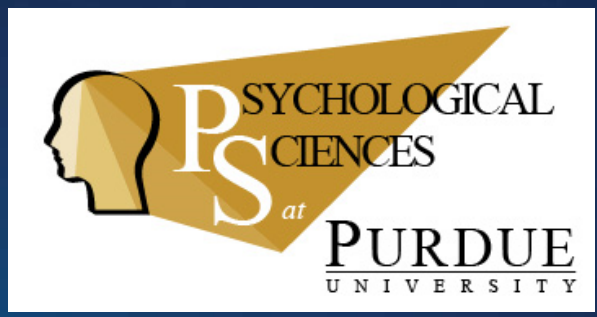




# A Human Information-Processing Analysis of Online Deception Detection

2/2/2016 NSA Quarterly Meeting

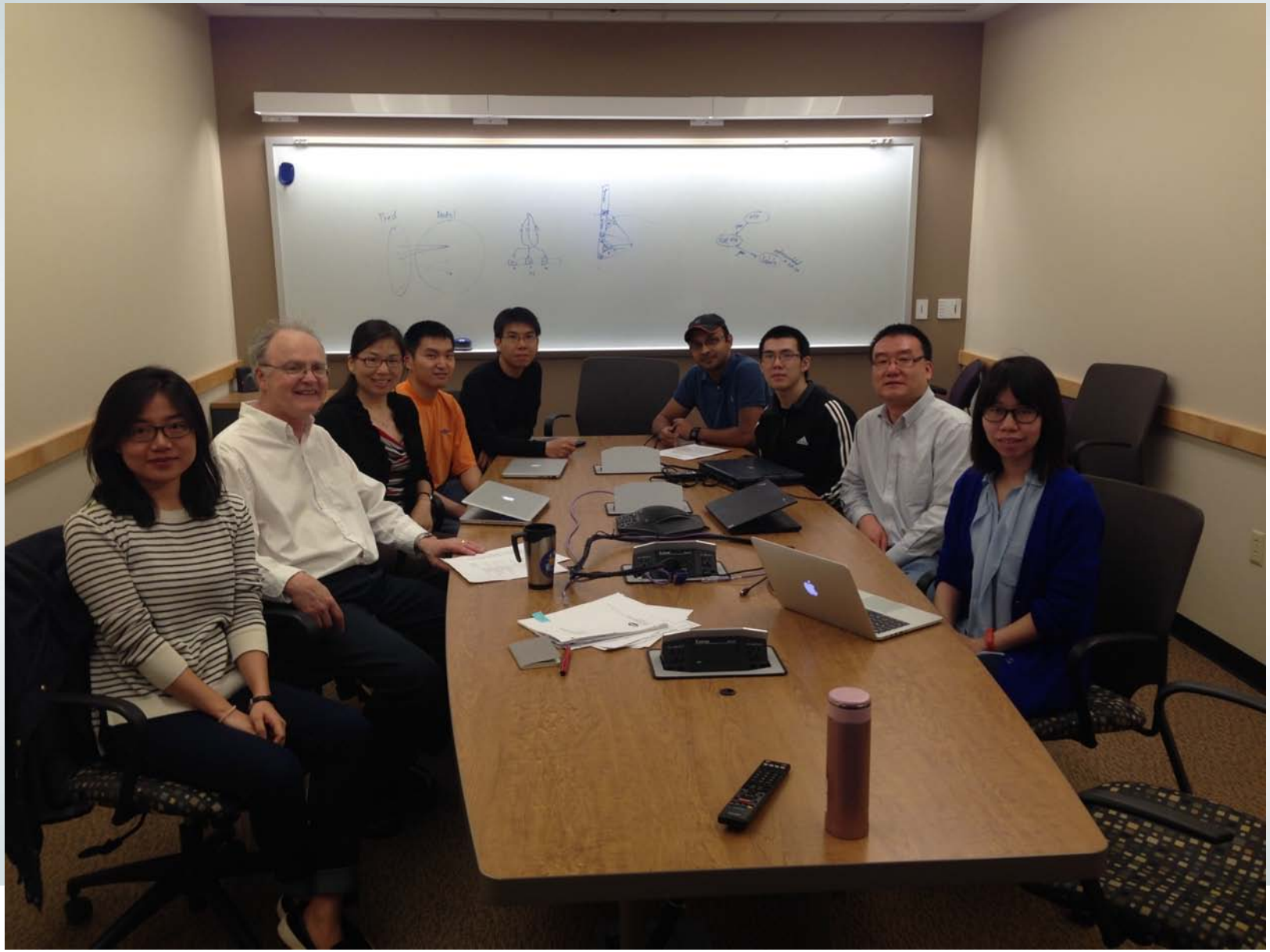
Robert Proctor

Funding for this research was provided by
the National Security Agency as part of a Science of
Security lablet through North Carolina State University

# Purpose

- Hard problem: **Understanding and Accounting for Human Behavior**
- User's decisions when faced with possible phishing attacks
  - A field experiment of phishing warning
  - An online study of domain-highlighting

# Interdisciplinary Team



- **Ninghui Li**, CS
- 5 CS grad students & postdoc
- 3 PSY/IE grad & undergrad students

# Research Topics

## App Selection Decisions

- Prior lablet grant and NSF grant
- When in the selection sequence, and how, to display risk information
- Combination of M-Turk studies and lab experiments in which participants have to make choices
- 2 journal articles and 3 conference proceedings papers.

## Scientific Understanding of Firewall Policies

- Security policies can be very complex, in the sense that they are difficult for humans to understand and update
- Developed an automatic tool for converting a firewall policy into modularized form
- Conducting an online study in which M-Turk Users receive training on typical or modularized firewall policies and then are tested on the respective type

# Phishing

## Phishing Lifecycle:

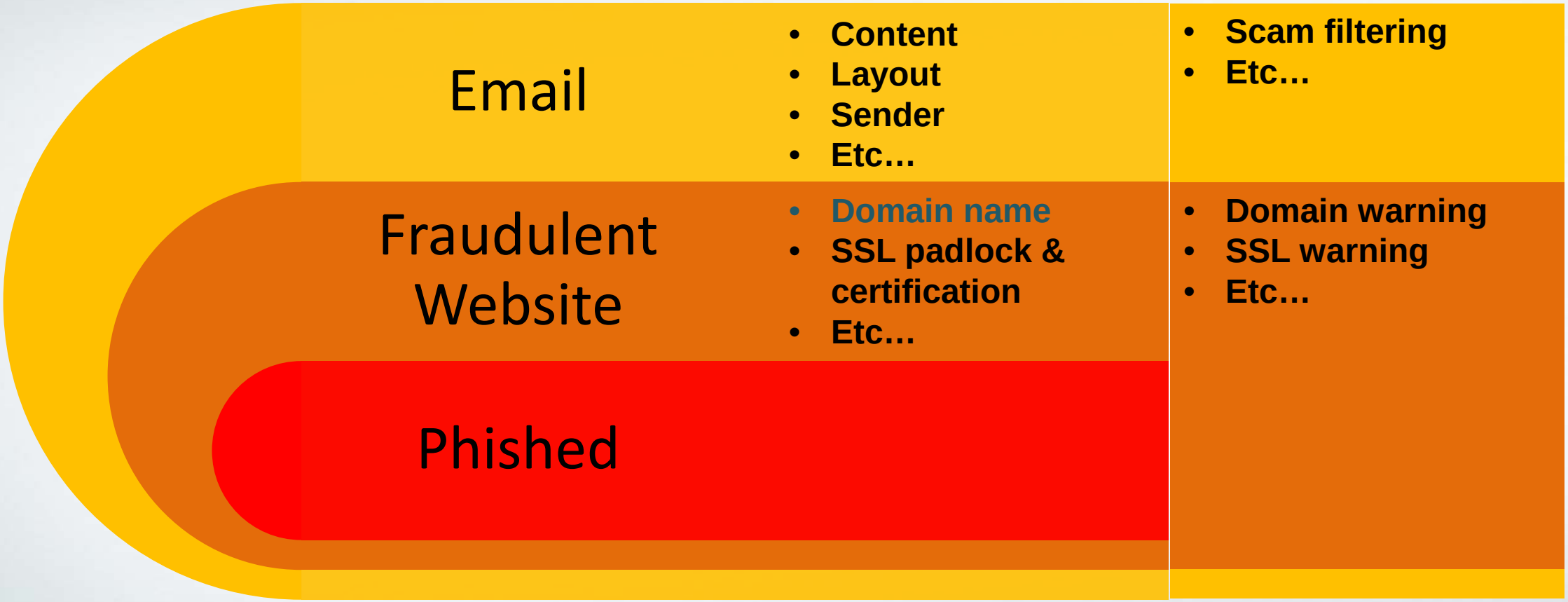
- starts from an unsolicited email sent by the deceiver posing as a legitimate party
- continues on the fraudulent webpage mimicking the authentic one after users' click on the link within the email
- ends with victims entering personal and credential information

# Research Against Phishing Attacks

- Computational methods for detecting phishing attacks have been ineffective
- Decision aid tools
  - Limited success, usability problems
- Studies have shown:
  - Users' attention is dominated by visual cues reinforcing webpage legitimacy, while ignoring the browser-based security cues;
  - Users are not familiar with phishing attacks and have difficulty understanding security warnings

# Phishing Detection and Prevention

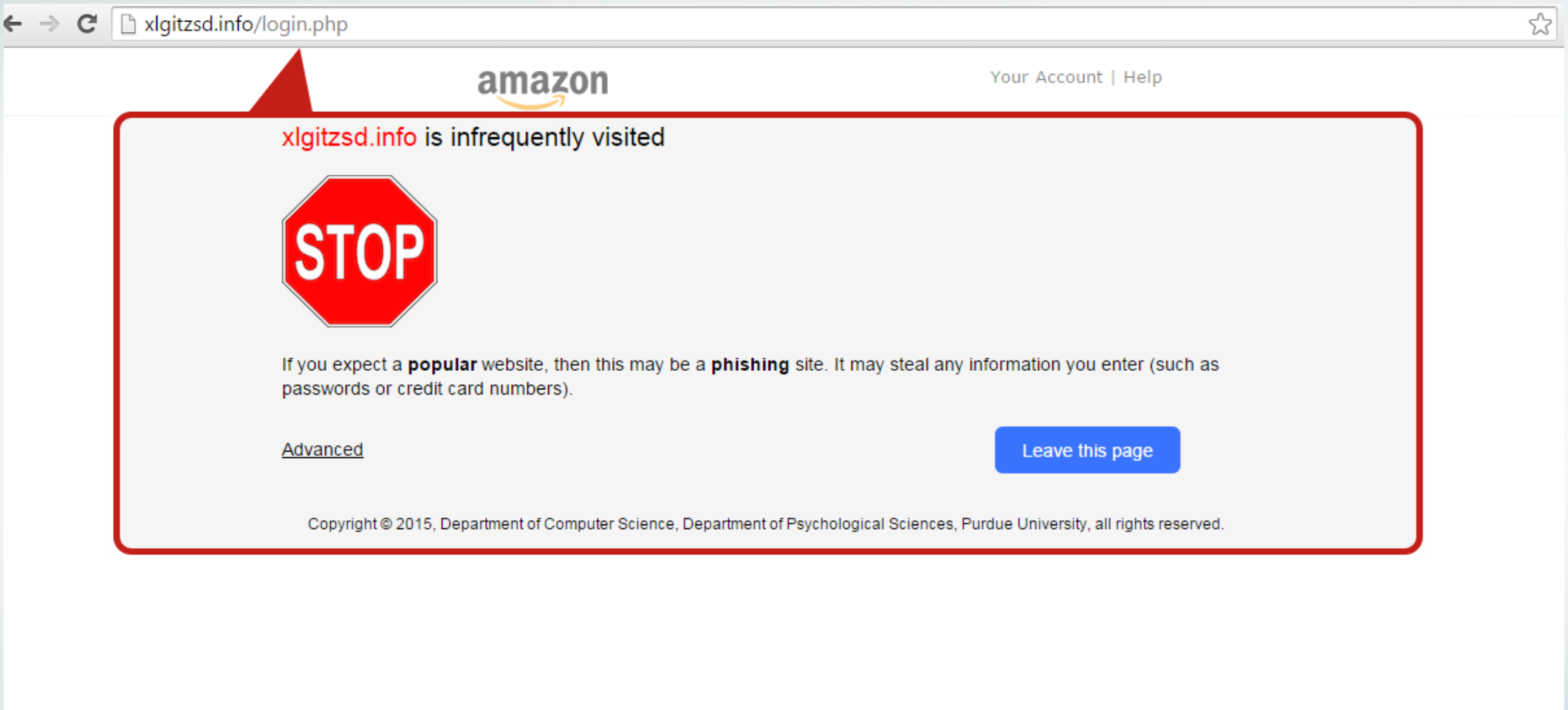
- Domain name is the most reliable telltale sign of a phishing website.



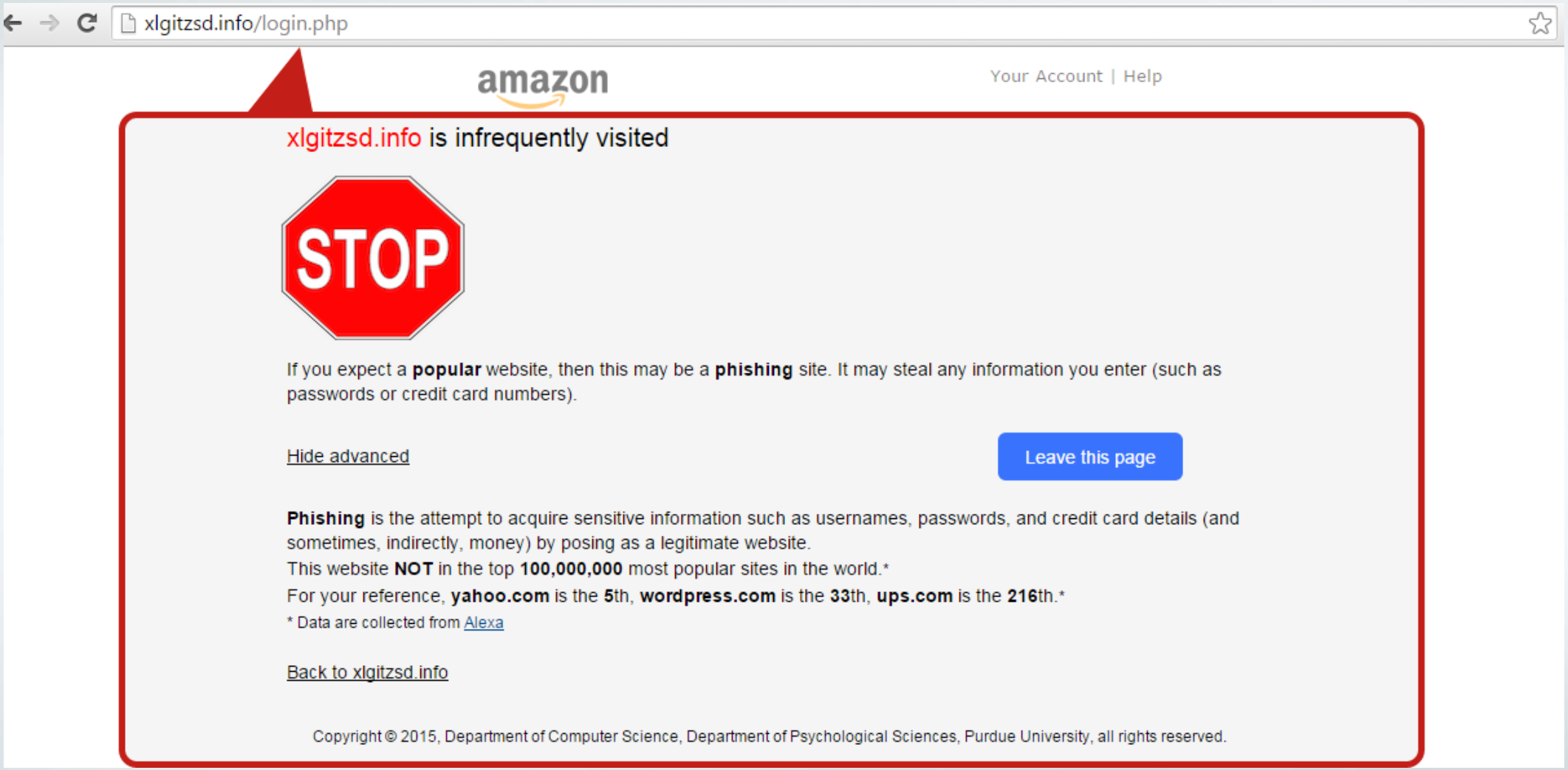
# A Chrome Extension Warning

- Domain popularity difference between phishing websites and legitimate popular websites
- Phishing sites are visited infrequently
  - More than 90% of them with a rank > 100,000
- Using that value as a criterion for displaying warning, have miss rate of only about 10% and acceptable, low false-alarm rate
- Active warning

# Warning Interface



# Warning Interface



# Phishing Detection and Prevention

**User's attention (STOP):**

- Stop sign to attract attention
- Domain name extracted from URL to aid user's decision about the website's legitimacy

**User's understanding (THINK):**

- Specific and complete identification of the risk
  - Without technical language
  - Not so lengthy that it takes time and effort to read the warning
- Explicit explanation of consequences if exposed to the risk

**User's action (ACTION):**

- Highlight the recommended action

# Phishing Education

## Included:

- Learn what phishing is
- Understand the implications of phishing
- Ways in which phishing can be detected & combatted

Few studies have combined education about phishing with other measures intended to protect users from phishing attacks

# Experiment Design

- A 3-week field experiment
  - The phishing warning Chrome extension installed for daily computer use
  - Participants (58) informed that they were taking part in a study about browser behavior of daily use
- 2 (phishing warning)*2 (phishing education) groups:
  - Group 1: phishing warning & phishing education
  - Group 2: phishing education
  - Group 3: phishing warning
  - Group 4: no phishing warning & no phishing education
- Phishing scenario that replicates a popular commercial website promotion
  - In week 3: email of Amazon gift card including links associated with a newly registered “phishing” domain maintained by us, simulating phishing attacks

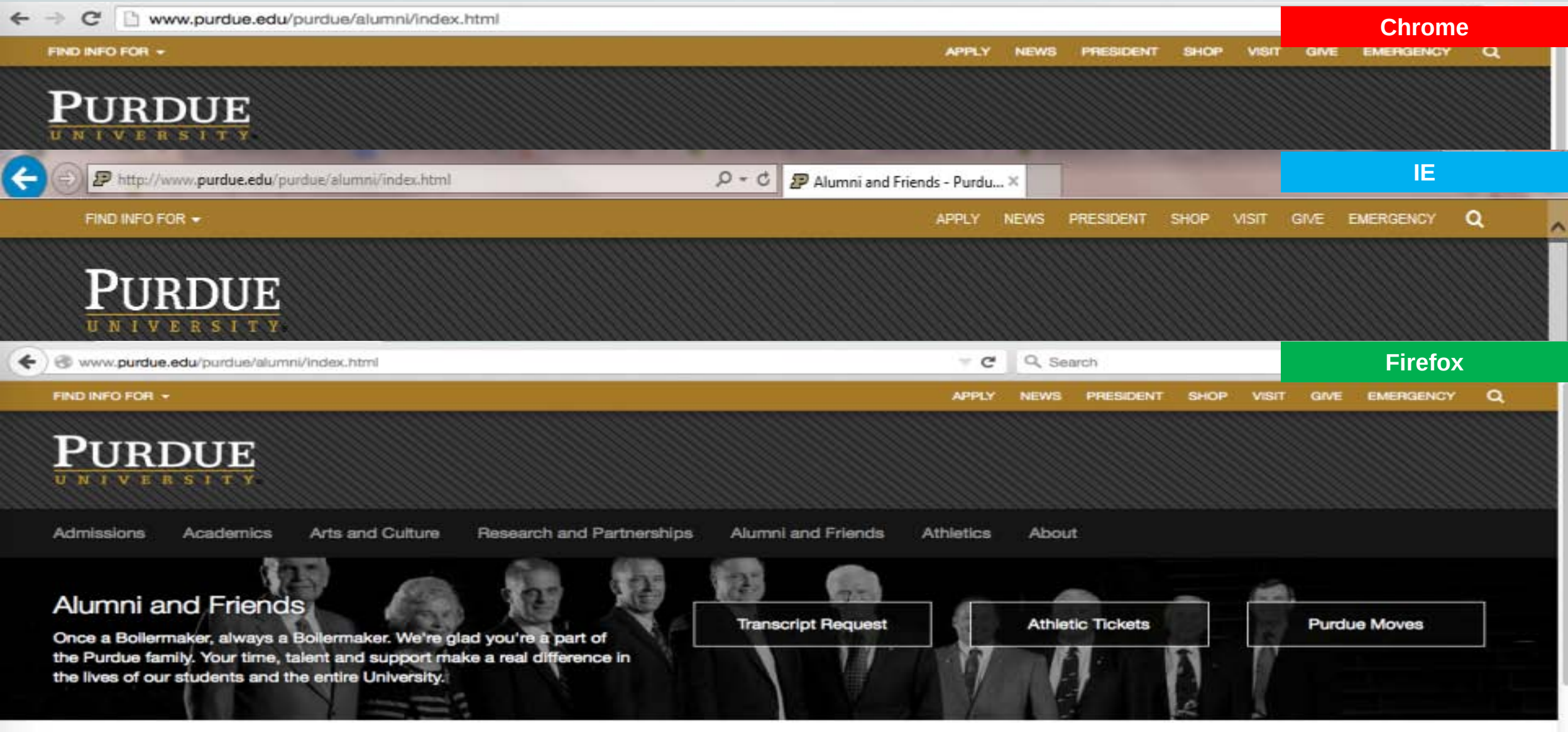
# Results

- % users who were phished:
  - 0 (0/19) for group 1 (phishing warning & phishing education)
  - 42% (7/17) for group 3 (phishing warning)
  - 78% (7/9) for group 2 (phishing education)
  - 69% (9/13) for group 4  (no education & no warning)
- No participant fell prey to our simulated phishing attack if they got the phishing education and saw the phishing warning.
- Phishing education was not effective alone, but it reduced the success of the attack when phishing warning was presented.

# Summary

- Warning can significantly reduce the number of users being phished
- Best performance obtained if users are warned and equipped with the knowledge about phishing

# Domain Highlighting



# Domain Highlighting

- Domain name highlighting is implemented to:
  - "… *make it easier for you to identify the site you visit*. This helps alert you to deceptive websites that try to trick you with misleading address and can help reduce the chances of compromising your personal information."
  - "…*quickly identify* what site you're on, and avoid getting tricked into think it's a different site."

# Three Assumptions

- Users will naturally attend to the address bar
- Users will use the domain name to judge the website's legitimacy
- Users can recognize legitimate domain names

# Effectiveness of Domain Highlighting

- Lin et al. (2011) conducted a laboratory experiment:
  - 22 participants
  - 16 webpages (8 legitimate and 8 fraudulent) based on a 5 point scale
  - 2 phases: no instructions were given for 1st phase
    directing users to look at the address bar for the 2nd phase
- Results showed a small benefit in correctly identifying fraudulent websites when participants were directed to look at address bar.

# Effectiveness of Domain Highlighting

- Limitations:
  - Only the domain highlighted webpages were used in their study without a control condition.
  - The experiment included a small number of participants.

# Current Study

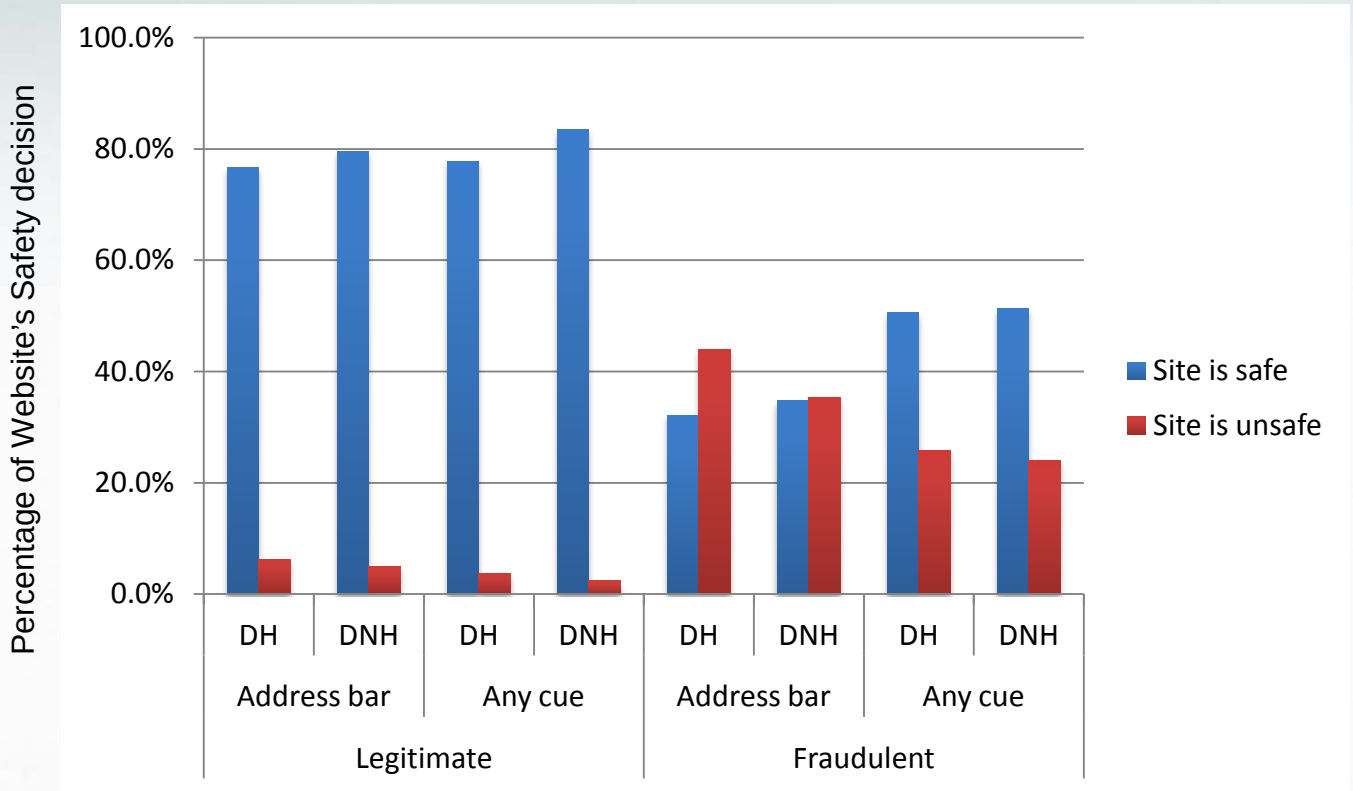
- MTurk online study
  - 320 participants recruited and restricted to the North America region
- Followed the 2 phase studies of Lin et al. (2011), except:
  - 12 webpages (2 groups of 6 frequently targeted website categories)
  - Domain highlighted or not was manipulated between two phases
  - Snapshots of login webpage

# Task Trial


Please gauge the safety of the above webpage with the following scale:
*(1 means unsafe and 5 means safe)*

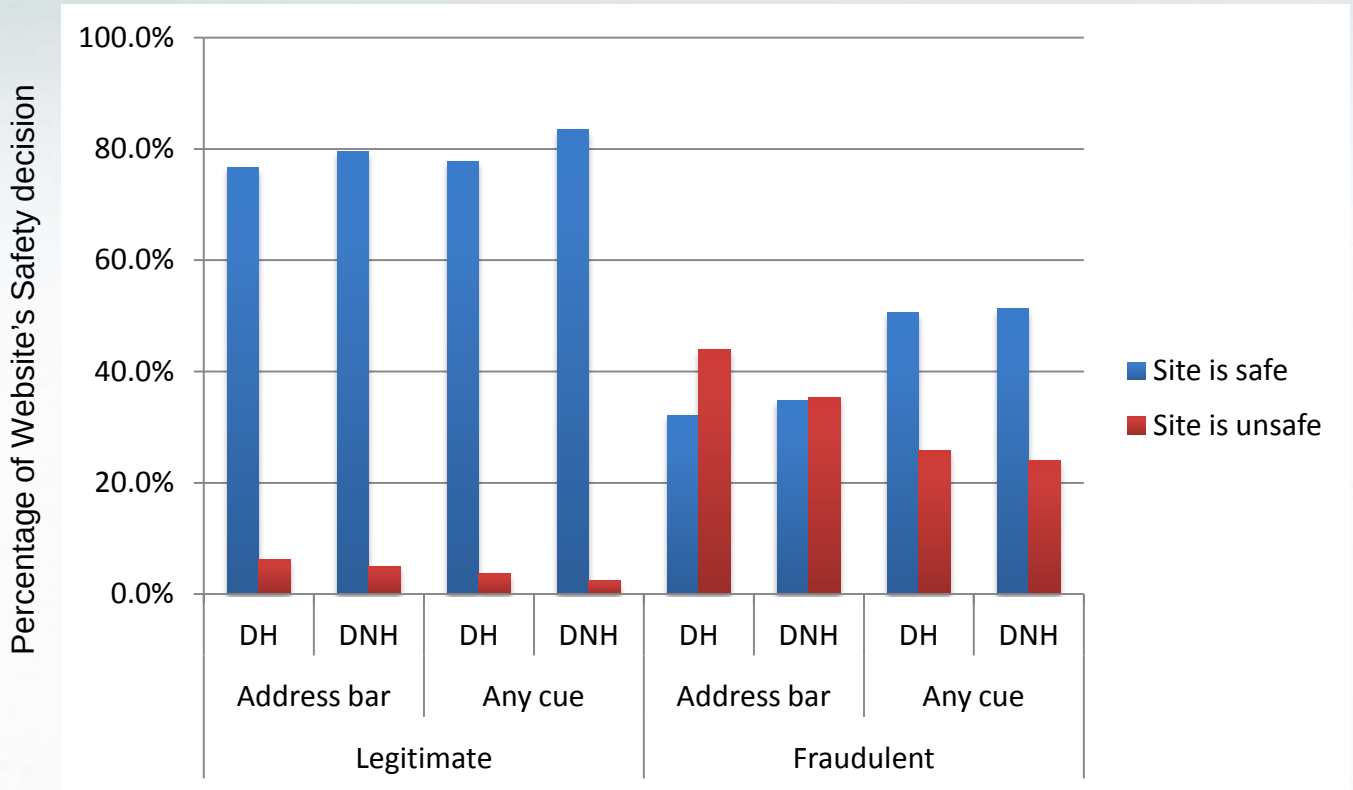
| ○ | ○ | ○ | ○ | ○ |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

# Results



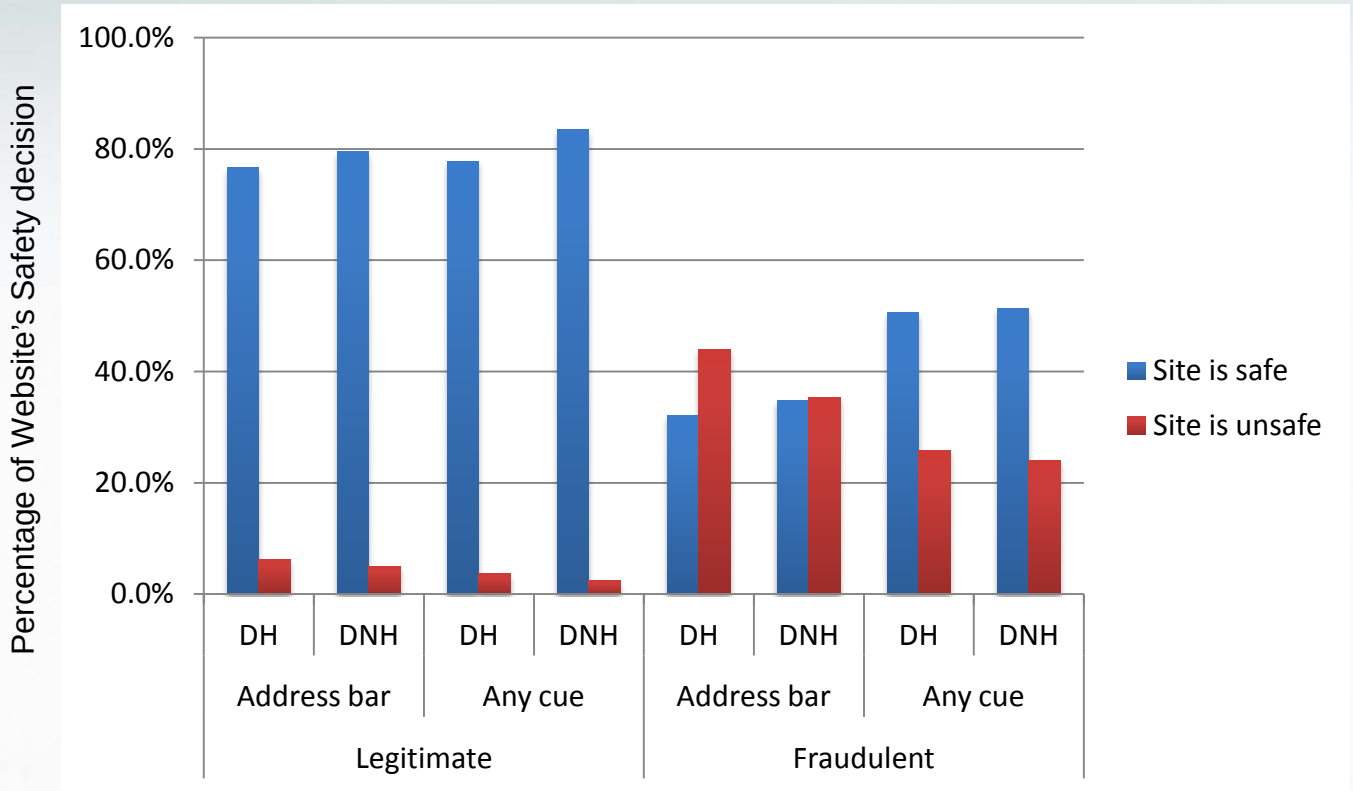

- A significant difference was found between the frequency of people's rating of web page type.

# Results




- The frequency of people's correct rating of fraudulent webpage was higher for the address bar than for the any cue.

# Results




- Added as an extra factor, domain highlighting showed no significant difference across all conditions.

# Post Session Questions:

- Did you know about the domain highlighting feature before this survey?

  Yes: 36%
  No: 64%

- Have you noticed the domain highlighting feature previously?

  Yes: 41%
  No: 59%

# Summary

- Directing user's attention to the address bar is slightly beneficial to help users detect phishing web pages.
- Domain highlighting gives almost no protection for the users to identify suspicious websites.
- Currently conducting eye-tracking study.

# Conclusion

- Drawing people’s attention to warning or domain name is not sufficient to protect users from phishing attacks.
- Both studies’ results showed the problems that could be associated with a lack of user’s knowledge of phishing.

# THE END