoofed senders much simpler in the time to come. Looking farther into the future, deeper knowledge-based models of the user and the types of prior relationships she may or may not have with different sites or organizations could also help fend off more sophisticated phishing attacks. Such techniques would likely build on ongoing research on federated identities and semantic web technologies [14]. In the meantime, however, we believe that using features such as those presented here can significantly help with detecting this class of phishing emails. We are currently in the process of building a live filtering solution based around PILFER, which we will start making available to users for testing for further validation.

REFERENCES


30. CBL. http://cbl.abuseat.org/ [July 2009].


