The Highly Insidious Extreme Phishing Attacks

Rui Zhao*, Samantha John†, Stacy Karas†, Cara Bussell†, Jennifer Roberts†, Daniel Six†, Brandon Gavett†, and Chuan Yue*

* Colorado School of Mines, † University of Colorado Colorado Springs

Introduction
- Phishing: uses spoofed websites to steal users’ passwords and online identities.
- Defense:
  - Blacklist-based
  - Heuristics-based
  - Whitelist-based
- Phishing reporting and verification services: APWG & PhishTank
- Phishing attacks have also been quickly evolving to evade the detection and defense.

Introduction – cont.
- First-layer context: a spoofed email or instant message
  - To lure users to the phishing websites
- The success is limited by two constraints
  - If phishing emails or instant messages are suspicious
    - Users would not click on phishing URLs
  - If phishing emails are captured by spam filters
    - Cannot even reach users in the first place

Introduction – cont.
- Second-layer context: look and feel similar to a targeted legitimate website
  - To lure users to submit their login credentials
- The success is limited by two constraints
  - If phishing websites trigger warnings if they are detected by browsers
  - If the look and feel of the undetected phishing websites are suspicious

Our Goal
- We explore the feasibility of the extreme of phishing attacks!
  - that have the almost identical look and feel as those of the targeted legitimate websites
- We evaluate the effectiveness of such phishing attacks by performing a user study

Metrics for Look and Feel
- We focus on the second-layer context

Web Single Sign-On (SSO)?
- Sign in multiple relying party (RP) websites using one single identity provider (IDP) account.
- Users are relieved from the huge burden of registering many online accounts and remembering many passwords.

Link Substitution
- Our toolkit needs to ensure that all the links on each phising webpage will be modified to point to the phishing website.
  - To keep existing visitors on a phishing website.
  - To maximize the chances of collecting their login credentials.
- Static Link Substitution:
  - Legitimate domain → phishing domain & HTTPS → HTTP & customisable rules for special links (in ‘head’ and ‘script’)
- Dynamic Link Substitution:
  - Injects JavaScript to intercept the dynamic link generation and modification events & legitimate domain → phishing domain & HTTPS → HTTP

Web SSO Login Window Generation
- More profitable and insidious because
  - The value of RP accounts is highly concentrated
  - The attack surface area is highly enlarged
  - The difficulty of phishing detection is highly increased

We achieve
- The automatic and dynamic construction
- The automatic inclusion of Web SSO phishing login windows

A Measurement of Existing Phishing Websites
- In 2015, measured and inspected 471 live phishing websites reported on PhishTank:
  - 85% do not contain any link
  - 22% contain invalid links
  - 17.6% contain links to the targeted legitimate websites
  - 24.4% contain links to other websites
- The majority of them, 449 (95%) of 471 – Simple phishing
- Only a handful of them – Advanced Phishing
  - 2 tokens, 5 tokens, and 11 other websites mostly similar
- 75% phishing websites support low-quality Web SSO phishing websites

Discussion - Defenses
- Heuristics-based solution
  - That only uses features extracted from visual or non-visual elements on webpages may fail
  - That uses features extracted from URLs may become either inaccurate or incorrect on the detection of Web SSO extreme phishing attacks
- The best web pages for these SSO phishing attacks on legitimate websites
- Blacklist-based solution
  - Indirectly affected if the construction of their blacklist relies on heuristics-based techniques or anti-phishing communities
- Whitelist-based solution - more robust

Discussion - Suggestions
- We suggest that researchers should seriously consider extreme phishing in their heuristics-based phishing detection solutions.
- We suggest that Web users should be trained to
  - Be aware of extreme phishing
  - Pay more attention to the domain name of a URL
  - Differentiate the spoofed Web SSO login “windows” from real ones

Conclusion
- Explored the extreme phishing attacks and investigated the techniques for constructing them
- Designed and implemented a concrete toolkit
  - Traditional phishing and Web SSO phishing
  - Automatically construct unlimited levels of phishing webpages based on user interactions
- Designed and performed a user study with 94 participants
  - Demonstrated that extreme phishing attacks are indeed highly effective
- Discussed the impacts of extreme phishing on existing phishing defense mechanisms
- Provided suggestions
  - This project is supported in part by the NSF grant CNS-1562149.