Certified Phishing

Taking a Look at Public Key Certificates of Phishing Websites

Monday, August 12

Vincent Drury, Ulrike Meyer
RWTH Aachen University
Research Group IT-Security
Why look at Certificates?

Increasing number of websites with HTTPS

Why look at Certificates?

Trend also observable in phishing websites

Phishing: Terms and Process

Attacker

Victim

Target
Phishing: Terms and Process

1. Copies Website

Attacker
Victim
Target
Phishing: Terms and Process

1. Copies Website
2. Sends Link
3. Enters Credentials

Attacker

Victim

Target
Phishing: Terms and Process

1. Copies Website
2. Sends Link
3. Enters Credentials

Attacker
Victim
Target

Copies Website
Sends Link
Enters Credentials
Certificates and CAs

Certificate Authority (CA)
Certificates and CAs

Domain Validation (DV)

<table>
<thead>
<tr>
<th>Issued To</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Name (CN)</td>
</tr>
<tr>
<td>Organization (O)</td>
</tr>
<tr>
<td>Organizational Unit (OU)</td>
</tr>
</tbody>
</table>

Certificate Authority (CA)
Certificates and CAs

Domain Validation (DV)

<table>
<thead>
<tr>
<th>Issued To</th>
<th><a href="http://www.letsencrypt.org">www.letsencrypt.org</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Name (CN)</td>
<td><a href="http://www.letsencrypt.org">www.letsencrypt.org</a></td>
</tr>
<tr>
<td>Organization (O)</td>
<td>&lt;Not Part Of Certificate&gt;</td>
</tr>
<tr>
<td>Organizational Unit (OU)</td>
<td>&lt;Not Part Of Certificate&gt;</td>
</tr>
<tr>
<td></td>
<td><a href="https://letsencrypt.org">https://letsencrypt.org</a></td>
</tr>
</tbody>
</table>

Organization Validation (OV)

<table>
<thead>
<tr>
<th>Issued To</th>
<th><a href="http://www.google.com">www.google.com</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Name (CN)</td>
<td><a href="http://www.google.com">www.google.com</a></td>
</tr>
<tr>
<td>Organization (O)</td>
<td>Google LLC</td>
</tr>
<tr>
<td>Organizational Unit (OU)</td>
<td>&lt;Not Part Of Certificate&gt;</td>
</tr>
<tr>
<td></td>
<td><a href="https://www.google.com">https://www.google.com</a></td>
</tr>
</tbody>
</table>
Certificates and CAs

Domain Validation (DV)

- **Issued To**
  - Common Name (CN): www.letsencrypt.org
  - Organization (O): <Not Part Of Certificate>
  - Organizational Unit (OU): <Not Part Of Certificate>

Organization Validation (OV)

- **Issued To**
  - Common Name (CN): www.google.com
  - Organization (O): Google LLC
  - Organizational Unit (OU): <Not Part Of Certificate>

Extended Validation (EV)

- **Issued To**
  - Common Name (CN): www.paypal.com
  - Organization (O): PayPal, Inc.
  - Organizational Unit (OU): CDN Support
Research Questions

Are there **general differences** between certificates of phishing and benign websites?
Research Questions

Are there **general differences** between certificates of phishing and benign websites?

Are there differences in comparison to a **specific target**?
Collection Process Overview

- 50,000 benign URLs from Alexa
- 31,264 phishing URLs from PhishTank (53 days)

Certificates:
- Benign: Collected 43,018
  - Duplicates: -698
  - Invalid: -2,842
  - Final: 39,478
- Phishing: 9,479

⇒ HTTPS does not mean safe
Collection Process Overview

- 50,000 benign URLs from Alexa
- 31,264 phishing URLs from PhishTank (53 days)
Collection Process Overview

- 50,000 benign URLs from Alexa
- 31,264 phishing URLs from PhishTank (53 days)

<table>
<thead>
<tr>
<th>Certificates</th>
<th>Benign</th>
<th>Phishing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collected</td>
<td>43,018</td>
<td>25,777</td>
</tr>
<tr>
<td>Duplicates</td>
<td>-698</td>
<td>-11,712</td>
</tr>
<tr>
<td>Invalid</td>
<td>-2,842</td>
<td>-4,586</td>
</tr>
<tr>
<td>Final</td>
<td>39,478</td>
<td>9,479</td>
</tr>
</tbody>
</table>

⇒ HTTPS does not mean safe
Collection Process Overview

- 50,000 benign URLs from Alexa
- 31,264 phishing URLs from PhishTank (53 days)

<table>
<thead>
<tr>
<th>Certificates</th>
<th>Benign</th>
<th>Phishing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collected</td>
<td>43,018</td>
<td>25777</td>
</tr>
<tr>
<td>Duplicates</td>
<td>-698</td>
<td>-11,712</td>
</tr>
<tr>
<td>Invalid</td>
<td>-2,842</td>
<td>-4,586</td>
</tr>
<tr>
<td>Final</td>
<td>39,478</td>
<td>9,479</td>
</tr>
</tbody>
</table>

⇒ HTTPS does not mean safe
General Differences

Distribution of Validation Types:

- **Benign**
- **Phishing**

The diagram shows the distribution of validation types for both benign and phishing certificates. The bars are color-coded: green for DV, orange for OV, and blue for EV. The proportions of each type are indicated by the segments within the bars.
General Differences

Distribution of (5 most common phishing) Issuers:

- Let’s Encrypt
- cPanel
- RapidSSL
- COMODO RSA
- COMODO ECC

⇒ Not generally possible to determine if phishing or benign from certificate alone
General Differences

Distribution of (5 most common phishing) Issuers:

⇒ Not generally possible to determine if phishing or benign from certificate alone
## Specific Target

<table>
<thead>
<tr>
<th>Rank</th>
<th>Target</th>
<th>Certificates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PayPal</td>
<td>1169</td>
</tr>
<tr>
<td>Rank</td>
<td>Target</td>
<td>Certificates</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>--------------</td>
</tr>
<tr>
<td>1</td>
<td>PayPal</td>
<td>1169</td>
</tr>
</tbody>
</table>

*: Hosted on the target's own infrastructure.

⇒ No evidence of active replication of certificate information, but abuse of target infrastructure possible.
## Specific Target

<table>
<thead>
<tr>
<th>Rank</th>
<th>Target</th>
<th>Certificates</th>
<th>Organization</th>
<th>URL DN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PayPal</td>
<td>1169</td>
<td>0</td>
<td>84</td>
</tr>
</tbody>
</table>

*: Hosted on the target's own infrastructure.

⇒ No evidence of active replication of certificate information, but abuse of target infrastructure possible.
### Specific Target

<table>
<thead>
<tr>
<th>Rank</th>
<th>Target</th>
<th>Certificates</th>
<th>Organization</th>
<th>URL DN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PayPal</td>
<td>1169</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>3</td>
<td>Microsoft</td>
<td>297</td>
<td>47*</td>
<td>10</td>
</tr>
<tr>
<td>12</td>
<td>Dropbox</td>
<td>37</td>
<td>1*</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>Google</td>
<td>33</td>
<td>1*</td>
<td>1</td>
</tr>
</tbody>
</table>

*: Hosted on the target’s own infrastructure.
## Specific Target

<table>
<thead>
<tr>
<th>Rank</th>
<th>Target</th>
<th>Certificates</th>
<th>Organization</th>
<th>URL DN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PayPal</td>
<td>1169</td>
<td>0</td>
<td>84</td>
</tr>
<tr>
<td>3</td>
<td>Microsoft</td>
<td>297</td>
<td>47*</td>
<td>10</td>
</tr>
<tr>
<td>12</td>
<td>Dropbox</td>
<td>37</td>
<td>1*</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>Google</td>
<td>33</td>
<td>1*</td>
<td>1</td>
</tr>
</tbody>
</table>

*: Hosted on the target’s own infrastructure.

⇒ No evidence of active replication of certificate information, but abuse of target infrastructure possible
Suitability of Certificates for Phishing Detection

● Discriminative power
  ● Benign websites without EV/OV certificate
Suitability of Certificates for Phishing Detection

- Discriminative power
  - Benign websites without EV/OV certificate

- Robustness
  - Trust in CAs?
Suitability of Certificates for Phishing Detection

- Discriminative power
  - Benign websites without EV/OV certificate

- Robustness
  - Trust in CAs?
  - Certificate Transparency
Suitability of Certificates for Phishing Detection

- Discriminative power
  - Benign websites without EV/OV certificate

- Robustness
  - Trust in CAs?
  - Certificate Transparency

- Other potential problems
  - TLS interception
Conclusion

- In general **hard to differentiate** certificates of benign and phishing websites
- Currently **no evidence that attackers actively replicate** the content of target certificates
  - But: hosting on target infrastructure sometimes possible
- Certificates as possible resource for future research:
  - Automated detection?
  - User Education?
Conclusion

- In general **hard to differentiate** certificates of benign and phishing websites

- Currently **no evidence that attackers actively replicate** the content of target certificates
  - But: hosting on target infrastructure sometimes possible

- Certificates as possible resource for future research:
  - Automated detection?
  - User Education?

Thank you for your attention!