SUMMARY REPORT
RESEARCH PAPER – RELATIVE INCIDENCE OF PHISHING AMONG DV, OV, AND EV ENCRYPTED WEBSITES

Chris Bailey and Kirk Hall, Entrust Datacard
Melih Abdulhayoğlu and Fatih Orhan, Comodo

April 16, 2018

This Summary Report updates the earlier statistics from our December 8, 2017 paper, distributed at the Ensuring Web PKI Integrity 2 meeting in New York in December 2017. The updated data and statistics reinforce our findings in that original paper.

1. Our Methodology

Entrust Datacard and Comodo have worked together to gather data from publicly trusted and valid SSL/TLS enabled phishing sites. The data in this update was collected between **March 16, 2018 and April 15, 2018** totaling **171,382 suspected phishing records**.

We double-checked our phishing results against Google's Safe Browsing API and marked the site as a phishing site if the same URL appeared on both our list and the GSB list. We got verification results back for 69,486 records from Google Safe Browsing. Of the 69,486 matching records, only 6,020 records showed a valid SSL/TLS connection UI. The rest of this study analyzes those 6,020 encrypted phishing sites.

Total suspected phishing records collected *	171,382	
And matched against Google Safe Browsing list	69,486	40.5% of suspected phishing records matched with Google Safe Browsing (GSB)
And has a valid publicly trusted SSL/TLS certificate	6,020	8.7% of the matched GSB phishing records had a valid SSL/TLS cert

Next, we determined the type of certificates (DV, OV, or EV) used on each of the 6,020 encrypted phishing sites we found.

2. Our Data

Here are the results of our initial data which show the 6,020 phishing sites we found with valid SSL/TLS certificates sorted by type, which we are first reporting in this study:

Cert Type	Number of Phishing Certs	Percent
EV	3	0.05%
OV	8	0.13%
DV	6,009	99.82%
Total	6,020	100.0%

Preliminary conclusion: Using these results, it seems our hypothesis that EV sites are safer than OV and DV sites is true. Also, our hypothesis that OV sites are safer than DV sites also appears to be true.

Further Analysis:

Reviewing the Data: DV websites make up the great majority of encrypted websites on the internet. How does this raw breakdown of encrypted phishing sites (EV, OV, and DV) look when compared to the entire population of certificates on the internet?

	Our Sample		The Internet	
Certificate Type	**Phishing Sites in Sample**	**Percent of Total Phishing Sites in Our Sample**	**Total Internet Certificate Population[1]**	**Percent of Total Cert Population**
EV	3	0.05%	195,409	0.7%
OV	8	0.13%	1,480,294	5.0%
DV	6,009	99.82%	27,822,384	94.3%
Total	6,020	100.00%	29,498,087	100.00%

When we compare the certificate type for the encrypted phishing sites in our *sample* against the entire certificate population on the *internet*, the results are revealing:

Certificate Type	**(1) Representation of Certificate Type in *Total Certificate Population***	**(2) Representation of Certificate Type Among *Phishing* Sites in Our *Sample***
EV	0.70%	0.05%
OV	5.00%	0.13%
DV	94.30%	99.82%

If phishing sites were equally distributed among each of the three certificate types, we would expect the numbers in Columns (1) and (2) to be the *same* – but they're not.

Instead, this table shows that the percentage of EV phishing sites in our sample (Column 2) was only 0.05% of all phishing sites in the sample, versus what we might have expected: 0.70% (Column 1) based on the representation of EV certs among the total cert population on the internet. This means the number of EV phishing sites in our sample is only **7%** of what we might have expected based on the number of EV certificates in the population generally (0.05%/0.70%), so by this measure EV sites are 93% safer for users than encrypted sites generally.

Likewise, the percentage of OV phishing sites in our sample (Column 2) was actually 0.13% versus the OV population of 5.00% (Column 1) among the total cert population on the internet. This means the number of OV phishing sites in our sample is only **2.6%** of what we might have expected based on the number of OV certificates in the population generally (0.13%/5.00%), so by this measure OV sites are 97% safer for users than encrypted sites generally

In contrast, the percentage of DV phishing sites in our sample (Column 1) was actually 99.82% versus the DV population of 94.3% (Column 1) based on the representation of DV certs among the total cert

[1] Based on Netcraft valid certificate population by certificate type as of April 2018.

population on the internet – meaning that by this measure, the number of DV phishing sites is 106% of the number we would have expected (99.82%/94.30%). Of course, the DV numbers are skewed because they now represent the overwhelming number of certs on the internet.

Making a rough comparison from this data set, a DV website is roughly 15 times more likely to be a phishing site than an EV site (106%/7%), and a DV website is roughly 41 times more likely to be a phishing site than an OV site is (106%/2.6%). We think we can greatly improve these numbers for OV and EV certificates as discussed in Section 3, thereby making these sites even safer. (We believe the apparent security advantage for OV sites in this study is due to the small number of OV and EV phishing sites found, which can skew the percentages.)

Our preliminary analysis of the OV and EV phishing sites in this sample indicates the same reasons why phishing content was present as we discussed in our December study. It comes down to three sources:

(1) Several phishing sites have shared content with OV certs that allow users to post phishing content,

(2) Several phishing sites have shared certificates where multiple SANs are listed and one or more of the independent sites included as SANS was flagged for phishing. These certs are often issued to web hosting companies, and

(3) Several phishing sites had been *compromised*. A phisher had taken over part of the site's directory and posted phishing content on the site where owner would not notice - the phishing URLs appeared to be "orphaned" URLs that are not reachable by scanning the site. Site compromise is generally the only reason we find phishing content on EV sites – the website owner is typically unaware of the content.

All of these phishing content sources can be addressed by the CAs who issued the certificates once the CAs are aware of the problem and contact the customer to redress. We plan to start doing that.

3. Our Conclusion. Based on these comparisons, OV and EV sites today are much safer for users than DV sites.

They can be made even safer by the voluntary, joint initiative that the largest commercial CAs will be discussing at the Ensuring Web PKI Integrity 3 conference in Salt Lake City on Thursday, April 26, 2018, and will be announcing at the next CA/Browser Forum Face-to-Face Meeting in London on June 6-7, 2018 – hence the name of the initiative, the "**London Protocol**." See attached draft London Protocol for more details.

We believe the results of this updated study also support our conclusion in the December 2017 version of this paper that browsers should recognize the greater users security from identity certificates (OV and EV), should modify and coordinate their current browser UI security indicators to provide users with this identity information, and should engage with CAs and others in renewed user training to recognize the updated UIs when making website security decisions.