Phorecasting Phishing Attacks: A New Approach for Predicting the Appearance of Phishing Websites

Brad Wardman¹, Dan Clemens², Joseph Wolnski²
PayPal Inc. – San Jose, United States¹
Shadow Dragon – United States²
brad.wardman@gmail.com¹

ABSTRACT
Phishing is a multi-billion dollar business. The anti-phishing industry continues to pursue new strategies to thwart and investigate phishing attacks. Detection strategies include algorithms to detect new phishing websites using email-, URL-, and content-based approaches. Investigative strategies include clustering algorithms using drop email addresses, WHOIS registration information, and website content, as well as subpoenaing phishers’ mailboxes to determine the amount of information phished from victims. Until now, information predicting the implementation of new phishing websites has seldom been shared with the public. In this research we present a set of methods that demonstrate a relationship between phishers and defacers. Highlighting this relationship assists in building substantial defenses and law enforcement cases against this threat and shows that the proposed strategy can be used to predict when and where new phishing websites and related attacks will surface next.

KEYWORDS
Phishing, defacements, cyber-crime, attribution, intelligence

1. INTRODUCTION
The commonly employed social engineering attack known as phishing lures victims into providing sensitive information such as usernames and passwords, credit card numbers, and identity-related information, by imitating websites of popular and trustworthy organizations (e.g. financial institutions, social networking services, and mail providers). The information obtained by the attacker (referred to as the phisher) is used to take control of the customer’s account, steal the individual’s sensitive information, or is sold on the underground market [1]. Phishing campaigns account for economic, operational, and social damages to the online communities they target [2]. The exact loss due to these targeted attacks is unknown; yet, industry security consultants RSA reported that in 2013 nearly 450,000 phishing websites accounted for $5.9 billion in losses[3]. Unknown and unreported damages cannot be accounted for; however, more accurate measurements of known phishing attacks are needed to determine the resources required to mitigate the effects of phishing on an organization.

Organizations whose clients and employees are targeted by phishing generally use takedown strategies that involve contacting Internet Service Providers (ISPs), registrars, and website administrators to have the abusive content taken offline [4]. Such strategies are effective in getting websites offline within a few hours [5]; nevertheless, these takedowns do little to act as a deterrence to reduce future attacks [6]. More aggressive strategies including law enforcement investigations and prosecutions provide deterrence to criminals because of possible prison sentences or other legal penalties. There are, however, issues behind getting law enforcement cases started. For example, it is difficult to measure the magnitude and financial impact of a single actor’s attacks. Other challenges include jurisdiction issues and developing effective attribution methods. To make matters worse, if attackers are not deterred or limited in their actions, any form of justice comes at an exceedingly high cost for the victims and related organizations.

Researchers and investigators have attempted to build larger cases and demonstrate
attrition through a number of techniques including phishing kits, drop email addresses, domain name registration, clustering algorithms, traditional investigative tradecraft, open source intelligence, and subpoenas. Phishing kits are common toolkits that phishers employ to automatically build phishing websites using a prepackaged set of website files [7]. The toolkits allow attackers to set up a working phishing website by putting the files in the proper directory. Drop email addresses are a common mechanism for the delivery of the stolen data to the attackers. These email addresses are commonly included within one or more of the website files, and are sometimes hidden using obfuscation techniques [8]. Spam, malware, and phishing researchers can use the WHOIS information collected by registrars during the registration process to identify candidates responsible for the activity. Weaver and Collins demonstrated that WHOIS information such as IP and network of data is a viable method for clustering phishing websites [9]. The aforementioned kits and drop email addresses have not only been used by researchers and investigators in the past to perform attribution but Wardman et al. and Britt et al. have also used these pieces of evidence for further intelligence and analysis through clustering [10][11]. Lastly, it is widely believed that subpoenas are generally a good way to build case numbers for a dedicated law enforcement response [8]. When law enforcement gets a subpoena for a drop email mailbox, they are potentially able to identify the number of accounts and information compromised as well as the IP addresses of the criminals behind the attacks through an analysis of the emails that was reporting the phished information back to the attacker.

Table 1: Examples of different categories of defacement reasons.

<table>
<thead>
<tr>
<th>Reason</th>
<th>Description</th>
</tr>
</thead>
</table>
| Security Fame | "I'm Sorry Admin... Hack By Black.Hack3r From Indian Dark Coders Team"
| Patriotism | "Message: Fuck the White House For the Liberation of Iran. I like Khomeni. v: 40. e: enIGGER!
| Religious Beliefs | "Hi Admin, First: sorry about hacking second: I'm hack you only to send my massage not to the Mony"

When investigating phisher drop email addresses, it is common to find that the email addresses link to information being reported on defacement webpages by defacers. Defacers are a set of cybercriminals who gain unauthorized access to web servers and related systems by attacking design, implementation, logic and configuration vulnerabilities in order to post illegitimate web content on that server. Defacers claim to break into web servers for a number of reasons including security fame, for fun, revenge, goodwill, patriotism, the challenge, and religious beliefs. See example defacements in Table 1. Many defacers choose to document their defacements using online defacement databases.
The defacers use these databases to get credit for, or maintain a history of, their defacements. Examples of defacement databases include: Zone-H, H4ck Mirror, Turk-H.org, Aljyyosh.net, Add-Attack.com, ZoneHC, FlashBack.se, ATTRITION, and HackDB.com. These databases help defacers establish credibility in their communities. Figure 1 demonstrates the growth that has been observed in defacement databases that these researchers have information on. The graph demonstrates the rising trend of announcing the compromise and defacement of publicly hosted domains. Note that these numbers do not represent the total number of defacements reported to the defacement databases. There are almost an equal number of potentially defaced domains that have been reported to the defacement databases that either did not capture the defacement web page before it was removed or was not reviewed by members of the defacement database for whatever reason.

The research presented in this paper was assembled to demonstrate the relationship between phishers and defacers, using the domains and timestamps in which phishing and defacement websites were reported. For our first experiment, we hypothesized that there may be a relationship between a phishing website and defacement if a defacement webpage was hosted on the same domain within plus or minus one week of the phishing website. The results suggested that there could be greater value when extending the time period of the domains being hosted to demonstrate larger trends in their relationships.

The rest of this paper is as follows. Section 2 presents literature on anti-phishing strategies, phishers, and clustering algorithms used for building larger cases against phishers. Section 3 describes the research experiments in detail. Sections 4 and 5 present the results of the experiments and a discussion of those results. Section 6 presents future directions that may increase the scope of this research. Finally, Section 7 concludes our findings.

2. LITERATURE REVIEW

Much of the literature published on phishing discusses the employment of detection strategies to mitigate phishing attacks. These detection strategies can be categorized into email-, URL-, and content-based approaches; however, some researchers have used blended approaches to yield better results. Email-based approaches utilize characteristics and features about phishing emails to determine patterns across emails used to phish victims. Saberi et
al. classified the words in the body of the email to label them as ham, spam, and scam using statistical approaches such as Poisson filters, K Nearest Neighbor, and Naïve Bayes probabilistic theory [12]. These approaches were able to classify 94.4% of scams, while misclassifying 0.08% of legitimate or spam emails as scams. As the spammers adapt and become sophisticated, such text-based approaches are believed to be less effective against phishing [12][13]. The attackers utilizing spamming methods are generating emails that are so similar to legitimate emails that false positive rates can soar. In response, researchers have investigated other features from emails to improve their accuracy.

Chandrasekaran et al. derived 25 features from emails, used information theory concepts to rank the features, and finally classified the emails using a Support Vector Machine (SVM) on those features [13]. These researchers used a custom collection of 200 phishing and 200 non-phishing emails. The results demonstrated their technique could achieve 95% detection rate. Fette et al. developed an algorithm named PILFER to identify phishing emails using ten features of the email with a machine-learning classification approach using random forests [14]. The features used by PILFER include IP-based URLs, age of the domain, non-matching URLs between the hyperlink and anchor tag (i.e. the visual and actual links), HTML and JavaScript presence, number of links and domains, numbers of periods in the URL, and spam filter output. PILFER was tested on 7,810 emails, in which 860 were phishing and identified 96% of the phish with a 0.1% false positive rate.

Some researchers have identified common patterns in the URLs to detect phishing campaigns. These approaches are referred to as URL-based. URL-based approaches parse features from the URL such as the number of dots in the URL, length of the hostnames, number of special characters, presence of hexadecimal characters or IP addresses, and length of the URL [15][16][17]. Such features can be used to classify URLs as malicious or benign using machine learning algorithms. For instance, Blum et.al parsed URLs into N-gram tokens and utilized an online learning algorithm that assigns an individual confidence to each feature or N-gram in order to accurately identify phishing URLs [18]. Another avenue that researchers investigated was using website hosting information utilizing WHOIS and zoning files to show the relationship between the registration date of a domain and the date in which the phishing websites was created [19]. URL-based approaches are limited though as many phishing URLs do not contain any identifying characteristics or deviations from normal URLs such as a phishing website hosted on a compromised website in the root directory as index.php.

Content-based approaches analyze the source code, images, and other associated files of the phishing website as a means of identification. For instance, Wardman & Warner presented an algorithm that compares the sets of known phishing website content files with the file sets of potential phish [20]. A similarity score is used to determine if a potential website is related closely enough related to the confirmed phish. If the similarity score is high enough, the potential website is labeled as a phish. Another content-based approach labeled Syntactical Fingerprinting compares sets of syntactical elements within known phishing websites to sets of syntactical elements in potential websites [6]. A significant enough similarity score allows the potential website to be labeled as a phish.

Much of the anti-phishing methodologies presented above are reactive in nature but researchers have taken more of a proactive approach by aggregating information about phishing incidents [9][21][22]. Some researchers have attempted to use clustering algorithms on the content of the email
Table 2: Examples of the various components of a URL referred to in the paper

<table>
<thead>
<tr>
<th>Phishing URL</th>
<th>Domain</th>
<th>Hostname</th>
</tr>
</thead>
</table>

Table 3: Illustration of the phishing domains to defacement URL comparisons

<table>
<thead>
<tr>
<th>Phishing Domain Name</th>
<th>Defacement Hostname/URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>aforcedirect.co.uk</td>
<td><a href="http://www.aforcedirect.co.uk/images/m.txt">http://www.aforcedirect.co.uk/images/m.txt</a></td>
</tr>
<tr>
<td>aidandavisofficial.com</td>
<td><a href="http://www.aidandavisofficial.com">http://www.aidandavisofficial.com</a></td>
</tr>
<tr>
<td>cmha.co.uk</td>
<td><a href="http://www.cmha.co.uk">http://www.cmha.co.uk</a></td>
</tr>
<tr>
<td>cocothemes.com</td>
<td><a href="http://cocothemes.com">http://cocothemes.com</a></td>
</tr>
</tbody>
</table>

messages [21][22]. These methods are ineffective, though, due to the short life of features that are extracted from the headers and the duplication found in the intended mimicry of the content. Another form of phishing information aggregation is the application of clustering algorithms to net blocks reported in phishing scams [9]. This approach suggests that the phishing websites hosted on the same net blocks or autonomous system number (ASN) are from the same phisher and attempts to estimate the extent of phishing and its losses. Other research such as that conducted by Chen et al. try to determine the severity of phishing attacks by applying textual classification and data mining techniques on phishing alerts from Millersmiles and financial information provided by the phished organizations [15]. The researchers claim to classify the risk level of phishing attacks based on the margin of loss from the phished organization and categorized text from the phishing alerts [15].

The approaches described above are, to this researcher’s knowledge, mostly propose technologies for aggregating phishing activity. To fully address this problem, new methodologies have been developed in order to find patterns between phishing websites that show relationships and even the provenance of phishing websites. The rest of this paper describes the development of anti-phishing technologies that can be utilized for the detection of phishing websites, the aggregation of phishing data, and to build evidence to show strong relationships between the individual attacks.

In this research we present a set of methods that demonstrate the ability to map sets of phishing attacks to a defacer. Investigators and law enforcement can use the data captured through this research to help prioritize their investigations. In addition, the results also indicate that the proposed methodology could be used to predict where and when phishing and other malicious websites will be hosted.

3. METHODOLOGY

Three data sets were used in this research. The first data set, D1, was provided by PayPal and included all confirmed phishing hostnames and timestamps from 2010 to 2012. This data set included over 200,000 hostnames. The second data set, D2, was collected from the Anti-Phishing Working Group, or APWG, and included over 550,000 hostnames and timestamps reported to the APWG from 2010 to 2012. There are around 2,000 distinct organizations in data set D2. The final data set, D3, consists of information collected from three defacement databases; however, the majority of the results come from one database with specifically enhanced data utilized for this case study. The information collected on each of these defacement databases includes the hostname, domain, or URL of the defacement, the defacer, the defacement date, and the IP address that hosted the defacement. For clarification, the term “defacement string” will be used for the

146
hostname, domain, or URL collected by the defacement databases.

Experiment 1: PayPal Phish

The first experiment compared the domains and timestamps in data set D1 with the defacement strings and dates in D3 to determine the defacement strings that were hosted within plus or minus one week on the same domain as a PayPal phishing website. The comparison was done using a substring algorithm to see if the phished domain provided by PayPal was a substring of the defacement string (i.e. a hostname, domain, or URL). Table 2 contains an example of each field within a URL in order to assist in terminology used throughout the paper. When the D1 domain was a substring of the D3 defacement string and the D1 and D3 dates were within one week of each other, the data was collected in the format of defacement string, defacer, hostname, and the difference in days between the D1 domain and the D3 defacement string. Table 3 shows some example phishing domains that were substrings of the defacement string. For the rest of this document, a value of -7 days refers to a defacement page being posted 7 days before a phishing page on the same domain, while a value of 3 days refers to a defacement page being posted 3 days after a phishing page on the same domain. The goal of this experiment was to demonstrate the relationship between phishers and defacers targeting a single organization.

Experiment 2: APWG Phish

The next experiment was more robust in that it compared the domains and timestamps in data set D2 (the APWG set) with the defacement strings and dates in D3 to determine the defacement strings that were hosted within one year (i.e. plus or minus) on the same domain as a phishing website reported to the APWG. Similar to the first experiment, the comparison was performed using a substring algorithm to see if the APWG phishing domain was a substring of the defacement string. As in Experiment 1, when the domain was a substring of the defacement string and the dates were within the tested timeframe of each other, the same data was collected. However, the difference is that Experiment 2 used one year as the timeframe compared to the one week in Experiment 1. The goal in this
experiment was not only to use more data to show a larger impact by these defacers but also to demonstrate that some phishers target multiple organizations, giving law enforcement more motivation to open an investigation on the phisher.

4. RESULTS

Experiment 1: PayPal Phish

The results of Experiment 1, using data sets D1 (i.e. PayPal phishing domains) and D3 (i.e. defacement strings), provided initial insights into what was observed as larger patterns. From 2010 to 2012, there were 1,029 defacers that had a defacement string within one week of a phishing website. Table 4 in the Appendix contains the list of the top 10 defacers relating to defacers in D3 with the phishing websites in Experiment 1. The full “Alias” has been redacted for operational security, so the first and last character of the defacer has been included along with a label that is consistent between Tables 4 and 5, both in the Appendix. The column Membership contains two values. The value “Group” refers to the alias being a group handle, while “Solo” signifies that the alias seemingly belongs to a single hacker. The Affiliation column refers to the country or region in the world that the defacer promotes in their defacements. The overall counts per defacer were slightly lower than anticipated. This will be discussed more thoroughly in Section 5.

A significant pattern found in Experiment 1 stemmed from the mapping of a histogram of the data. Figure 2 is a histogram of the timescale relationship between the defacement and phishing websites. The blue bars cover the time span of -7 days to -1 day. These values represent the count in which a defacement page is advertised between 1 to 7 days before a phishing page appears on the same domain. The orange bar represents the count where the defacement and phishing pages occur on the same day. Finally, the green bar represents the time span (i.e. 1 to 7 days) when the phishing page is up before the defacement page. There is a noticeable pattern with a gradual increase from the value -7 days to 0 days, as well as the steep decrease from 0 days to 7 days. It is also important to note that the largest count in this histogram occurs on 0 days. In addition, the second highest count was -1 days.

The results of this initial histogram encouraged further research. Using the same sources, we adjusted the window of time from plus or minus one week to plus or minus three weeks and produced the histogram in Figure 3. It is interesting to note that the same trend can be seen in both figures despite a change in the time span, goes up before using the domain to host a PayPal phishing website.
The results of Experiment 1 raised questions on if these same patterns hold true for other phishing attacks. Therefore, Experiment 2 was organized to compare the defacement strings to other targeted organizations.

**Experiment 2: APWG Phish**

Experiment 2 consisted of two data sets: D2 (APWG phishing domains) and D3 (defacement strings). The comparison of D2 to D3 produced 61,507 overlapping values. Table 5 in the Appendix contains a list of the top 10 overlaps of a defacer related domain to a phishing domain hosted within one year of each other. The aliases iskorpi7x, Hmeit7, kaMtiEz, and 1923Turk are top defacers related to phishing in both Tables 4 and 5.

The pattern reflecting the days to phishing before or after defacement in Experiment 2 is very similar to the pattern captured in Experiment 1; however, it expands over a longer period of time. Figure 4 contains the count of phishing-to-defacement pairs per day from one year prior to one year after a phishing website. The orange line is representative of the count of pairs in which the phish is hosted on the same day as the defacement. The green lines represent the time buckets when the defacement was posted from 1 to 30 days after the phishing website. There is a clear pattern, like the patterns observed in Figures 2 and 3, indicating that the majority of the defacements occur days before, or on the day of, the phishing website. Observations of Figure 4 suggest that the defacement domain could be used as an indicator of suspiciousness for a greater period of time than one week. A deeper analysis of this figure and its implications will be described in Section 5.

5. DISCUSSION

**Potential Scenarios of Results**

The results of Experiments 1 and 2, illustrated in the associated histograms, demonstrate a pattern that suggests defacers want to capture their defacement before using the server for phishing or other functionalities. It is very important to note that in order for the defacer to get credit for their defacement, the defacer needs the defacement database to obtain a live
A screenshot taken by the defacer is not sufficient proof for credit. The defacer’s need for a live screenshot may help explain the pattern observed in the histograms. The scenarios presented below indicate the potential relationships between phishers and defacers that would exhibit this pattern.

The first scenario is that the phisher is the defacer. It is believed that this is the case in a vast majority of the websites that are -1 or 0 days. In addition, as investigators of phishing incidents, we have observed that phishers occasionally set up their phishing web page a couple of days before they start spamming. In fact, in this instance if website identification is sufficient and takedown services act quickly, some phishing websites are down before they are spammed.

The second scenario is that the phisher and defacer are part of the same organized group or crime unit. In this case, it is possible that no sole culprit is responsible for the whole attack. For instance, a defacer may claim credit for compromising a website; however, the phisher drop email address receiving the stolen information could be linked to another member of the group that the defacer is part of. Hence, suggesting that the phisher may be a specialized unit within the larger organization used to get credentials, but not necessarily do other tasks such as compromising the server or even monetizing the stolen information.

The third scenario involves multiple attackers compromising the originally defaced website after this information is shared within a public forum (e.g. Twitter, defacement databases, pastebin, etc…). In many compromises it is common to find more than one attacker leveraging the same or multiple vulnerabilities on the same host.

The final scenario is that some defacers sell credentials on underground markets. The phishers purchase the server credentials at a minimal cost of $5 to $10. This supports speculations that some phishers are non-technical criminals that leverage the work of others by using phishing kits, spamming tools, and social engineering to generate income.

**Signature Generation**

Figure 3 provides deep insight into the potential applicability of this research at a larger scale. It can be observed that the trending of the defacement being announced around -275 days before the phish has a greater count magnitude than the +30 day value (i.e. those values to the right of the green bars). This may indicate that normal behavior of the overlaps between defacements and phishing websites occur before -275 days and after +30 days. If this is the case, then one could create signatures to protect users from the domains posted in defacement databases with a Time to Live, or TTL, of 305 days. These signatures could be used by browser blacklists, spam filters, intrusion detection systems, and anti-virus vendors, to name a few, to protect corporate employees and daily users from malicious content being distributed from compromised web servers.

It is incredibly important to mention that these findings indicate that the level of abuse on a compromised web server seems to persist for an extended period of time after defacements. It is likely that the presence of a phishing website or other malicious content such as an exploit kit causes the webserver owner or hosting provider to take some action on the server that reduces the likelihood of staying compromised. We hypothesize that the pattern observed in Figure 3 will be present when comparing defacement domains to other attacks being distributed from compromised web servers as presented as future work in Section 6.

One limitation of this research which is described in Future Work includes missed matches due to the comparison of phishing domains to the defacement domains. Different
sources capture domain information differently. For instance, some defacement databases store the DNS name of the defacement page, while others store the hostname. Similar data inconsistencies occur in the APWG data set. An example scenario is when both the defacement database and the phishing data set contained different hostnames that were hosted on the same DNS name; hence, causing no overlap between the two hostnames even though they were hosted on the same DNS name. Future work to better capture the relationship is addressed in the following section.

The data supports that there is an opportunity to use the posted information by the defacers as an early warning system for malicious activities on those domains such as identifying new phishing attacks. Phished organizations will now be a little more aware of the possibility for phishing sites to be hosted on those domains. Anti-spam vendors can develop rules to further inspect URLs hosted on the defaced domains for a certain amount of time. Further research, as discussed in the Future Work section, may even demonstrate that a similar study could be done on malware-related domains. And if the same data holds true for defacers and malware-related domains, AV companies can develop signatures and IDS rules can be written to prevent malware being dropped by those domains.

6. FUTURE WORK

Current trends suggest that the use of compromised domains for phishing as well as other attacks will only continue to grow as the APWG reports that around 70% of all phishing domains are from hacked servers [23]. To help address this issue, this work could be expanded on in a variety of ways. The first is to obtain more data on defacements and defacers. A number of defacement and threat intelligence databases exist that could be used for further experimentation. We used three defacement databases and two phishing URL collections in this research; however, there are a number of other databases containing information on defacements, phishing URLs, and other types of malicious content.

Next, we believe that there will be an even higher correlation through the use of the IP addresses that the domains are hosted, which may reduce the hit rate on matches discussed as a limitation above. Hundreds, even thousands, of domains can be hosted on a single server; therefore, a similar study using IP lookups and comparisons may yield better results. An example of such an experiment could be to perform a DNS lookup on the defacement domain to obtain the IP address of the hosting machine, run a reverse DNS on that IP address to get the other domains on the IP, and lastly, perform a DNS name comparison of those domains with the phishing domains. This extra reverse DNS step would account for the possibility that some entities do not save the IP address where the phishing websites are hosted. The reason that we did not use this method in this research was that the data was not being captured in the historical feeds we used. Therefore, we have made changes to current collection strategies to collect such data for future experiments.

Another direction this research could take is to perform this comparison over known domains used for malware-related activities such as redirects and hosting exploit kits. We hypothesize that this would result in a similar pattern as the one found in the histogram presented in this research, which would further validate the theory that the defacers’ need to take credit for their conquests can be used as a mechanism for predicting malicious acts.

Lastly, we would like to see how the clusters of websites based on defacers in this research map to other phishing clustering algorithms.
such as those presented in the 2010 publication by Wardman et al. [24] The combination of the sets of algorithms may help to differentiate between phishers or groups directly related to phishers and defacers who sell their servers on underground markets.

7. CONCLUSIONS
This research demonstrates a new methodology using defacement artifacts to prioritize investigations on phishers. The data indicates that there is a substantial relationship between the domains defaced by defacers and those domains used to host phishing websites. Patterns in the results also suggest that defacement data can be used as a proactive measure for predicting when and where future malicious activity such as phishing will occur.

REFERENCES


### Table 4: Top 10 Defacers Relating to PayPal Phish

<table>
<thead>
<tr>
<th>Count</th>
<th>Alias : Label</th>
<th>Membership</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>98</td>
<td>1---k : Defacer A</td>
<td>Group</td>
<td>Turkey</td>
</tr>
<tr>
<td>91</td>
<td>i---x : Defacer B</td>
<td>Solo</td>
<td>Turkey</td>
</tr>
<tr>
<td>88</td>
<td>H---7 : Defacer C</td>
<td>Solo</td>
<td>Indonesia</td>
</tr>
<tr>
<td>49</td>
<td>G---y : Defacer D</td>
<td>Solo</td>
<td>Indonesia</td>
</tr>
<tr>
<td>47</td>
<td>a---r : Defacer E</td>
<td>Solo</td>
<td>Israel</td>
</tr>
<tr>
<td>45</td>
<td>z---a : Defacer F</td>
<td>Group</td>
<td>Gaza</td>
</tr>
<tr>
<td>44</td>
<td>k---z : Defacer G</td>
<td>Solo</td>
<td>Indonesia</td>
</tr>
<tr>
<td>43</td>
<td>S---D : Defacer H</td>
<td>Solo</td>
<td>Iraqi Kurdistan</td>
</tr>
<tr>
<td>27</td>
<td>M---N : Defacer I</td>
<td>Solo</td>
<td>Iran</td>
</tr>
<tr>
<td>27</td>
<td>D---D : Defacer J</td>
<td>Solo</td>
<td>N/A</td>
</tr>
</tbody>
</table>

### Table 5: Top 10 Defacers Relating to APWG Phish

<table>
<thead>
<tr>
<th>Count</th>
<th>Alias : Label</th>
<th>Membership</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>5751</td>
<td>i---x : Defacer B</td>
<td>Solo</td>
<td>Turkey</td>
</tr>
<tr>
<td>4423</td>
<td>H---7 : Defacer C</td>
<td>Solo</td>
<td>Indonesia</td>
</tr>
<tr>
<td>2165</td>
<td>l---k : Defacer A</td>
<td>Group</td>
<td>Turkey</td>
</tr>
<tr>
<td>1839</td>
<td>k---z : Defacer G</td>
<td>Solo</td>
<td>Indonesia</td>
</tr>
<tr>
<td>1467</td>
<td>G---1 : Defacer K</td>
<td>Solo</td>
<td>Turkey</td>
</tr>
<tr>
<td>1146</td>
<td>m---r : Defacer L</td>
<td>Solo</td>
<td>Turkey</td>
</tr>
<tr>
<td>1072</td>
<td>c---4 : Defacer M</td>
<td>Solo</td>
<td>Kurdish</td>
</tr>
<tr>
<td>900</td>
<td>A---m : Defacer N</td>
<td>Group</td>
<td>Iran</td>
</tr>
<tr>
<td>752</td>
<td>T---d : Defacer O</td>
<td>Solo</td>
<td>Turkey</td>
</tr>
<tr>
<td>749</td>
<td>k---L : Defacer P</td>
<td>Solo</td>
<td>Saudi Arabia</td>
</tr>
</tbody>
</table>