Johns Hopkins University
Security Privacy Applied Research Lab

A Framework for Detection and Measurement of Phishing Attacks

Technical Report
SPAR-JHU:SD-NP-MC-AR:251206

Submitted as SPAR Technical Report on December 6th, 2006

A Framework for Detection and Measurement of Phishing Attacks

Sujata Doshi
Johns Hopkins University
sdoshi@cs.jhu.edu

Niels Provos
Google Inc.
niels@google.com

Monica Chew
Google Inc.
mmc@google.com

Aviel D. Rubin
Johns Hopkins University
rubin@jhu.edu

Abstract

Phishing is combination of social engineering techniques and sophisticated attack vectors used to harvest personal identity and financial information from unsuspecting consumers. Most often a phisher tries to lure her victim into clicking a URL pointing to the rogue page. In this paper, we focus on studying the URLs employed in various phishing attacks. We find that it is often possible to tell that a URL belongs to a phishing attack without requiring any knowledge of the corresponding page data and describe several features that can be used to distinguish a phishing URL from a non-phishing one. These features are used to model a logistic regression filter that is efficient and has a high accuracy. Finally, we use this filter to analyze traffic and report the state of phishing on the Internet today.

1 Introduction

The term *phishing* originates from the analogy that Internet criminals use email baits to *fish* for passwords and financial data from a sea of unaware consumers [21]. The term was coined around 1996 in a Usenet article and relates to the theft of AOL passwords and corresponding accounts [4].

Over the decade the definition of phishing has expanded. In addition to the phishing emails, attackers now use attack vectors like fake web sites, Trojan horse key loggers and man in the middle data proxies to trick the users. More often the victim is a user of online banking, payment services such as PayPal, and online e-commerce sites.

Phishing attacks are growing rapidly by the day. The Anti Phishing Work Group detected a total of 14,191 unique phishing attacks in July 2006 [12]. Sophos, an anti-virus company, claims that freely downloadable do-it-yourself phishing kits exist [27]. Consequently anyone surfing the web can now get their hands on these kits and launch their own phishing attack. These kits are supposed to contain all the graphics, web code and text required to construct bogus web sites designed to have the same look-and-feel as legitimate online banking sites. They also include spamming software which enables potential fraudsters to send out hundreds of thousands of phishing emails as bait for potential victims. These numbers and technology indicate the need for improved phishing detection and prevention and also a need for increased awareness amongst the target masses.

> Subject: FPA NOTICE : eBay User Suspension Dear eBay user,
> We regret to inform you that your eBay registration has been suspended for an indefinite amount of time due to the violation of our site policy below:
> . . .
> Please provide us the necessary information in order to prove the ownership of the account, by logging directly to our safe and trust link. **https://signin.ebay.com/ws/eBayISAPI.dll?SignIn**
>
> . . .
> Information will be provided at the request of law enforcement agencies to ensure that perpetrators are prosecuted to the fullest extent of the law.
> Best Wishes,
> eBay Trust and Safety
> eBay International AG

Figure 1: A Typical Phishing Email

In this paper we address these problems by presenting a framework of identifying phishing attacks in URLs. We categorize the various types of phishing attacks in URLs and determine several features which can distinguish a phishing URL from a non-phishing one. These features are used in a logistic regression classifier. Our classifier achieves accuracy up to 97.3%, indicating that URL analysis alone can obtain a high degree of accuracy in phishing detection. Further, we use our classifier to perform detailed measurements on 12 days of Google Toolbar URLs (from August 20th to August 31st 2006) to understand the status of phishing in the Internet today.

2 Background on Phishing

2.1 The Phishing Cycle

Phishing can be described as a three step process: *Planning, Attack*, and *Fraud*. Each step is described as follows.

1. **Planning.** In this phase the attacker determines the victim to attack; the information to be obtained from the victim; and how to obtain this information. Social engineering techniques are employed to gain information about the target victim. Various media, for example phone, instant messenging, clients, email, and the Internet, can be used to gain this information.

2. **Attack.** This phase involves delivery of the phishing message and luring the victim to give up his/her credentials. Email is a popular method used to deliver the phishing message to the target. The layout of a typical phishing email is shown in Figure 1. This mail is targeted at gaining financial information from eBay clients.

3. **Fraud.** The final step of the attacker is fraud. The attacker uses the information obtained in the attack phase to buy goods, steal money from the victims account and identity theft.

This three-step process does not stop after one attack. It is a continuing process wherein the attacker repeats the same steps with another unsuspecting victim.

Dear eBay user,
. . .
Please provide us the necessary information in order to prove the ownership of the account, by logging directly to our safe and trust link.
<a target =“ _blank ”
href=“**http://84.244.5.117/www.ebay.com/ws/verify.html**”
>
https://signin.ebay.com/ws/eBayISAPI.dll?SignIn </a>
. . .
Information will be provided at the request of law enforcement agencies to ensure that perpetrators are prosecuted to the fullest extent of the law.
Best Wishes,
eBay Trust and Safety
eBay International AG

Figure 2: HTML content of the Phishing Email

Table 1: Commonly Used URL Obfuscation techniques

Obfuscation Type	**Descriptive Examples**
Type I	`http://210.80.154.30/~test3/.signin.ebay.com/ebayisapidllsignin.html` `http://0xd3.0xe9.0x27.0x91:8080/.www.paypal.com/uk/login.html`
Type II	`http://21photo.cn/https://cgi3.ca.ebay.com/eBayISAPI.dllSignIn.php` `http://2-mad.com/hsbc.co.uk/index.html`
Type III	`http://www.volksbank.de.custsupportref1007.dllconf.info/r1/vm/` `http://sparkasse.de.redirector.webservices.aktuell.lasord.info/`

Type IV	`http://www.wamuweb.com/IdentityManagement/`
	`http://www.walspring.com/update-wells-info/index.html`
	`http://mujweb.cz/Cestovani/iom3/SignIn.html?r=7785`

In Figure 2 we depict the HTML content for the phishing email in Figure 1. An important point to notice is that the target URL is being obfuscated using host-based obfuscation. In our paper we identify several such URL obfuscation attacks which are detailed in the following Section.

2.2 Categories of Phishing Attacks in URLS

There are varied methods of constructing a phishing URL. Each method involves some form of obfuscation. Below we describe the most prominent techniques used.

- **Type I: Obfuscating the Host with an IP address.** In this form of attack the url's hostname is replaced with an IP address and usually the organization being phished is placed in the path. Very often the IP address is also represented in hex or decimal rather than the dotted quad form.
- **Type II: Obfuscating the Host with another Domain.** In this form of attack the url's host contains a valid looking domain name and the path contains the organization being

phished. This form of attack usually tries to imitate urls containing a redirect so as to make it appear valid.

- **Type III: Obfuscating with large host names.** This form of attack has the organization being phished in the host but appends a large string of words and domains after the host name.
- **Type IV: Domain unknown or misspelled.** Here the URL either does not indicate which organization is being phished or the domain name is misspelled.

Table 1 gives descriptive examples for each type mentioned above. These obfuscation types are also used as features in our classifier.

3 Our Approach

Our approach towards phishing detection involved identifying certain distinguishing features of phishing URLs and using these features in a logistic regression model. For training the model we created a training blacklist and whitelist as follows:

Training Blacklist: The current phishing protection in Firefox uses a blacklist based approach [26]. The blacklist is maintained by Google and is continuously updated. It contains a list of known phishing URLS. We use 1245 urls from this blacklist as our training blacklist.

Training Whitelist: We used a list of the top 1000 most popular URLs as the basis of our training white list set. We further added some URLs which had lower popularity and were not phishing urls to the training set. This brought our training white list to 1263 urls.

We further used the following Google infrastructure to obtain some of our features:

White Domain Table: This table maintains a white list of top level domains. We further use this table to create a list, the *Target Organization List*, of organizations which are known phishing targets. The *Target Organization List* includes various banks and also organizations like Paypal, Ebay, Amazon. We also use the *White Domain Table* in one of our URL features which will be explained in Section 3.1

Google's Index Infrastructure: We use Google's index [1] infrastructure to obtain features of a page. These features include page rank, page index and page quality scores. These are precomputed during Google's crawl phase and are currently stored in a table maintained by Google. The table is indexed by URL. We will call this table the *Crawl Database* henceforth.

In the following section we detail the various features we have identified.

3.1 Feature Analysis

A phishing URL and the corresponding page have several features which can differentiate it from a non-phishing URL. We have categorized our features into 4 groups; Page Based Features, Domain Based Features, Type Based Features and Word Based Features.

Page Based Features In this section we describe features which are properties of a phishing page.

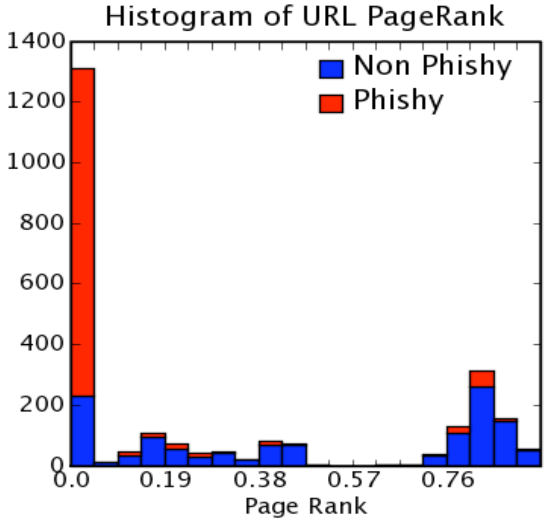

Figure 3: URL Page Rank

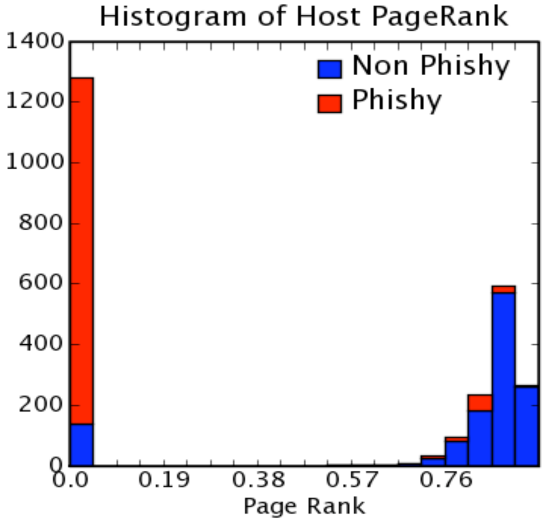

Figure 4: Host Page Rank

- PageRank. PageRank [22] is a numeric value on a scale of [0,1] that represents how important a page is on the web. The higher the PageRank, the more important the page. Phishing web pages most often either have a very low page rank or their page rank does not exist in the *Crawl Database*. Very few phishing pages manage to increase their page rank, possibly by using link spamming techniques. In this paper we consider link spamming as an orthogonal problem. We have identified 3 page rank features — Page Rank of URL, Page Rank of Host and whether the Page Rank is present in *Crawl Database*.

  Figures 3 and 4 show the histograms for the first two features on our training data[1]. Note that almost all of the non-phishing URLs in Figure 4 have a hostname PageRank value in the range of [0.75,1] indicating that hostname PageRank value is a strong feature for identifying if a URL is non-phishing.

- Page Index. Phishing web pages usually are accessible for a short period only and hence many might not be found in the index. We use this property as a feature; that is given a URL we use the *Crawl Database* to test if the corresponding page is found in the index or not. The histogram in Figure 5 for this feature confirms our conjecture that phishing pages are not long lived.

- Page Quality In [11] Google recommends quality guidelines that a site can use to aviod being removed entirely from the Google index. The *Crawl Database* also maintains scores that quantify the quality of the page. The higher the score the better the page quality. We found that phishing pages would either have a low quality score or have no scores at all. These scores are used as features in our classifier.

Domain Based Features This category contains only one feature; is the URL's domain name found in the *White Domain Table*? Phishing url domains are usually either obfuscated or unknown.

[1] Note that for those URLs where the page rank could not be determined, we assign them a page rank of 0.

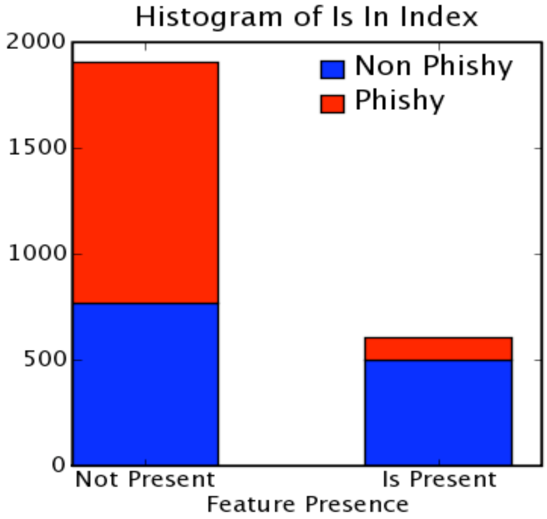

Figure 5: Is the Page In Index

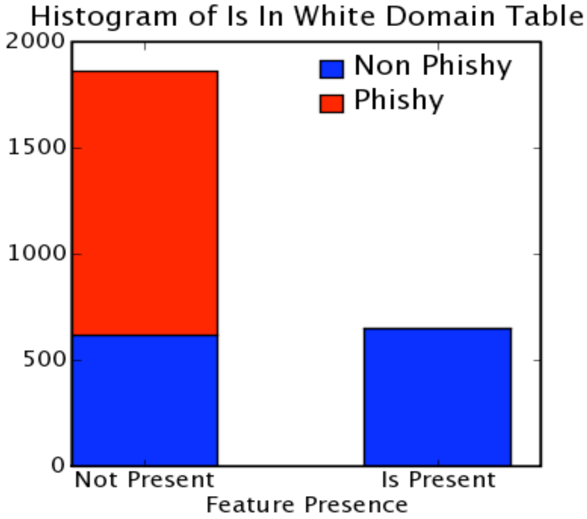

Figure 6: Is the URL's domain in the *white domain table*

The histogram in Figure 6 confirms that the domains of all the phishing URLs in our training data were not found in the *White Domain Table*, indicating that presence in the table is a strong feature for identifying if a URL is not phishing.

Type Based Features In this section we describe features which were based of the various categories of obfuscation identified in Section 2.2.

1. *Host is obfuscated with IP address.* The histogram for this feature is shown in Figure 7. It can be observed that all non-phishing URLs in our training data do not contain host obfuscation and a significant portion of the phishing URLs are host obfuscated with an IP address. This feature is very useful in identifying if a URL is phishing.

2. *Organization being phished is present in the path.* We iterate through the *Target Organization list*, and check if any of the organization are found in the URLs path but not in its host. Note that the *Target Organization List* is organized by length of the organization so as to perform a search for the longest organization first. The histogram for this feature is in Figure 8

3. *Number of characters present after organization in host.* This feature is used to identify URLs that fall under Type III (refer to Table 1). We obtain the URLs hostname and use the *Target Organization List* to check if an organization is found in the hostname. If it is then we determine the number of characters from the end of the organization to the end of the hostname. From our training data we found that usually a non-phishing URL has a path separator after the organization, `http://by124fd.bay124.hotmail.msn.com/cgi-bin/getmsg` is such a URL where the number of characters after msn.com in the hostname is 0. On the other hand Type III phishing URLs, in our training set, are found to have a large number of characters after the organization. `http://www.volksbank.de.`

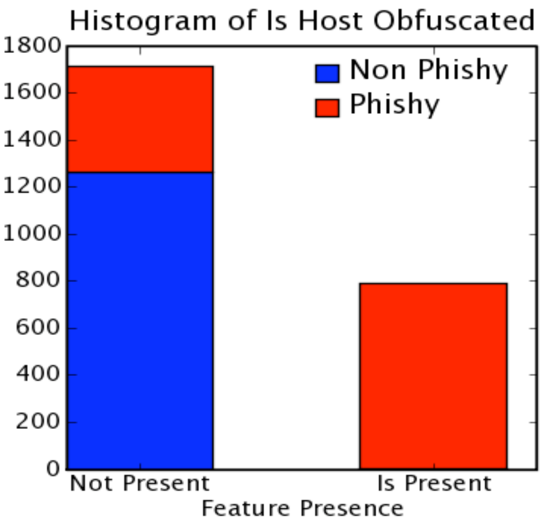

Figure 7: Is URL's host obfuscated with IP address

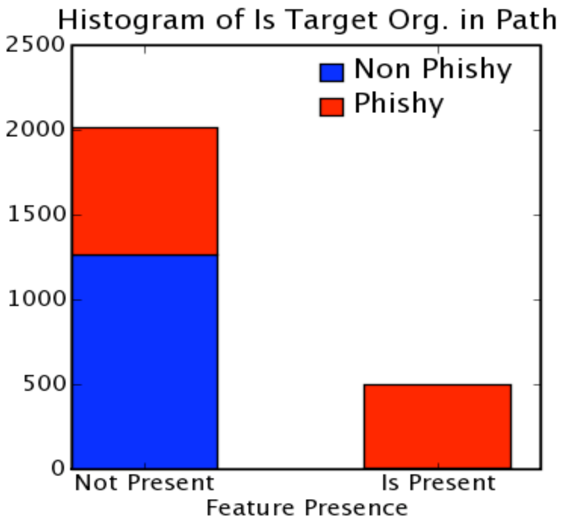

Figure 8: Is organization being phished found in path

`anmeldung-neueschutzmassnahmen.inetbanking.gad.de.storephp.com/index.html` is one such URL where 63 characters are found in the URLs hostname after volksbank.de. Table 2 gives the minimum, maximum and average number of characters for our training data.

Table 2: Number of characters after organization in host

Data Set	Min	Max	Average
Black List	0	63	7.34
White List	0	14	0.21

Word Based Features Phishing URLs are found to contain several suggestive word tokens. For example the words `login` and `signin` are very often found in a phishing URL. In order to obtain all such tokens we used substring extraction [23] over the URLs in our training blacklist. Using the substring extraction algorithm in [23] we obtained 150 tokens and their frequencies. We discarded all tokens with length < 5 since they contained several common URL parts such as `http://, www, com`. Some tokens also contained organizations like ebay and paypal. We discarded these too, since they were already covered by our *Target Organization List*. We further removed query parameters such as `&UsingSSl=, &errmsg=`. From the remainder we chose the tokens `webscr, secure, banking, ebayisapi, account, confirm, login and signin` as features in our classifier. Our word based features are boolean, that is we test if the given word is present or absent in a URL. Note that these features specially enhance the detection of Type IV phishing URLs which do not contain a target organization or the organization is misspelled. Table 3 shows the distribution of these features in our training set.

We find that the features `login` and `signin` are very prominent in our blacklist. The feature `webscr` is very prominent in Paypal URLs and `ebayisapi` is found in several Ebay URLs both phishing and non-phishing.

Table 3: Number of URLS in training blacklist/whitelist with Word Based Feature present/absent.

Feature	Feature absent		Feature present	
	White List	Black List	White List	Black List
webscr	1259	1068	4	177
secure	1261	1122	2	123
banking	1252	1146	11	99
ebayisapi	1244	1072	19	173
account	1244	1184	19	61
confirm	1260	1192	3	53
login	1230	977	33	268
signin	1251	955	12	290

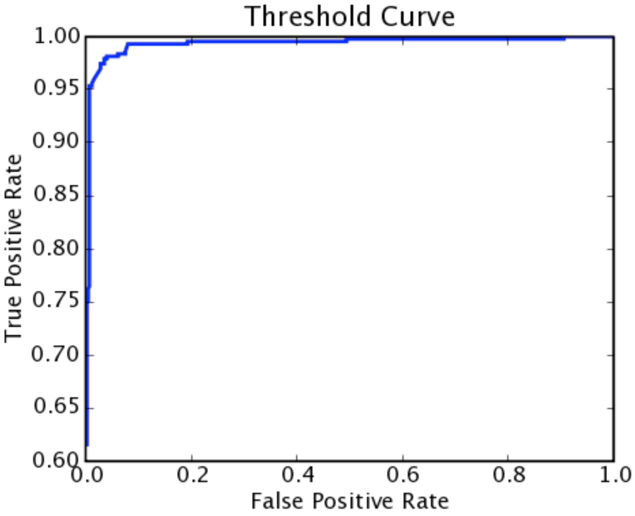

Figure 9: Threshold curve for the phishing URL class.

3.2 Training With Features

In this section we describe the training methodology we took with the features identified in Section 3.1. We used the Weka data mining library [29] for our analysis. Our labeled data consisted of 2508 URLs of which 1245 were phishing URLs and 1263 were non-phishing URLs, as described in Section 3. Phishing URLS were placed under the positive class and Non-phishing ones were under the negative class. We chose a random split of the labeled data — 66% were used for training and the remaining 34% were used as the test set. These sets were disjoint.

We chose the Logistic Regression learning algorithm [8] and used Weka to obtain the coefficients on our 18 features. These coefficients are listed in Table 4 along with the odds ratio. Host obfuscation was found to be the most useful in detecting if a URL is phishing, as indicated by the high odds ratio. Similarly presence in the *White Domain Table* is a strong indicator that a URL is not phishing. We evaluated the trained model on the 34% test set split. This evaluation gave us an accuracy of 97.31% with a True Positive Rate of 95.8 % and False Positive Rate of 1.2% [2].

In order to obtain a better metric of accuracy we also examined the area under the Threshold Curve. This is popularly known as the ROC (Receiver Operating Characteristic) curve [24] and reports each type of error at different probability thresholds. Note that the probability thresholds denote probability of a URL being phishing and for logistic regression they are computed using Equation 1.

$$P(PhishingURL) = \frac{e^{score}}{1 + e^{score}} \tag{1}$$

score denotes the weighted score obtained from the feature values and the regression coefficients.

Figure 9 shows the ROC curve for our logistic regression classifier for the positive class , that is the phishing URL class. The area under the curve is 0.9923 indicating that our classifier has a

[2] Note that we repeated the training multiple times over different random 66% splits and observed that the accuracy and coefficients were consistent over these multiple runs.

Table 4: Features and their Coefficients

Feature	Logistic Coefficient	Odds Ratio $e^{coefficient}$
Is URL in *White Domain Table*	-3.82	0.0219
Quality Score II	-1.9543	0.1417
PageRank of Host	-1.8812	0.1524
PageRank of URL	-1.2606	0.2835
PageRank in *Crawl Database*	-0.536	0.5851
Quality Score I	0.0443	1.0453
Number of characters after organization in host	0.2306	1.2594
Word *secure* presence	0.3328	1.3949
Word *account* presence	0.8589	2.3605
Is Page in Index	0.8738	2.3961
Word *webscr* presence	0.9969	2.7099
Word *login* presence	1.8587	6.4155
Word *ebayisapi* presence	2.1659	8.7221
Word *signin* presence	2.5404	12.685
Word *banking* presence	2.6361	13.9593
Word *confirm* presence	2.7586	15.777
Is *target organization* in path but not in host	2.9464	19.0378
Is host obfuscated with IP	6.3933	597.8151
Constant	-0.5881	

high accuracy of phishing URL detection.

We used the threshold curve to deduce the probability threshold at which the false positives are very low. Specifically, we set the probability threshold to 0.985 at which the False Positive Rate is 0.7% and the True Positive Rate is 88%. and the accuracy is 93.4%.

We used this filter to perform analysis on URLs obtained from Google Safe Browsing toolbar. Note that since our probability threshold is high, we have very few false positives. We detail our experiments and results in the following section.

4 Analysis

In this section we describe our analysis using the filter described in the previous section. Section 4.1 describes our data sources and Section 4.2 describes our experiments and results.

4.1 Data Description

We analyzed URLs from the Google Safe Browsing toolbar during the time period of August 20th to 31st August 2006. The data consisted of two main components:

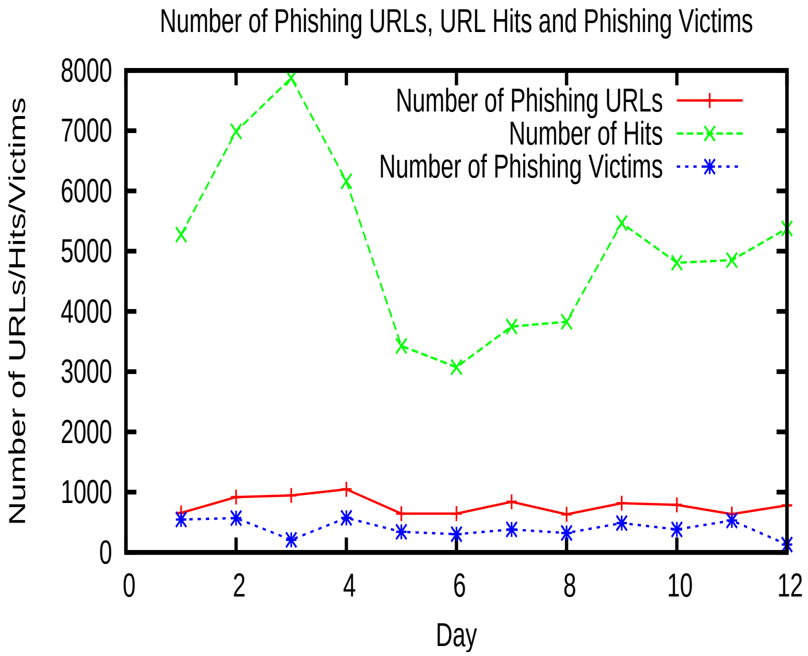

Figure 10: Number of Phishing URLs, Hits to the URLs and Phishing Victims from August 20th to August 31st

- Data Source I: Toolbar URLs. This data included the unique URLs which had been visited each day and the number of visits to each URL. Over 12 days, this amounted to many million URLs.
- Data Source II: Toolbar Lookup Requests. This data included consecutive lookup requests from the toolbar. Each request contained the URL the browser navigated to. To determine which events involved user interaction at a phishing page, we examined all the requests where two consecutive URLs remained on the the same domain. During the time period that we analyze, this amounted to several million URLs.

4.2 Experiments and Results

We conducted 5 main experiments on the data sets described in the previous section. We detail the experiment and the results obtained in the following sections.

4.2.1 Experiment I: Average Phishing URLs Per Day

We used the Logistic Regression Classifier described in Section 3.2 and classified the URLs in Data Source I. Figure 10 shows phishing URLs identified and also the number of viewers (hits) to Phishing URLs on each day.

On average there are 777 URL phishing attacks in a day and from our data set maximum phishing attacks were observed on Day 4 (August 23rd), 2006 and maximum number of viewers

Table 5: Distribution of Obfuscation Types Used

Obfuscation Type	Total Number of URLs	Percentage
Type I	3110	33.32%
Type II	1615	17.30%
Type III	4337	46.46%
Type IV	273	2.9%

Table 6: Phishing URLs (top 10) active across all 12 days

URL	Obfuscation Type
213.92.8.142/Redirect/cgi.ebay.it/ws/eBayISAPI.dll	Type I
211.100.16.141/shopping/PaypalReturn.aspx	Type I
203.143.16.149/icons/small/www.paypal.com/SecureInfo/paypal/index.php	Type I
210.83.203.118/https:/cgi3.ca.ebay.com/eBayISAPI.dllSignIn.php	Type I
www.53.com.wps.portal.secure.pool34.st/context	Type III
210.77.218.175/rdcwwyh/www.paypal.com/update/cgi-bin/index.php	Type I
212.23.177.35:8080/paypal/ReponseDemandeAccordPaypal.do	Type I
www.volksbank.de.vr-web.networld.onlinebanking.glas11.st/anmelden.cgi	Type III
66.207.71.141/signin.ebay.com/Members_Log-in.htm	Type I
www.volksbank.de.vr-web.networld.onlinebanking.orange13.st/anmelden.cgi	Type III

(hits) were observed on Day 3 (August 22nd). The graph in Figure 10 also depicts Phishing Victims. These are explained in Section 4.2.3.

4.2.2 Experiment II: Distribution of Obfuscation Types Used

We used the Phishing URLs identified in Experiment I and determined which obfuscation category (Refer to Table 1) each URL falls in. Table 5 shows the distribution. The results are aggregated over all 12 days.

As indicated by the distribution, Type I and Type III are the popular obfuscation techniques used in phishing. The high percentage of host obfuscation attacks (Type I) is also consistent with our feature analysis, which gave the Host Obfuscation feature a high logistic coefficient.

Additionally, we examined the phishing URLs over each day to determine the phishing URLs which stay active across all the 12 days. In other words we determine those phishing URLs whose site has been accessed on each day during the observation period. The top 10 such URLs and their Obfuscation Types after eliminating a few false positives are depicted in Table 6.

4.2.3 Experiment III: Average Phishing Victims Per Day

We used the Logistic Regression Classifier on the Toolbar Lookup Requests (Data Source II) to determine how many users interact with a phishing page and fall for phishing every day. As indicated in Section 4.1, the URLs from this data source are those where a user has interacted with the web site pointed to by the URL. In other words, the user has entered some information at this web site.

We use our classifier to determine which of these URLs are phishing URLs, and consequently

Table 7: Success Rate of Phisher from 20th-31st August

Day	Date	Success Rate
1	08/20/06	10.35%
2	08/21/06	8.15%
3	08/22/06	2.65%
4	08/23/06	9.29%
5	08/24/06	9.95%
6	08/25/06	9.75%
7	08/26/06	10.13%
8	08/27/06	8.44%
9	08/28/06	8.87%
10	08/29/06	7.94%
11	08/30/06	10.94%
12	08/31/06	2.43%
Average		**8.24%**

Table 8: Geographical Distribution of Phishing in a day

Country	Number of Phishing URLs
United States	393
Sao Tome and Principe	33
Belize	25
China	16
Germany	7
Taiwan	6
United Kingdom	5
Russian Federation	5
Romania	4

determine how many users have fallen for a phishing page. On average we find that 397 users fall for phishing in a day [3]. Figure 10 indicates the number of users that fell for phishing on each day during the observation period. As can be observed, Day 4 (August 23rd) has the maximum number of phishing victims. In Table 7 we present the Success Rate of the phisher which is computed as $\frac{PhishingVictimCount}{NumberofHits} * 100$. Based on our results we find that on average, 8.24% of the number of hits to a phishing page a phisher succeeds in his attack.

4.2.4 Experiment IV: Distribution of Phishing by Organization

We further analysed our results from Experiment I and Experiment II to determine which organizations were popular phishing targets. We used the *Target Organization List* to determine which organization was the target. Table 9 presents the number of Phishing URL's by organization as well as the number of Phishing Victims. These are averaged over the 12 days of data and the top 20 organizations are displayed. The Unknown category indicates that the organization could not be determined from the URL. Most URLs that fall under this category are the Type IV URLs where the target organization is not in the list or the organization has been misspelled.

As can be observed Ebay and Paypal are the top target organizations and a large percentage of the phishing victims fall for attacks against these two organizations.

4.2.5 Experiment V: Geographical Distribution of Phishing

We used the phishing URLs identified from Experiment I to determine the Geographical distribution of Phishing in a day. To determine country that hosts a particular phishing URL, we used Google's IP to Geo-Location infrastructure. This distribution is depicted in Table 8[4]. The results indicate that a large percentage of phishing attacks are in the United States.

[3] We point out that the results from this experiment are not exact but can be used as a ballpark estimate.

[4] Note that we have displayed the results for the top 10 countries. For some URLs we were unable to determine the origin country — these are not included in our distribution.

Table 9: Distribution of Phishing by Organization.

Organization	Phishing URL	Phishing Victim
Ebay	231	224
Paypal	211	105
Fifth Third Bank	61	0
Unknown	60	32
Bank Of Scotland	32	18
Volksbank	29	0
Wells Fargo	29	0
Bank of America	28	4
Sparkasse	14	3
Private Banking	13	0
HSBC	7	0
Chase	5	1
Amazon	4	1
Banamex	4	0
Barclays	4	0
Credit Union	4	0
E-gold	3	0
Visa	2	0
Deutsche Bank	2	0
Commbank	1	1

We point out that since we set a very high threshold probability for our classifier (Refer to Section 3.2), our analysis is conservative. In other words, as per our true positive rate, 88% of the phishing attacks present in our data sets will be caught by our filter. Further, since our false positive rate is negligible (0.7%), we can infer that the numbers obtained in our analysis are a lower bound on the phishing attacks present in our datasets. This analysis hence confirms that phishing attacks are a very serious threat today.

5 Related Work

There exists a plethora of different techniques in phishing detection. One of the popular methods of detection is using add-in toolbars for the browser. Chou *et al.* introduced one such tool called SpoofGuard which warns users when the visited web site has a high probability of being a spoof [2]. The tool uses domain name, url, link and images to evaluate the spoof probability. Given a downloaded web-page, the plug-in applies a series of tests, each resulting in a number in the range [0,1]. The total score is a weighted average of the individual test results.

Gabber *et al.* present a tool, Lucent Personalized Web Assistant (LPWA) [10, 17], which can protect a clients identity information. They define client persona in terms of username, password and email address and introduce a function which provides a client with different persona for different servers it visits. A similar concept was later presented by Ross *et al.* in [25] in a tool PwdHash.

PwdHash replaces a users password with a one way hash of the password and the domain name

Table 10: Anti-Phishing Tools

Tool	**Primary Feature**	**Limitations**
Google Tool Bar	Uses a blacklist of Phishing URLs to identify a phishing site	Might not recognize phishing sites not present in the blacklist
eBay Tool Bar	Uses Account Guard which warns users through a colored tab which turns red on a spoofed site	Applicable only to Ebay and PayPal
SpoofStick	Provides basic domain information; on Ebay it will display *You are on ebay.com*, on a spoofed site it will display *You are on 20.2.40.10*	Not very effective against spoofed sites opened in multiple frames
NetCraft	Risk rating system used. Dominant factor in computing risk is age of the domain name.	Part of their technique involves using a database of sites, and hence might not recognize new phishing sites successfully.
SiteAdvisor	Primarily protects against spyware and adware attacks. Based on using bots to create a huge database of malware and test results on them to provide ratings for a site	As in the case of NetCraft, if a new phishing site does not have a rating in their database it might not be caught by this tool.

[25]. While this is a simple technique to protect against password phishing, it is not secure against offline dictionary attacks, key logger attacks, DNS cache poisoning attacks and cannot be securely applied when the user does not have the privileges to install the tool on the computer. Other anti-phishing tools include Google Safe Browsing [26], eBay Tool bar [9], SpoofStick [5], NetCraft tool bar [20] and SiteAdvisor [18]. These are summarized in Table 10.

Dhamija *et al.* propose Dynamic Security Skins [6], another class of browser based defences against phishing. This solution is an implementation of their prior work on Human Interactive Proofs [7], which allows a human to distinguish between legitimate and spooofed web sites. Dynamic Security Skins allow a remote server to prove its identity in a way which supports easy verification by humans, but is hard to spoof by attackers [6].

This is achieved by having a trusted browser password window and using secure remote password protocol (SRP), a verifier based authentication protocol. Spoofing of the trusted window is made difficult by providing an image which is a shared secret between the user and his browser. The image is chosen by the user (or randomly assigned) and overlaid across the window and text entry boxes. Further more during each transaction the server generates an image which is used to create the skin for the browser. To verify the server, the user has to visually verify if images match. Currently this protocol does not provide security for situations where the user login is from a public terminal. Further more it does not protect against malware and relies on the browser to store important verifier information during the SRP authentication.

In [13], Herzberg *et al.* propose TrustBar, a third party certification solution against phishing. The authors propose creating a Trusted Credentials Area (TCA). The TCA controls a significant area, located at the top of every browser window, and large enough to contain highly visible logos and other graphical icons for credentials identifying a legitimate page. While their solution does not rely on complex security factors, it does not prevent against spoofing attacks. Specifically, since the logos of websites do not change, they can be used by an attacker to create a look alike TCA in an untrusted web page.

Several authentication mechanisms have been deployed against the phishing problem. These mechanisms broadly fall into user authentication, server authentication and email authentication. AOL Passcode is one such user authentication system designed to protect against password phishing. It uses a device which generates a unique six digit numeric code every 60 seconds for login to the AOL web site.

Microsoft implemented an email authentication protocol called SenderID in January 2005 [19]. This addresses the problem of domain spoofing. In the SenderID Framework, domain administrators publish Sender of Policy Framework (SPF) records in the Domain Name System (DNS) which identify authorized outbound e-mail servers. Receiving e-mail systems verify whether messages originate from properly authorized outbound e-mail servers. If the sending e-mail server's IP address matches any of the IP addresses that are published in the SPF record then the email is considered to be authenticated and is delivered to the recipient.

Yahoo implemented a domain level email authentication protocol called DomainKeys [15, 14]. The domain owner generates the public/private key pair used for signing all outgoing messages. The receiving email system applies local policies based on the results of the signature test. If the domain is verified and other anti-spam tests don't catch it, the email can be delivered to the user's inbox. Furthermore Yahoo and Cisco have collaborated their DomainKeys and Internet Identified Mail technology to form DomainKeys Identified Mail [3]. While this unification might seem to be a big step towards defending against email spoofing attacks one must question its scalability. The protocol is only applicable within closely knit domains which prevents it from being globally deployed.

Visual similarity is another method used for detecting phishing pages. Liu *et al.* [28] compare legitimate and spoofed web pages and definesimilarity metrics. The spoofed web page is reported as a phishing attack if the visual similarity is higher than its corresponding preset threshold.

More recently, Jakobsson introduced a new model, called a *phishing graph*, to visualize the flow of information in a phishing attack [16]. In a phishing graph the nodes correspond to knowledge or access rights and directed edges correspond to means of obtaining information or access rights from existing information or access rights. The edges can also be labeled with the cost of traversing the edge. One of the nodes corresponds to the target resource to be accessed. An attack is said to be successful if there exists a path from the starting state to this target node. The author describes several example attacks and the corresponding phishing graphs for these attacks.

While this model is not in its essence a defensive technique, it is the first step towards developing an abstract model for visualizing phishing attacks. Given a phishing graph one can perform economic analysis of an attack and quantify the cost in terms of total traversal from the start node to the target node. By introducing this model the ability to understand the threat involved increases.

6 Conclusion

In this paper we have identified several new features for identifying phishing URLs. We show that by only using URL analysis a high percentage of phishing attacks can be caught. In comparison with other URL classification techniques, ours includes more generic features such as PageRank and Index information which enhance the detection accuracy. We find the strongest feature for detecting if a URL is phishing is examining if the host has been obfuscated.

We have used our classifier to quantify the amount of phishing on the Internet today. On average we found around 777 phishing pages per day and around 397 users per day are tricked by phishing. Additionaly, Type I (Host Obfuscation) and Type III (Large Host Names) are the popularly used obfuscation techniques today. Ebay and Paypal were the top phishing targets and most of the phishing sites were concentrated in the US. These measurements validate the seriousness of phishing as a threat today.

While our detection mechanism can catch a considerable percentage of phishing URLs, we hope to improve it more in the future. One possible area of open research is to determine features from HTML content of phishing web sites. Furthermore, we seek to explore methods of combining our feature analysis with distinguishing features of phishing emails, to enhance detection of phishing emails.

7 Acknowledgments

This work was carried out at Google Inc., Mountain View. We thank all members of the Safe Browsing team at Google, for their useful feedback. We thank Peter Norvig and Fabian Monrose for reviewing the paper and providing insightful comments on this work.

References

[1] Nancy Blachman. Google guide, making searching even easier. `http://www.googleguide.com/google_works.html`.

[2] Neil Chou, Robert Ledesma, Yuka Teraguchi, Dan Boneh, and John Mitchell. Client-side defense against web-based identity theft. In *11th Annual Network and Distributed System Security Symposium (NDSS '04), San Diego*, 2004. `http://crypto.stanford.edu/SpoofGuard/#publications`.

[3] CISCO. Domainkeys identified mail available royalty-free to the industry-at-large. `http://newsroom.cisco.com/dlls/2005/prod_060105d.html`.

[4] Richard Clayton. Insecure real world authentication protocols (or why is phishing so profitable), 2005. `http://www.cl.cam.ac.uk/users/rnc1/phishproto.pdf`.

[5] CoreStreet. Spoofstick. `http://www.corestreet.com/spoofstick`.

[6] Rachna Dhamija and J. D. Tygar. The battle against phishing: Dynamic security skins. In *SOUPS '05: Proceedings of the 2005 symposium on Usable privacy and security*, pages 77–88, New York, NY, USA, 2005. ACM Press.

[7] Rachna Dhamija and J. D. Tygar. Phish and hips: Human interactive proofs to detect phishing attacks. In *Human Interactive Proofs*, pages 127–141, 2005.

[8] D.W.Hosmer and S. Lemeshow. *Applied Logistic Regression.* Wiley, New York, USA, 1989.

[9] eBay. ebay toolbar. http://pages.ebay.com/ebay_toolbar/.

[10] Eran Gabber, Phillip B. Gibbons, David M. Kristol, Yossi Matias, and Alain Mayer. Consistent, yet anonymous, web access with lpwa. *Commun. ACM*, 42(2):42–47, 1999.

[11] Google. Webmaster guidelines. http://www.google.com/support/webmasters/bin/answer.py?answer=35769.

[12] Anti Phishing Work Group. Phishing activity trends report, July 2006. http://www.antiphishing.org/reports/apwg_report_july_2006.pdf.

[13] Amir Herzberg and Ahmad Gbara. Trustbar: Protecting (even nae) web users from spoofing and phishing attacks. Cryptology ePrint Archive, Report 2004/155, 2004. http://eprint.iacr.org/.

[14] Yahoo! Inc. Domain-based email authentication using public-keys advertised in the dns (domainkeys). Internet Draft. http://www.ietf.org/internet-drafts/draft-delany-domainkeys-base-06.txt.

[15] Yahoo! Inc. Domainkeys: Proving and protecting email sender identity. http://antispam.yahoo.com/domainkeys.

[16] Markus Jakobsson. Modeling and preventing phishing attacks. *Phishing Panel of Financial Cryptography*, 2005.

[17] D. Kristol, E. Gabber, P. Gibbons, Y. Matias, and A. Mayer. Design and implementation of the lucent personalized web assistant, 1998.

[18] McAfee. Mcafee siteadvisor. http://www.siteadvisor.com/.

[19] Microsoft. Microsoft delivers new tools to help reduce spam, 2005. http://www.wwwcoder.com/main/parentid/282/site/5204/266/default.aspx.

[20] NetCraft. Netcraft anti-phishing tool bar. http://toolbar.netcraft.com/.

[21] Gunter Ollmann. The phishing guide, understanding and preventing phishing attacks, 2004. www.nextgenss.com/papers/NISR-WP-Phishing.pdf.

[22] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project, 1998.

[23] Niels Provos, Joe McClain, and Ke Wang. Search worms. In *WORM '06: Proceedings of the 4th ACM workshop on Recurring malcode*, pages 1–8, New York, NY, USA, 2006. ACM Press.

[24] Foster J. Provost, Tom Fawcett, and Ron Kohavi. The case against accuracy estimation for comparing induction algorithms. In *ICML*, pages 445–453, 1998.

[25] Blake Ross, Collin Jackson, Nicholas Miyake, Dan Boneh, and John Mitchell. A browser plug-in solution to the unique password problem. In *Proceedings of 2005 USENIX Security Symposium*, 2005. http://crypto.stanford.edu/PwdHash/pwdhash.pdf.

[26] Fritz Schneider, Niels Provos, Raphael Moll, Monica Chew, and Brian Rakowski. Phishing protection design documentation, 2006. `http://wiki.mozilla.org/Phishing_Protection:_Design_Documentation`.

[27] Sophos. Do-it-yourself phishing kits found on the internet, reveals sophos, 2004. `http://www.sophos.com/pressoffice/news/articles/2004/08/sa_diyphishing.html`.

[28] Liu Wenyin, Guanglin Huang, Liu Xiaoyue, Zhang Min, and Xiaotie Deng. Detection of phishing webpages based on visual similarity. In *WWW '05: Special interest tracks and posters of the 14th international conference on World Wide Web*, pages 1060–1061, New York, NY, USA, 2005. ACM Press.

[29] Ian H. Witten and Eibe Frank. *Data Mining: Practical machine learning tools and techniques.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.