Sunrise to Sunset: Analyzing the End-to-end Life Cycle and Effectiveness of Phishing Attacks at Scale

Adam Oest, Penghui Zhang, Brad Wardman, Eric Nunes, Jakub Burgis, Ali Zand, Kurt Thomas, Adam Doupé, Gail-Joon Ahn

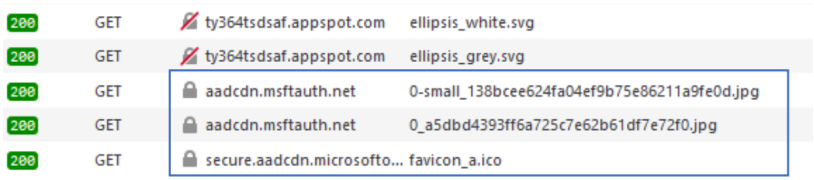
USENIX Security 2020

Presented by Nikita Borisov, January 2026

Motivation

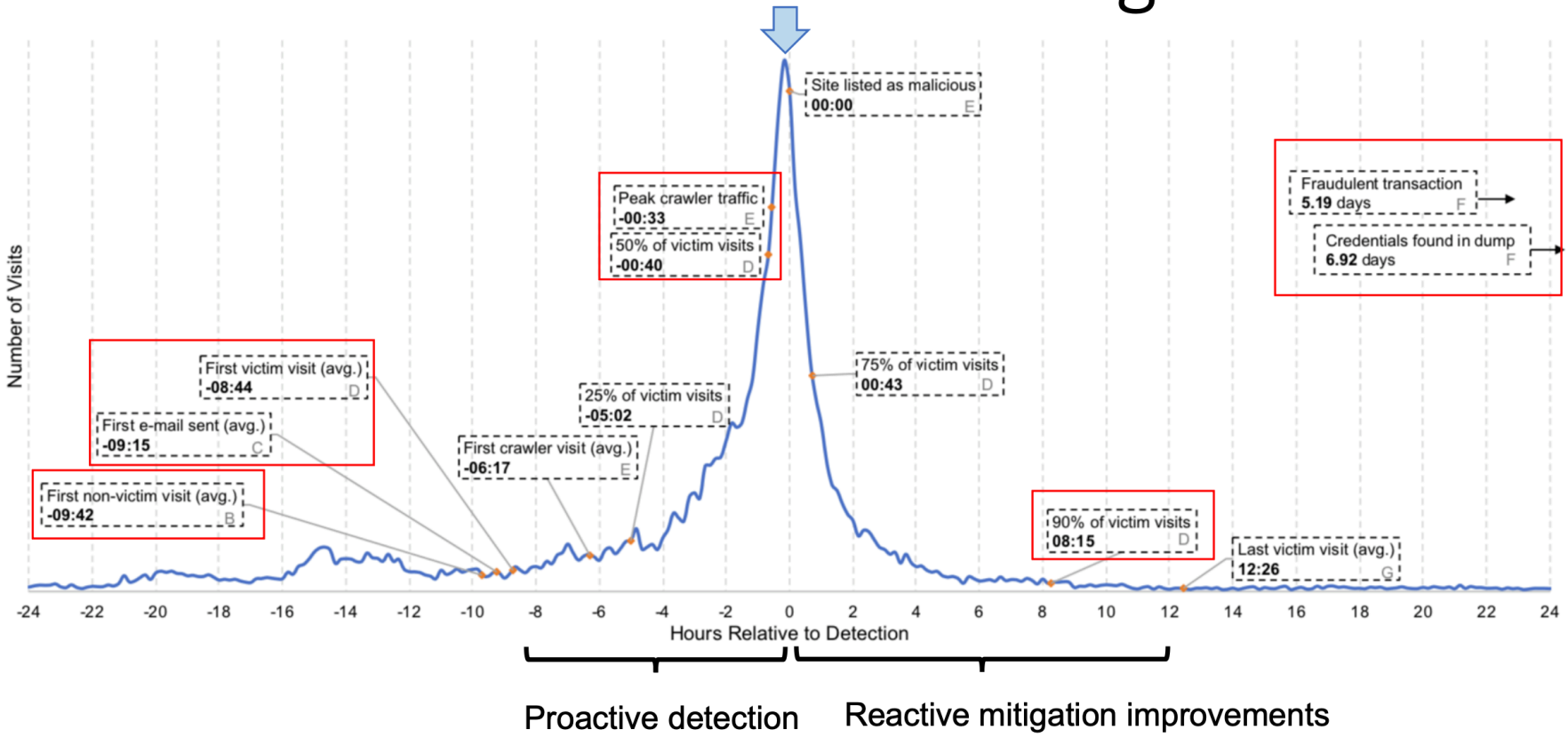
- Phishing has low click-through rates (5–8%) and low conversion rates (9%)
- All browsers incorporate anti-phishing blacklists
- *And yet phishing persists*
  - Cloaking defenses & redirections help avoid email-based defenses
  - Moving target helps stay ahead of blacklists
  - Scale

Approach



- Detect phishing through embedded resource loading
  - Phishing page embeds a link to real page
  - Browsers leave `Referer` *[sic]* headers pointing to the phishing site
- Log visits to page, and track campaign progress
  - Correlate with browser-based defenses
- Correlate with
  - Fraud data: does this lead to account abuse?
  - Email data: timing of spam with phishing links, reporting

End-to-end Timeline of Phishing



Data Sets Coming Together

- URL / referer data set (PayPal)
- Fraud detection for end-to-end (PayPal)
- Email detection (Google)
- Safe browsing data set (Google)

Interesting Results

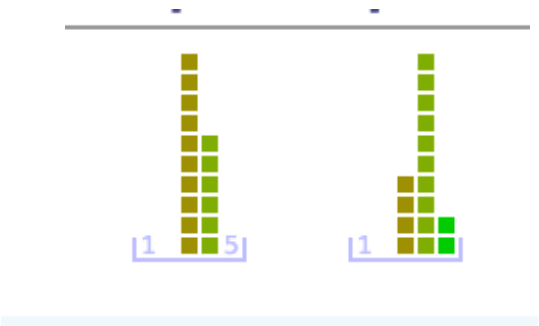
- Attacks are short-lived but *campaigns* last longer
  - Not enough domain-level takedowns?
- 62.73% victims lured before detection (within about 9 hours)
  - How can we detect faster?
- 7.42% of visitors lose their credentials
  - Phishing still works!
- Campaigns follow a Pareto distribution
  - What makes the top attackers more successful? Scale? Skill?

Safe Browsing Effectiveness

Strengths and weaknesses

- Large scale, e2e data set
- Novel methodology
- Interesting findings
- Clear, well-written

- Visibility issues
- Attacker adaptation
- Generalizability of data set beyond PayPal
- Causal analysis



Visibility

- Is 40% visibility good or bad?
- How does the missing 60% affect validity?
- Do you agree/disagree?

  The mere fact that so many of phishing websites in our dataset embed third-party resources shows that attackers do not fear being detected by certain organizations

- Is this approach useful as a detection, rather than measurement technique?

Limitations and Improvements

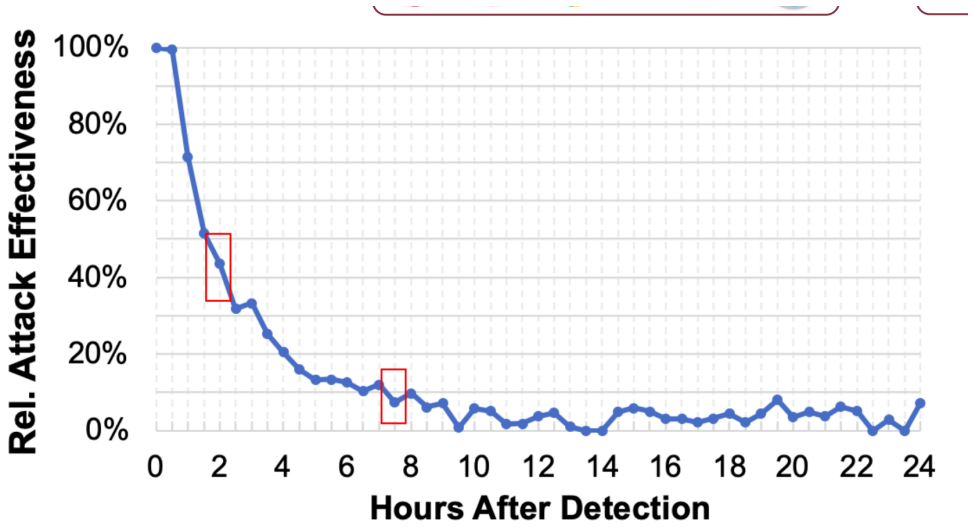
- “Only” one organization studied
  - What could be found out from a broader analysis
  - How would phishing differ in other contexts
- Other potential ways to slice the data?
- Other data sources to include?

Contributions

- Is this a good type of paper?
  - Novel methodology, but will it last?
  - Large scale data, but with limitations?
  - Point-in-time, proprietary

Better detection / mitigation

- How can we improve the 9-hour detection window?
- How can we help the 37.73% of victims who get phished *after* detection?



Data Sharing and Privacy

- Are there opportunities for better data sharing among organizations?
- Is the data currently being collected too privacy-invasive?

Other discussion?