# Alice in Battlefield:

# An Evaluation of the Effectiveness of Various UI Phishing Warnings

Ranran(Monica) Bian
(ID: 1706953)
The University of Auckland
Department of Computer
Science
Auckland, New Zealand
rbia002@aucklanduni.ac.nz

## ABSTRACT

Phishing warning is one of the most popular and important web browser warnings. Hence, there has been a variety of User Interfaces (UI) designed and implemented for web browser phishing warnings. Some of them are still in use while others are obsolete and deprecated. By comparing and contrasting three experiments conducted by three major American universities, this paper explores the effectiveness of various UI phishing warnings as well as evolvement of phishing warning UI designs and implementations. All of the three most well-known web browsers are discussed during the exploration, which are Google Chrome, Internet Explorer and Firefox.

## KEYWORDS

Web browser phishing warnings, UI, effectiveness, evolvement, Google Chrome, Internet Explorer, Mozilla Firefox, story-telling, Alice, Stanford, MIT.

## 1. INTRODUCTION

Phishing attack usually uses legitimate-looking but fake websites and emails to deceive users into disclosing personal or financial information such as usernames, passwords, credit card details(and sometimes, indirectly, money) to the attacker. This type of attack is also capable of tricking users to download and install hostile software, which scans the user's personal computer or monitors user's online shopping activities in the purpose of stealing vital private information. In

the past decades, phishing attack has indeed become a significant threat to internet users.

Each of the attacks listed below represents one phishing attack technique that has been recorded by APWG (Anti-Phishing Working Group) [1]:

- Homograph Attack: A phishing webpage with the URL displayed in the address bar, which is superficial similar to legitimate website's domain name. This is widely recognized as a very effective phishing technique. One example is the attack site www.bankofthevvest.com for Bank of the West. Internet Explorer 7 has a gap between the two "v" letters in the address bar as compared to the font used in the address bar for Mac Firefox which has no gap between the two characters. As a result, Mac Firefox makes it quite difficult for the end users to detect the phishing attack.
- Picture-in-picture Attack: A phishing webpage that shows a fake browser window which looks like showing of the legitimate webpage. According to C. Jackson et. all [2007], picture-in-picture attack has high success rate and is as effective as the similar-name attack. Figure 1 shows a picture-in-picture attack example.


**Fig. 1.** Picture-in-picture attack.

- IP-address Attack: A phishing webpage that displays an IP address instead of website's domain name in the address bar in the purpose of obscuring a server's identity.
- Hijacked-server Attack: A phishing webpage which is supported by server at a legitimate company but the server is already hijacked by the phishing attackers.
- Popup-window Attack: A phishing site that uses borderless popup-windows on top of real site webpage to request the user's private information.

So far, a surprising large number of phishing web sites have been reported to APWG and phishing attacks have cost financial institutes and card issuers billions of dollars [5]. To tackle this issue, phishing warnings were introduced, which aim to prevent internet users from visiting websites that intend to deceive users for their private information or offer malicious executables. Unfortunately, *while the priest climbs a foot, the devil climbs ten,* like the scene in "Casino Royale" movie where the drug dealers used statistic report produced by the United Nations

organization as their own marketing expanding proposal, phishing attackers also instantly adopt research-proven effective techniques to improve their web credibility. A funny but possible assumption is that phishing attackers learn about all the published researches and experiments on the effectiveness of phishing attacks and phishing warnings to keep their techniques updated all the time.

Having set the scene, by comparing and contrasting three experiments conducted on different phishing warning user interfaces provided by world top three used web browsers (Figure 2), this papers adopts a story-telling mode to explore whether these phishing warning UIs are effective at helping users to differentiate legitimate websites from phishing websites and preventing users from entering their credentials into phishing websites. In 2005, M. Wu et. all [4] from MIT(Massachusetts Institute of Technology) conducted two studies of Internet Explorer's three types of security toolbars and some other security indicators to find out whether they are effective at preventing phishing attacks (one of the three characters in the story and referred as **MIT** (in bold) in the rest of this paper). A year later, C. Jackson et. all [3] from Stanford University did an experiment to evaluate the effectiveness of Extended Validation (EV) in Internet Explorer 7 against picture-in-picture attacks (one of the three characters in the story and referred as **Stanford** (in bold) in the rest of this paper). In 2013, D. Akhawe and A. P. Felt [2] conducted a large-scale field study of Google Chrome and Firefox's security warning effectiveness (primary character in the story and referred as **Alice** (in bold) in the rest of this paper).

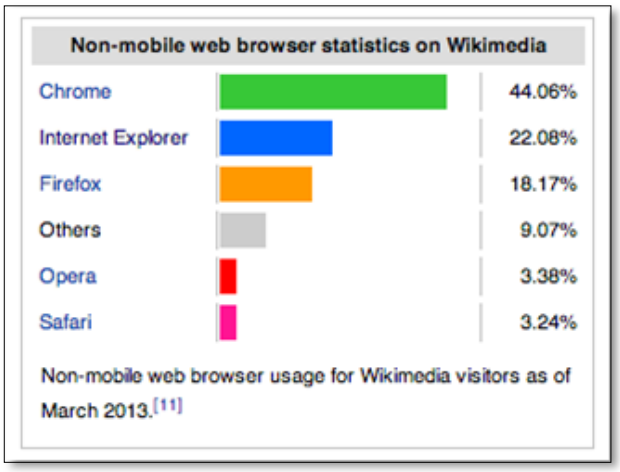


**Fig. 2.** Web browser statistics.

There are two reasons that the story is named "Alice in Battlefield". First and foremost, it is a brutal battle for all three web browsers against phishing attacks. Second equal important reason is that only **Alice** demonstrated that browser security warnings can be effective in practice while both **MIT** and **Stanford** concluded the ineffectiveness of browser anti-phishing defenses. In some way, **MIT** and **Stanford** support the popular opinion "Given a choice between dancing pigs and security, the user will pick dancing pigs every time [6]". Ironically, that is the exact theory **Alice** tries very hard to overturn.

The rest of this paper is organized as follows. We begin by introducing each of our three leading

characters (MIT**, Stanford and Alice**) in more details. The next section tells the "Alice in Battlefield" story in 12 small chapters, which includes analysis of the effectiveness of various UI phishing warnings, similarities and differences presented by three studies; appreciations and criticisms on three experiments; revolutionary development of browser phishing warning designs and implementations between **MIT**(2005) and **Alice**(2013). Next, some possible future work is discussed. At last, we conclude this paper with some highlights from our story.

# 2. THE DESCRIPTION OF OUR THREE LEADING CHARACTERS

This section describes each of our three leading characters from three important aspects.

## 2.1 MIT

- Focused Phishing Warning UI: Three simulated security toolbars (Figure 3) in Internet Explorer by grouping the features and types of information displayed of five existing toolbars.


**Fig. 3.** The three simulated toolbars used in MIT.

- Methodology and Experiment: The stimulated man-in-the-middle phishing attack whose content is a perfect copy of the legitimate site was studied. User's sensitive information was captured illegally during transportation of user's submitted data to the real site. Experiment was carried out in a lab environment, where 30 subjects were randomly divided into three groups and each of the three security toolbars was tested on 10 subjects. Some efforts were made to hide from the subjects that security is the primary goal of the experiment. Tutorial emails were distributed to subjects in the middle of the study. Spoof rates ("the fraction of simulated attacks that successfully obtain user's username and password or other sensitive information without raising user's suspicion." [4]) for the following four scenarios were measured and analyzed: different toolbars; before and after the tutorial; wish-list attacks

at different attack positions; a regular browser interface and the blocking warning box.

- Main Findings: Many subjects failed to look at the security toolbars; others explained away the security toolbars' warnings as long as the content of web pages looked legitimate. The usability of the anti-phishing defenses studied was unsatisfactory as they did not significantly help users detect high-quality phishing attacks.

## 2.2 Stanford

- Focused Phishing Warning UI: Passive indicator, which non-interruptively show the status of the HTTP(S) connection in Internet Explorer 7 user interface. Extended validation (EV) is actually the passive indicator in the experiment, which unlike normal certificates, not only verify that the owner controls a specific domain name but also attest to the identity of a legal business. IE 7 shows the presence of the EV by turning the address bar green and providing more information about the certificate owner, as shown in Figure 4.


**Fig. 4.** Extended validation (EV) indicators

- Methodology and Experiment: The experiment was conducted in an artificial lab environment, where 27 subjects were randomly divided into three different groups and each subject classified 12 web sites as phishing or legitimate. Three groups are differentiated from each other by information level of EV that the users received before and during the experiment. The effect of passive EV indicator and the effect of reading a training help file about security features in IE 7 were both measured at the end of the study.
- Main Findings: Passive EV indicator did not help prevent users from phishing attacks; passive EV indicator did not help untrained users classify a legitimate site; training document made more real and phishing sites to be considered as legitimate.

## 2.3 Alice

- Focused Phishing Warning UI: Contemporary browser-provided, full-page interstitial phishing warnings in both Google Chrome (Figure 5) and Mozilla Firefox (Figure 6) that discourage the user from proceeding.

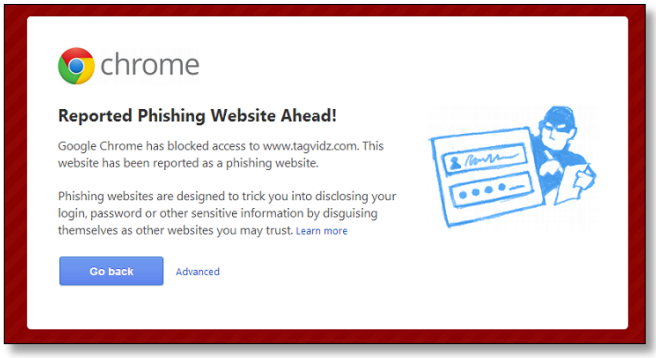


**Fig. 5.** Phishing warning for Google Chrome

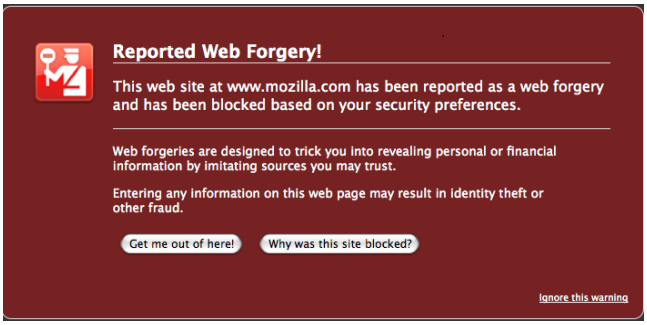


**Fig. 6.** Phishing warning for Mozilla Firefox

- Methodology and Experiment: A large-scale field study of user decisions after seeing browser phishing warnings was performed in May and June 2013. Data was collected using the browsers' telemetry frameworks, "which are a mechanism for browser vendors to collect pseudonymous data from end users" [2] without disturbing end users' normal browsing activities. Clickthrough rates ("describes the proportion of users who clicked through a warning type." [2]) were computed to evaluate the effectiveness of this kind of phishing warning UI.
- Main Findings: Security warnings can be very effective in practice; internet security experts and system architects should never dismiss the goal of communicating security information to end users.

| Browser Warning Type | Clickthrough Rate |
|---|---|
| Mozilla Firefox Phishing Warning | 9.1% |
| Google Chrome Phishing Warning | 18.0% |

**Fig. 7.** Main data findings of **Alice**

# 3. THE SOTRY:

# ALICE IN BATTLEFIELD

Now, buckle up your seatbelt and sit tight for the upcoming story. Each chapter in the story represents one unique aspect for comparing or contrasting, appreciating and criticizing **MIT**, **Stanford** and **Alice**.

***Chapter 1***

***Effective Phishing Attacks***

One of **Stanford**'s key findings is that both homograph attacks and picture-in-picture attacks are very effective internet phishing attacks. **MIT** consolidates this finding by presenting the similar-name attack had the highest spoof rate, 50%, among all five simulated phishing attack techniques studied.

***Chapter 2.***

***What's the Priority of Security for Users?***

As mentioned in previous section, the theory that **Alice** wants to overturn is that "Given a choice between dancing pigs and security, the user will pick dancing pigs every time [6]". Although this theory can't be proven true in all possible circumstances, however, in some way, it shows that the primary goal of end users isn't usually security. **MIT** claims that "in real life, security is rarely a user's primary goal" [7] as end users are more concerned with other tasks, such as online shopping, email reading, document editing and so on. Avoiding disclosure of private information is

probably considered as important, but it isn't foremost in end users' mind. **Stanford** seconded **MIT**'s claim by mentioning that "In actual usage scenarios, security is rarely a user's primary goal" [7]. It is undoubted that security matters a lot to computer scientists and internet security experts, but just not that much to everyday users.

***Chapter 3:***

***Accuracy of Browser Phishing Warnings***

This is one of the areas that **Alice** had noticeable discrepancy from **MIT** and **Stanford**. According to **Stanford**, it is very difficult for phishing filters which are integrated into web browsers to maintain perfect accuracy, as a result, web browser training documentation should be carefully designed and not to give users the impression that the phishing filter is 100% accurate. **MIT** piled on this opinion by declaring that different security toolbars have different levels of accuracy. On the contrary, **Alice** claims that the false positive rate for all Safe Browsing warnings is low enough to be negligible. One way to explain this discrepancy is that both **MIT** and **Stanford** were being too conservative on the matter. However, a more possible explanation is that with the rapid development of information technologies over the last decade, web browser vendors advanced their techniques to sustain the perfect accuracy of the phishing filter, which is flawless to some extent.

***Chapter 4:***

***Different Warning UIs Studied***

As described in Section 2, both **MIT** and **Stanford** studied passive indicators of browser phishing warnings while the focused warning UI of **Alice** is full-page interstitial warning box. **MIT** mainly experimented with three types of security toolbars in Internet Explorer and **Stanford** focused on status of IE address bar in the presence of extended validation.

***Chapter 5:***

***Different Experiment Settings***

Both **MIT** and **Stanford** conducted their experiments in artificial laboratory environments, where some inevitable biases existed. During their experiments, some efforts were made to divert subjects' attentions so that they wouldn't instinctively realize that security is the primary concern of the experiments. However, the effectiveness of those efforts remains as a big question mark because there were still some clues about the primary goal given throughout the experiments. Among all three, **Alice** apparently had the best ideal experiment setting, where the telemetry frameworks integrated in Chrome and Firefox were adopted to collect genuine user data without disturbing their normal browsing activities. Once again, thanks to the rapid development of technologies. The telemetry framework is appreciated more as compared to the experiment settings **MIT** and **Stanford** had and it kind of raises a comparability issue (which

this paper chooses to ignore) for the three experiments.

***Chapter 6:***

***Different Rates Cared***

Again, as described in Section 2, the three experiments were interested in different kinds of rates. **Alice** used "Clickthrough rates" as the primary benchmark while **MIT** constantly measured "Spoof rates" throughout the experiments. The formal definitions of these two rates are presented above in Section 2.

***Chapter 7:***

***Opposite Main Findings***

Apparently, both **MIT** and **Stanford** took the side which declares the ineffectiveness of two phishing warning UIs in Internet Explorer. At the opposite side, **Alice** showed strong confidence on modern full-page interstitial warning user interface. An interesting correlation between **MIT** and **Alice** is that **MIT** admitted that compared to the passive security toolbars, it is more effective interface for getting user's attention about a phishing site to actually block access to it by popping up a modal dialog box when the phishing site is visited. The conclusion drawn from **MIT**'s follow-up study is that warning box blocking the phishing webpage is a much stronger signal than the passive indicators. All of these **MIT**'s (2005) statements are considered as correlations because they describe the same warning user interface as described in **Alice** (2013). To some level, **MIT** foreseen and suggested the phishing warning UI designs and implementations for the next six years. An interesting discrepancy between **MIT** and **Alice** is that **MIT** considered 10% spoof rate with a blocking warning box as ineffective after admitting it's much better than spoof rates for passive security toolbars. On the other hand, as shown in Figure 7, the clcikthrough rates for Firefox and Chrome phishing warnings were 9% and 18% respectively and they were considered quite effective by **Alice**. So the question comes down to who's the king to define how effective is effective? No matter how perfect the phishing warning UI is, there is no way that we can expect 0% spoof rate or clickthrough rate because each user has their own interpretations and unique user behaviors on phishing warnings. One of **Alice**'s main purposes is to overturn the popular suggested opinion that browser security warnings are just ineffective. This paper agrees Alice in some way and believes that the opinion is not accurate at all circumstances but it is true positive in some cases as declared by **MIT** and **Stanford**.

***Chapter 8:***

***Who's Job to Fight Against Phishing Attacks?***

**Stanford** discussed that the effectiveness of extended validation would increase as the technique is adopted by more financial web sites and public awareness grows. **MIT** mentioned that the security toolbars became more effective as the experiment went along and subjects had more experiences with them. It sounds like both

**Stanford** and **MIT** were suggesting that we should rely on end users for the effectiveness of phishing warnings. This is an invalid suggestion as no one should expect an 80-year-old lady to learn about phishing warnings and obey them to protect herself from phishing attacks. End users could easily make wrong judgment and decision based on their own interpretations and existing experiences on the site or the warning. For example, it would be extremely difficult for end users and probably out of their capability to use their naked eyes to detect visually deceptive domain names. Luckily, both **MIT** and **Alice** listed some parties that should be responsible for fighting against phishing attacks:

- Internet security experts
- Security warning UI designers and researchers
- Internet companies should follow some standard practices to better distinguish their sites from malicious phishing sites.
- E-commerce firms should adopt good web practices to make phishing attacks less likely to succeed.
- Operators should not use domain names that are vague, inconsistent, or otherwise unrelated to their brands.
- Organizations should not make their outsourcing relationships directly visible to Internet users.

***Chapter 9:***

***Suggestions for Browser Warning Designers***

Below are some suggestions for browser warning designers made by **Alice** and **MIT**:

- Eye-catching warning appearance to immediately get user's attention or to force them to stop for a second
- Providing remembering exception mechanism
- Avoiding Warning Fatigue
- Do not hide important information from the main page of phishing warning UI
- Proposing an alternative paths (e.g., directing users to the real intended site) with the warning and allowing users to complete their intended task safely eventually.
- Integrating security concerns into critical path of user intended tasks so that the security warning cannot be ignored anymore.

***Chapter 10:***

***False Positives***

All of the three characters agreed on the negative effect of false positives on end user's trust on browser phishing warnings.

***Chapter 11:***

***Explanatory Links***

The large-scale field study conducted in **Alice** shows that users rarely click on the explanatory links such as "Learn More" or "More Information". **MIT** consolidated this finding by presenting that only 7 out of 30 subjects clicked the toolbar's "What's this?" explanatory link.

***Chapter 12:***

***Color***

Both **MIT** and **Stanford** use colors (Green, Yellow and Red) in their phishing warning UIs to represent different security statuses. However, what if the end user is color-blind and would it still be noticeable and useful? Does the meaning of the color system remain unchanged and consistent forever? Does the color system make much sense for untrained end users? Would it be more user-friendly if a smiley face scale system (Figure 8) was used?

**Fig. 8.** Smiley face scale system

# 4. FUTURE WORK

Having compared and contrasted phishing warning UIs in Internet Explorer (2005 & 2006) with those provided by Google Chrome and Mozilla Firefox (2013) in the story, it will be quite interesting to find out what is happening with contemporary IE phishing warning UI designs and implementations. A wild guess is that IE has also adopted full-page interstitial warning UI like Chrome and Firefox did.

As quoted in the Introduction section, *while the priest climbs a foot, the devil climbs ten*, internet security experts, web browser warning designers and researchers should always watch out and monitor the latest and state-of-the-art phishing techniques and adjust existing warning UI designs and implementations accordingly in order to provide more effective and sound protections to end users.

# 5. CONCLUSIONS

By telling the "Alice in Battlefield" story, this paper evaluates the effectiveness of various UI phishing warnings and explores evolvement of phishing warning UI designs and implementations over the last decade. In some level, it is believable that the effectiveness presented by **Alice** derived from the ineffectiveness demonstrated by both **MIT** and **Stanford**. According to APWG latest report released on July 31, 2013: phishing attack numbers dropped 20% from historical highs [1], the effectiveness of contemporary phishing warning UI as shown in **Alice** could have contributed greatly to this fact. Although due to the limited number of literatures reviewed, this paper is only capable of providing a partial answer to the interested research topic; however, we hope that you did enjoy the story.

# 6. REFERENCES


[1] Anti-phishing working group: http://www.antiphishing.org

[2] Akhawe, D., & Felt, A. P. (2013, August). Alice in warningland: A large-scale field study of browser security warning effectiveness. In Proceedings of the 22th USENIX Security Symposium. Available: http: //www.cs.berkeley.edu/devdatta/papers/alice-in-warningland.pdf

[3] Jackson, C., Simon, D. R., Tan, D. S., & Barth, A. (2007). An evaluation of extended validation and picture-in-picture phishing attacks. In Financial Cryptography and Data Security (pp. 281-293). Springer Berlin Heidelberg. DOI: 10.1007/978-3-540-77366-5_27 Available: http: //dx.doi.org.ezproxy.auckland.ac.nz/10.1007/978-3-540-77366-5 27

[4] Wu, M., Miller, R. C., & Garfinkel, S. L. (2006, April). Do security toolbars actually prevent phishing attacks? In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 601-610). ACM. DOI: 10.1145/1124772.1124863

[5] Loftesness, S.: Responding to "Phishing" Attacks" (2004), http://www.glenbrook.com/opinions/phishing.htm

[6] McGraw, G. (2006). Software security: building security in (Vol. 1). Addison-Wesley Professional.

[7] Whitten, A., & Tygar, J. D. (1999, August). Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In Proceedings of the 8th USENIX Security Symposium (Vol. 99). McGraw-Hill. Available: http://www.usenix.org/events/sec99/full_papers/whitten/whitten.ps