PhishGuru – Embedded Training

- Users get sent periodic training emails that look like phishing attacks
- If a user clicks a “phishing” link, they are shown succinct and engaging information on protecting themselves from phishing

User study results

- Sending training materials through normal email is ineffective but users are motivated to learn after falling for fake phishing attack
- Users retain and transfer knowledge better when trained after falling for fake phishing attack versus getting training materials in normal email
- Real world studies confirm our lab study results

Anti-Phishing Phil – A Training Game

- A web-based interactive game to teach people how to avoid phish by paying attention to URLs
- Players move Phil around screen to examine bait and identify URLs as phishing or legitimate

User study results

- Lab study showed people made better decisions after playing our game versus reading online training materials
- 4,517 people participated in online study
- Online study demonstrated that playing our game helped people learn to make better decisions
- People retained the knowledge after one week

Online study results: false negative before, immediate, and after one week delay. Novices that were initially poor at identifying phish are much better at identifying phish after playing our game.

This work is funded by the US National Science Foundation under Grant No. 0524189, Fundação para a Ciência e Tecnologia (FCT) Portugal under a grant from the Information and Communications Technology Institute (ICTI) at CMU, and ARO/CyLab. Principal Investigators: Lorrie Cranor, Alessandro Acquisti, Julie Downs, Jason Hong, and Norman Sadeh
Soon to be commercially available

For more information: http://cups.cs.cmu.edu/trust
or contact Lorrie Cranor <lorrie@cmu.edu>

User education is important to prepare users for evolving attacks. However, automated detection can significantly reduce the burden on end users.

## PILFER – Powerful Email Filtering

**Rationale:** Spam filters let a large number of phishing emails slip through

**Solution:** Use advanced machine learning in combination with features specifically targeting phishing emails

**Implementation:** Can work standalone or in combination with spam filter – e.g. available as Spam Assassin (SA) plugin

**Evaluation:** 90,000 emails

<table>
<thead>
<tr>
<th></th>
<th>False positives</th>
<th>False negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>PILFER</td>
<td>0.13%</td>
<td>4.79%</td>
</tr>
<tr>
<td>SA</td>
<td>3.19%</td>
<td>7.68%</td>
</tr>
</tbody>
</table>

- Catches more phish than Spam Assassin
- 25 x fewer false positives
- 20 x faster

## CANTINA – Powerful Web Filtering

**Rationale:** Blacklists are slow to update and only contain large-scale phishing attacks

**Solution:**
- Create fingerprint for a suspicious web page and use search engines and machine learning to identify
- No human intervention needed
- Protects against spear-phishing too

**Implementation:** Makes use of search engines and locally running machine learning algorithms

**Evaluation:**
- 90% true positive rate in correctly identifying phish
- 5% false positive rate in mis-identifying legitimate sites as phish

## Selected Publications


