A Framework for Detection and Measurement of Phishing Attacks

Sujata Garera
Johns Hopkins University
Baltimore, MD 21218
sgarera@cs.jhu.edu

Niels Provos
Google Inc.
Mountain View CA 94043
niels@google.com

Monica Chew
Google Inc.
Mountain View CA 94043
mmc@google.com

Aviel D. Rubin
Johns Hopkins University
Baltimore, MD 21218
rubin@jhu.edu

ABSTRACT

Phishing is form of identity theft that combines social engineering techniques and sophisticated attack vectors to harvest financial information from unsuspecting consumers. Often a phisher tries to lure her victim into clicking a URL pointing to a rogue page. In this paper, we focus on studying the structure of URLs employed in various phishing attacks. We find that it is often possible to tell whether or not a URL belongs to a phishing attack without requiring any knowledge of the corresponding page data. We describe several features that can be used to distinguish a phishing URL from a benign one. These features are used to model a logistic regression filter that is efficient and has a high accuracy. We use this filter to perform thorough measurements on several million URLs and quantify the prevalence of phishing on the Internet today.

Categories and Subject Descriptors

K.4 [**Computers and Society**]: Electronic Commerce; K.4.4 [**Electronic Commerce**]: [Security]

General Terms

Measurement, Security

Keywords

Phishing, URL Obfuscation

1. INTRODUCTION

Phishing, a term coined in 1996, was originally used to describe theft of AOL passwords and corresponding accounts [3]. However, over the decade the definition of phishing has expanded. Attackers today use attack vectors such as email,

trojan horse key loggers and man in the middle data proxies to trick the users. More often than not, the victim is a user of online banking sites or 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 27,221 unique phishing URLs in January 2007. Sophos, an anti-virus company, claims that freely downloadable do-it-yourself phishing kits exist [22]. 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.

We find that existing anti-phishing techniques, whether third party certification based [12], password based [20] or URL based [16] are not robust enough for phishing detection. In this paper we focus on improving URL based detection of phishing sites. While techniques that rely on retrieving content pointed to by a URL might have better detection rates, it is important to note that these techniques can result in undesired side effects. For instance, fetching content could potentially trigger actions (such as signing up to a mailing list or even acknowledging receipt of a credit card) on behalf of the user. Such unwanted events can be avoided through relying on classification schemes that are based only on examining the URL.

In this paper we identify several fine-grained heuristics that can be used to distinguish between a phishing URL and a benign URL. These heuristics are used to model a logistic regression classifier [9]. In addition to obfuscation style heuristics, which have been considered in previous work, our classifier also incorporates several general heuristics based on Google's Index Infrastructure. As a result, our classification technique achieves an accuracy of 97.3%, thus demonstrating that URL analysis alone can achieve a high degree of accuracy in phishing detection.

In an effort to better understand the prevalence of phishing on the Internet, we use our classifier to perform detailed

Table 1: Commonly Used URL Obfuscation techniques

Type	Descriptive Examples
I	`http://210.80.154.30/~test3/.signin.ebay.com/ebayisapidllsignin.html`
	`http://0xd3.0xe9.0x27.0x91:8080/.www.paypal.com/uk/login.html`
II	`http://21photo.cn/https://cgi3.ca.ebay.com/eBayISAPI.dllSignIn.php`
	`http://2-mad.com/hsbc.co.uk/index.html`
III	`http://www.volksbank.de.custsupportref1007.dllconf.info/r1/vm/`
	`http://sparkasse.de.redirector.webservices.aktuell.lasord.info`
IV	`http://www.wamuweb.com/IdentityManagement/`
	`http://mujweb.cz/Cestovani/iom3/SignIn.html?r=7785`

measurements on 12 days of Google Toolbar URLs (from August 20th to August 31st 2006). Our data set comprises of several million URLs. We quantify the average number of phishing attacks and the number of potential phishing victims in a day. We find that current literature in phishing lacks such a large scale and thorough measurement study. Our findings confirm the seriousness of the phishing threat today.

2. PHISHING URL TYPES

In a phishing attack, an adversary typically lures the victim into clicking a URL pointing to the phishing site. The adversary usually obfuscates this URL through varied methods. To determine the popular obfuscation techniques currently in use, we examined a black list of phishing URLs maintained by Google. This black list is used to provide phishing protection in Firefox [21]. The prominent obfuscation techniques are:

- **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 there is no apparent relationship to the organization being phished or the domain name is misspelled.

Table 1 provides some illustrative examples for each of the aforementioned types. These obfuscation types are also used as features in our classifier, described in Section 3.1.

3. MODELING PHISHING URLS

Based on the categories of phishing URLs found in the previous section, we identify distinguishing features of phishing URLs and use the features to model a logistic regression classifier. We choose logistic regression because it is computationally efficient to evaluate. For training the model we created a training black list and white list as follows:

Training Black list: As indicated previously, Google maintains a black list of URLs for providing phishing protection to Firefox users [21]. This black list is continuously updated through commercial sources as well as internal Google sources. It is thoroughly examined by human experts to eliminate any noisy (non-phishing) URLs. We use 1245 URLs from this list as our training black list.

Training White list: We used a list of the top 1000 most popular URLs as the basis of our training white list set. This list was collected through a variety of means, including manual selection as well as algorithmic techniques that collate signals to detect high quality URLs. We further added URLs which had lower popularity to the training set. These were mainly non-phishing URLs that contained redirects. This brought our training white list to 1263 URLs.

Additionally, we take advantage of Google's infrastructure to create two data structures:

White Domain Table: This table maintains a white list of top level domains. We use this table to create the *Target Organization List*, a list of organizations which are known phishing targets. This includes various banks and also organizations like Ebay.

Crawl Database: We use Google's index [1] infrastructure to obtain features of a page. These are precomputed during Google's crawl phase and maintained internally by Google in the crawl database. This table is indexed by URL.

3.1 Feature Analysis

A phishing URL and the corresponding page have several features which can differentiate it from a benign URL. These include obfuscation style features which were described in Section 2. It is important to note that an adversary might be able to easily evade URL detection mechanisms that rely on only obfuscation based features. He would accomplish this by crafting the URL very carefully. Our classifier, however also incorporates several generic features and hence is more robust against an adversary seeking to avoid detection. We categorize our features into four groups: Page Based, Domain Based, Type Based and Word Based features.

Page Based. These include the Page Rank [17] of a web page, its presence in the index and its quality. We discuss each in turn below:

- Page Rank [17] is a numeric value on a scale of [0,1] that represents the relative importance of a page within a set of web pages. The higher the Page Rank, the more im-

portant the page. Phishing web pages are short lived and thus either have a very low Page Rank or their Page Rank does not exist in the *Crawl Database*. Three page rank features that provide discriminatory power are the Page Rank of URL, Page Rank of Host and whether the Page Rank is present in *Crawl Database.*
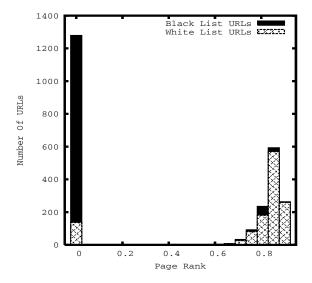

Figure 1: Page Rank distribution for the white list and black list URLs hostname

Figure 1 shows the histogram for the host page rank on our training data[1]. Almost all of the white list URLs in Figure 1 have a hostname Page Rank value in the range of [0.75,1], indicating that hostname Page Rank value is a strong feature for identifying if a URL is non-phishing.

- Phishing web pages are usually accessible for only a short period of time; thus, many might not be found in the index. We use this property as a feature. Given a URL, we use the *Crawl Database* to test if the corresponding page is found in the index or not. We found that 39.5% of the white list URLs are present in the index. On the other hand, only 8.2% of the black list URLs are found in the index. These numbers confirm our conjecture that phishing pages are short lived.

- Google recommends quality guidelines [11] that a site can use to avoid being removed entirely from the Google index. The *Crawl Database* also maintains scores that quantify the quality of the page. These scores measure if the web masters comply with the guidelines in [11]. 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.

Domain Based . This category contains only one feature: whether or not the URL's domain name can be found in the *White Domain Table.* Phishing URL domains are usually either obfuscated (Type I, II and III) or unknown (Type IV). As suspected, a considerable percentage (51.2%) of the white list URLs were present in the table and only 0.2% of

---

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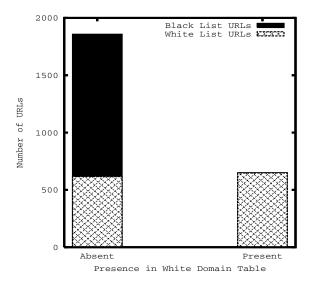

Figure 2: Histogram for the Domain Based feature. The figure depicts the number of URLs whose domain is present or absent in the *White Domain Table*

the black list URLs were found in this table. The results are depicted in the histogram in Figure 2. These results indicate that presence in the *White Domain Table*, is a strong feature for identifying if a URL is not phishing.

Type Based. The obfuscation types identified in Section 2 can also be used as features:

- *Type I URL.* Figure 3 shows the histogram for host based obfuscation. Almost all non-phishing (white list) 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.

- *Type II URL.* We iterate through the *Target Organization list* to check if any of the organizations are found in the URL's path but not in its host. The *Target Organization List* is sorted by length of the organization so as to perform a search for the longest organization first. As depicted in Figure 3 considerable portion of the black list URLs are Type II URLs.

- *Type III URL.* To identify this category of URLs we determine the number of characters present after an organization in the hostname. We obtain the URL's hostname and use the *Target Organization List* to check if an organization is found in the hostname. We then determine the number of characters from the end of the organization to the end of the hostname.

From our training data we found that a non-phishing URL usually has a path separator after the organization, e.g. `http://by124fd.bay124.hotmail.msn.com/cgi-bin/getmsg` has 0 characters after `msn.com` and before the path separator. Furthermore, the maximum number noticed in a white list URL are 14 characters and the average across all white list URLs is 0.21 characters.

On the other hand Type III phishing URLs in our training data have 7.34 characters (on average) after the target or-
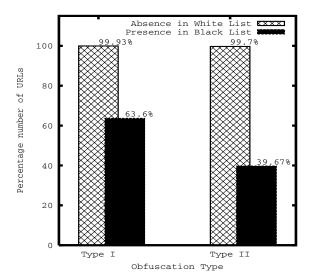
Figure 3: Distribution of Type I and Type II URLs in the training data. Observe that while Type I,II obfuscation is present in a significant percentage of black list URLs, it is practically absent in the white list URLs.

ganization and before the path separator and a maximum of 63 characters.

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 sub-string extraction algorithm by Provos *et al.* [18] over the URLs in our training black list. We obtained 150 tokens and their frequencies. We discarded all tokens with length < 5 since they contained several common URL parts such as `http://`, and `www`. We discarded organization name tokens like Ebay, 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.

Table 2: Distribution of Word Based Features.

Feature	Distribution of Feature Presence	
	White List (%)	Black List (%)
confirm	0.23	4.25
account	1.5	4.9
banking	0.87	7.95
secure	0.16	9.88
ebayisapi	1.5	13.9
webscr	0.32	14.2
login	2.61	21.53
signin	0.95	23.29

Our word based features are boolean, we test if the given word is present or absent in a URL. These features specially enhance the detection of Type IV phishing URLs which do not contain a target organization or the organization is mis-

Table 3: Features used in classification

Feature	Odds Ratio e^{coeff}	Types Addressed
Page Based		
Quality Score II	0.141	All
Host Page Rank	0.152	All
URL Page Rank	0.283	All
Page Rank Presence	0.585	All
Quality Score I	1.045	All
Page in Index	2.396	All
Domain Based		
In *White Domain Table*	**0.022**	I, II, III
Type Based		
Is Type III	1.259	III
Is Type II	19.038	II
Is Type I	**597.815**	I
Word Based		
secure	1.395	IV
account	2.361	IV
webscr	2.710	IV
login	6.416	IV
ebayisapi	8.722	IV
signin	12.685	IV
banking	13.959	IV
confirm	15.777	IV

spelled. Table 2 shows the distribution of these features in our training set.

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

3.2 Training With Features

We chose the logistic regression classification technique. Logistic regression is a method for determining whether a set of independent variables has a unique predictive relationship to a binary dependent variable. Specifically, logistic regression is instrumental in evaluating the probability of the dependant variable occurring. In our setting, the dependent variable is whether a URL is a phishing URL, and the independent variables are the features we identified previously. For a given URL, and the values for each feature with respect to that URL, our logistic regression classifier estimates the probability of the URL being a phishing URL.

To train the logistic regression classifier the identified features were computed on the training data. Our labeled data consisted of 2508 URLs of which 1245 were phishing URLs and 1263 were benign URLs, as described in Section 3. Phishing URLs were placed under the positive (true) class and non-phishing ones were under the negative (false) class. We chose a random split of the labeled data — 66% of URLs 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 by Hosmer *et al.* [9] and used the data mining library in [23] to obtain the coefficients of our 18 features. To indicate the relative strength of each feature in identifying a Phishing URL we report the corresponding odds ratios, $e^{coefficient}$, in Table 3. Odds ratio is defined as the ratio of the odds

of an event occurring in the positive class (phishing) to the odds of it occurring in the negative class (not phishing). An odds ratio of 1 indicates that a feature is equally useful in identification of both classes. An odds ratio greater than 1 implies the corresponding feature is more useful in identifying the positive class.

Host obfuscation was found to be the most useful in detecting if a URL is phishing, as indicated by the very high odds ratio. Similarly presence in the *White Domain Table* is a strong indicator that a URL is not phishing. We point out that, even though these two features alone are strong discriminators, they are not sufficient in identifying all types of phishing URLs. Hence we also use the other features listed in our model.

We evaluated the trained model on the 34% test set split. We performed our evaluation over multiple runs with randomized partioning. This evaluation gave us an average accuracy of 97.31% with a True Positive Rate of 95.8 % and False Positive Rate of 1.2%.

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 [19] 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 the equation

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

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

We find that the area under the ROC curve for the positive class is 0.9923, thus confirming that our classifier has a high accuracy of phishing URL detection. We used the ROC 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%, True Positive Rate is 88% and the accuracy is 93.4%.

The resulting filter was used to analyze URLs obtained from Google Safe Browsing toolbar. Since our probability threshold is high, we have very few false positives. We detail our measurements in the following section.

4. ANALYSIS AND FINDINGS

We use the model described in Section 3 to measure frequency, type and origin of phishing attacks based on data obtained from the Google Safe Browsing toolbar. We collected several million URLs from August 20th to August 31 2006. The data consisted of two main components, unique URLs which are visited each day, and consecutive lookup requests to these URLs.

Average Phishing URLs per day.
We used our classifier, on this data set to determine the average number of phishing URLs which have been visited from Google's toolbar in a day. Figure 4 shows the distribution of phishing attacks on each day of our study. From our analysis we find that on average there are 777 URL phishing attacks in a day and an average of 5073 viewers to a phishing page. We observed that maximum phishing attacks occurred on Day 4 (August 23rd) and maximum number of viewers (hits) occurred on Day 3 (August 22nd). The graph
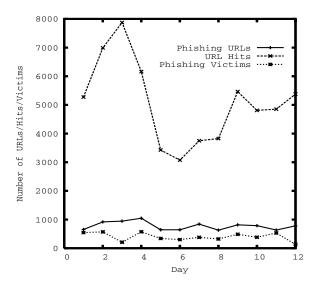

Figure 4: Phishing Attacks Observed

in Figure 4 also depicts the phishing victims which we will describe shortly.

We further analyzed the phishing URLs identified and categorize them based on the obfuscation category. Table 4 shows this distribution.

Table 4: Distribution of Obfuscation Types

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%

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. The considerable percentage of Type II and Type III attacks also confirms our previous argument that despite the high odds ratio, host obfuscation alone is not sufficient to identify phishing URLs.

Additionally, we examined the phishing URLs identified to determine the 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 are depicted in Table 5. We find that Type I and Type III obfuscation are prominent amongst the top 10 URLs.

Average Potential Phishing Victims per day.
We used the Logistic Regression Classifier on the lookup requests to each toolbar URL to determine how many users interact with a phishing page. Each lookup request consisted of 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

Table 5: Phishing URLs (top 10) active across all 12 days

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

on the same domain. A user that has any interaction at a site classified as phishing is regarded as a potential phishing victim. We point out that since we only look at URL based features we cannot determine the specific information entered by a potential victim. Extending our classifier towards content based analysis, such as that presented by Wu *et al.* [24], would improve the detection of phishing victims significantly.

Using our classifier, we found that on average there are 397 potential phishing victims in a day. Figure 4 indicates the number of users that fell for phishing on each day during the observation period. Day 4 (August 23rd) has the maximum number of potential phishing victims. Based on the number of users who view phishing pages in a day, we further can infer Potential Success Rate of a phisher as follows:

$$Potential\ Succes\ Rate = \frac{Number\ of\ Potential\ Victims}{Number\ of\ User\ Views} * 100$$

Our analysis indicates that on average, this success rate is 8.24%, implying that on average, 8.24% of the viewers to a phishing page are potential phishing victims.

Distribution of Phishing by Organization.

We further analyzed our results to determine which organizations were popular phishing targets. We used the *Target Organization List* to determine which organization was the target. Table 6 presents the number of unique phishing URLs by organization. The table also indicates the Potential Success Rate of a phisher for each organization. The results are averaged over the 12 days of data and the top 15 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. Ebay and Paypal are the top target organizations and phishers attract several potential victims to the corresponding phishing sites.

Geographical Distribution of Phishing.

We used the phishing URLs identified to determine the geographical distribution of phishing per 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 7.[2] The results indicate that a large percentage of phishing attacks are in the United States.

[2]The URLs were distributed across 59 countries. We have displayed the results for the top 9 countries.

Table 6: Distribution by Organization.

Organization	Unique Phishing URL	Potential Success Rate (%)
Ebay	231	14.8
Paypal	211	7.6
Fifth Third Bank	61	0
Unknown	60	8.2
Bank Of Scotland	32	8.6
Volksbank	29	0
Wells Fargo	29	0
Bank of America	28	2
Sparkasse	14	3
Private Banking	13	0
HSBC	7	0
Chase	5	3
Amazon	4	4
Banamex	4	0
Barclays	4	0

Table 7: Geographical Distribution

Country	Phishing URLs(%)
United States	70.3
Sao Tome and Principe	6
Belize	4.5
China	2.9
Germany	1.3
Taiwan	1.2
United Kingdom	1
Russian Federation	1
Romania	0.9

Discussion.

We point out that since we set a very high threshold probability for our classifier as discussed in Section 3.2, the analysis and results presented 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 and 12% (False Negative Rate) will be missed. 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 data sets. Our analysis confirms that phishing attacks are a very serious threat today.

Table 8: Anti-Phishing Tools

Tool	Primary Feature	Limitations
Google Safe Browsing	Uses a blacklist of phishing URLs to identify a phishing site	Might not recognize phishing sites not present in the blacklist
NetCraft Tool Bar	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.
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
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.

5. RELATED WORK

While there exists a plethora of different techniques in phishing detection, we find that existing research lacks a thorough measurement study like the one we present in this paper. In this section we briefly discuss some of the detection methods presented previously.

One of the popular methods of detection is using add-in toolbars for the browser. Chou *et al.* introduced one such tool SpoofGuard [2], that determines if a web page is legitimate based on a series of domain and URL based tests. Gabber *et al.* present a tool [10], that tries to protect a clients identity and password 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 [20] in a tool PwdHash.

Other anti-phishing tools include Google Safe Browsing [21], NetCraft tool bar [16], SpoofStick [4] and SiteAdvisor [14]. These are summarized in Table 8. Most of these tools bars rely only on black list information and might not correctly identify new phishing attacks. In comparison, our classifier uses several generic features and hence enhances the detection of phishing URLs.

Recently, Zhang *et al.* [25] conducted a thorough analysis on several anti-phishing tool bars to evaluate their effectiveness. Their evaluation suggests that SpoofGuard was very effective in detection of phishing sites, however it did have a very high false positive rate. This is probably due to the fact that SpoofGuard does not rely on any black list/white list information. Their results indicate that the freshness of the phishing URL affects the performance of several anti-phishing tools that rely solely on black list information.

Another approach towards phishing detection, is visual appearance of the web site. Dhamija *et al.* present a technique Dynamic Security Skins [6], which is an extended implementation of their work on Human Interactive Proofs [7]. This technique, uses a shared secret image that allows a remote server to prove its identity to a user, in a way that supports easy verification by humans and is hard to spoof by attackers. This protocol, however does not provide security for situations where the user login is from a public terminal.

In [12], 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 also 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 that generates a unique six digit numeric code every 60 seconds. Microsoft, on the other hand implemented SenderID [15], an email authentication protocol, to address the problem of domain spoofing.

Many researchers have also explored the question as to why users are susceptible to phishing attacks. Jakobsson introduced a new model, called a *phishing graph*, to visualize the flow of information in a phishing attack [13]. While this model is not in its essence a defensive technique, it is the first step towards developing an abstract model for visualizing phishing. A phishing graph enhances the ability to analyze and understand the course of a phishing attack.

More recently, Dhamija *et al.* [5] analyzed 200 phishing attacks from the Anti-Phishing Work Group database and identified several reasons, ranging from pure lack of computer system knowledge, to visual deception tricks used by adversaries, due to which users fall for phishing attacks. They further conducted a usability study with 22 participants and found that 23% of the participants failed to look at security indicators against phishing attacks and as a result 40% of the time they were susceptible to a phishing attack. Based on their analysis, the authors suggest that it is important re-think the design of security systems, particularly by taking usability issues into consideration.

6. CONCLUSION

In this paper we have identified several new and generic features for identifying phishing URLs. We use our features in a logistic regression classifier that achieves a very high accuracy. One of the major contributions of this work is a large scale measurement study conducted on Google Toolbar URLs. We use our classifier and analyze several million URLs. On average we found around 777 unique phishing

pages per day and on average 8.24% of the number users who view phishing pages are potential phishing victims. Our analysis validates the seriousness of phishing as a threat today. For our full findings, please see the technical report [8].

7. ACKNOWLEDGMENTS

8. 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.

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

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

[5] Rachna Dhamija, Doug Tygar, and Marti Hearst. Why phishing works. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 581–590. ACM Special Interest Group on Computer-Human Interaction, January 2006.

[6] Rachna Dhamija and J. D. Tygar. The battle against phishing: Dynamic Security Skins. In *SOUPS '05*, 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] Sujata Doshi, Niels Provos, Monica Chew, and Aviel D. Rubin. A Framework for Detection and Measurement of Phishing Attacks. Technical report, Johns Hopkins University, SPAR, December 2006. `http://www.cs.jhu.edu/~sdoshi/index_files/phish_measurement.pdf`.

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

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

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

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

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

[14] McAfee. Mcafee siteadvisor. `http://www.siteadvisor.com/`.

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

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

[17] 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.

[18] Niels Provos, Joe McClain, and Ke Wang. Search worms. In *WORM '06*, pages 1–8, New York, NY, USA, 2006. ACM Press.

[19] Foster J. Provost, Tom Fawcett, and Ron Kohavi. The Case against Accuracy Estimation for Comparing Induction Algorithms. In *ICML*, pages 445–453, 1998.

[20] 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.

[21] Fritz Schneider, Niels Provos, Raphael Moll, Monica Chew, and Brian Rakowski. Phishing Protection Design Documentation, 2006. `http://wiki.mozilla.org/Phishing_Protection:_Design_Documentation`.

[22] 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`.

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

[24] Min Wu, Robert C. Miller, and Greg Little. Web Wallet: preventing phishing attacks by revealing user intentions. In *SOUPS '06*, pages 102–113, New York, USA, 2006. ACM Press.

[25] Yue Zhang, Serge Egelman, Lorrie Faith Cranor, and Jason Hong. Phinding Phish: Evaluating Anti-Phishing Tools.