

Using Contextual Information to Improve Phishing Warning Effectiveness

by

Satyabrata Sharma

A Thesis Presented in Partial Fulfillment
of the Requirement for the Degree
Master of Science

Approved April 2015 by the
Graduate Supervisory Committee:

Rida Bazzi, Chair
Erin Walker
Ashraf Gaffar

ARIZONA STATE UNIVERSITY

May 2015

ABSTRACT

Internet browsers are today capable of warning internet users of a potential phishing attack. Browsers identify these websites by referring to blacklists of reported phishing websites maintained by trusted organizations like Google, Phishtank etc. On identifying a Unified Resource Locator (URL) requested by a user as a reported phishing URL, browsers like Mozilla Firefox and Google Chrome display an ‘active’ warning message in an attempt to stop the user from making a potentially dangerous decision of visiting the website and sharing confidential information like username-password, credit card information, social security number etc.

However, these warnings are not always successful at safeguarding the user from a phishing attack. On several occasions, users ignore these warnings and ‘click through’ them, eventually landing at the potentially dangerous website and giving away confidential information. Failure to understand the warning, failure to differentiate different types of browser warnings, diminishing trust on browser warnings due to repeated encounter are some of the reasons that make users ignore these warnings. It is important to address these factors in order to eventually improve a users reaction to these warnings.

In this thesis, I propose a novel design to improve the effectiveness and reliability of phishing warning messages. This design utilizes the name of the target website that a fake website is mimicking, to display a simple, easy to understand and interactive warning message with the primary objective of keeping the user away from a potentially spoof website.

To Maa, Papa, Debu and Koka

ACKNOWLEDGEMENTS

I would like to thank many people for their constant support and guidance during the course of this work. First of all, I thank my advisor, Dr. Rida Bazzi, who guided me at every step in the project with his experience and helped me look into minute details of things, and also proposed the idea of using the target website information in a warning design. Thanks to my committee members, Dr. Erin Walker and Dr. Ashraf Gaffar, who were available for discussion and guidance at all times. Thanks to the Arizona State University for providing me an opportunity to work with such great minds and technology. Thanks to my colleagues in Dr. Bazzi's lab who, though not associated with this project, were always available for feedback and ideas whenever I needed their input in anything. Brainstorming sessions with them helped me tremendously in making decisions. And finally, thanks to my family, and many friends who endured this long process with me, always offering support and love.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vi
LIST OF FIGURES	vii
CHAPTER	
1. INTRODUCTION AND MOTIVATION	1
1.1 Motivation	2
1.2 Thesis	3
2. RELATED WORK: WARNING LITERATURE	5
2.1 Warning System	5
2.2 Browser Warning Effectiveness	6
2.3 Warning Ignorance Factors	8
2.4 Browser Warning Improvement	9
2.5 Phishing Warnings	12
3. PHISHING ATTACK: USING TARGET WEBSITE INFORMATION TO IMPROVE WARNING DESIGN	13
3.1 Technical Feasibility	14
3.1.1 Implementation Overview	15
3.1.2 Technical Details	16
3.1.3 Warning Design	20
3.1.4 Limitation	22
3.2 Usability Issues	23
3.2.1 Warning Design	24
3.2.2 User Understanding	25
3.2.3 User Trust	26
4. STUDY AND RESULTS	27

CHAPTER	Page
4.1 Hypothesis	27
4.2 Results	28
4.3 Limitations.....	31
5. CONCLUSION AND FUTURE WORK.....	32
5.1 Summary	32
5.2 Future Work	33
REFERENCES	34
APPENDIX	
A. SURVEY QUESTIONS	36
B. INSTITUTIONAL REVIEW BOARD APPROVAL	45

LIST OF TABLES

Table	Page
3.1 Top 10 Targets Identified by Phishtank	17
3.2 Breakdown of Targets Identified by Add-on	18
3.3 Unavailable URLs Breakdown.....	19
3.4 Breakdown of Targets with Multi Target URLs.....	19
3.5 Top Phishing Targets by Kaspersky	20
3.6 User's Desired Action Based on Understanding of Warning.....	26

LIST OF FIGURES

Figure	Page
3.1 Mozilla Firefox 34 Phishing Warning Page	14
3.2 Google Chrome 39 Phishing Warning Page.....	15
3.3 Proposed Warning Message - Stage 1	21
3.4 Proposed Warning Message - Stage 2	22
3.5 Proposed Warning Message - Stage 3	23
3.6 Proposed Warning Message - Stage 4	23
A.1 Appendix- Amazon Receipt.....	38
A.2 Appendix- Warning 1	38
A.3 Appendix- Warning 2	39
A.4 Appendix- Warning 3	40
A.5 Appendix- Warning 3	41
A.6 Appendix- Warning 3_2	42
A.7 Appendix- Warning 3_3	42
A.8 Appendix- Warning 3_4	43
B.1 Appendix- IRB Email Approval	46
B.2 Appendix- IRB Approval ERA Website.....	46
B.3 Appendix- IRB Approval Correspondence.....	47

Chapter 1

INTRODUCTION AND MOTIVATION

With growing number of internet users around the world, the attack plane for malicious attackers is getting bigger and it is a constant challenge for the security community to device methods to foil the various types of attacks used to steal private and confidential information from people. Today, it is widely recognized that the security and privacy of a system is not entirely dependent on technology. Consideration of human factors is an important part of any security system design. This idea revolves around the fact that a system can have users of different backgrounds and technical prowess and thus the ‘usability’ of the system is a high priority requirement today. Usability is defined as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” (ISO 9241). The ease with which a user learns how to use a system, the rate at which she makes errors while using it, the ease with which she can retain the knowledge attained by previous usage of it, the satisfaction of using it, are some factors that constitute the usability of a system Garfinkel and Lipford (2014). Usability has become a requirement that needs attention throughout the development cycle of a system rather than at the end. In developing security systems too, usable privacy and security (UPS) has become a high priority requirement. To make security effective, security mechanisms should be usable by all types of people using a system, ranging from non-technical users to experts. Without usable security, the user is more likely to become vulnerable and get compromised by an attacker.

To protect a user from possible threats, many security mechanisms have been developed. Browser warning is one such mechanism, which is incorporated in today’s

browsers to protect an internet user from threats like identity theft, theft of confidential information, exposure to malware etc. Organizations like Google and Phishtank maintain blacklists of reported phishing and malware websites. To protect a user from these websites, browsers use these lists and check every URL being requested by the user against them. For every URL that is found in these lists, the browser presents a warning message to the user about the potentially malicious website. The warnings give users the option to either stop from proceeding or ignore and bypass them. These warnings are intended to protect the users from malicious websites but studies have shown that these warnings are not very effective in stopping users from visiting potentially dangerous websites. Users have been observed to ignore browser warnings fairly regularly Dhamija *et al.* (2006) as heeding a warning involves spending time reading and understanding it, which deviates from their original task. Commonly, users get habituated to browser warnings and this too leads to users not taking the warnings seriously Bravo-Lillo *et al.* (2013), Bravo-Lillo *et al.* (2014). A well-known epigram to succinctly describe this observed behavior of users goes as follows: given a choice between dancing pigs and security, users will pick dancing pigs every time McGraw *et al.* (1999)m. Thus, there is a need to make users understand warnings better and help them heed them.

1.1 Motivation

With browser warnings, including phishing warnings, it is desired that every time a warning is displayed to an internet user who is trying to visit a certain website, the user heeds the warning message and refrains from visiting the potentially malicious website. Exception to this are cases where the user is working under a controlled and safe environment, or the requested website has been incorrectly tagged as a malicious website and the user is absolutely certain about it. But current browser warnings

are not able to achieve this near 100% success rate. At an average, about 20% of the Internet users that encounter a phishing warning message, are found to click-through to the spoof website Symantec (2013). Other warning messages face similar rates of ‘ignorance’, which will be discussed in the next chapter. Thus, as mentioned earlier, there is a need to design better warning messages that can improve the users understanding of the warning and the threats associated, and also the users adherence to the warnings.

1.2 Thesis

This thesis work focuses on phishing attack warnings and their effectiveness. The goal of this work is to understand how users’ reactions to phishing warnings can be improved so that they are more likely to make safe decisions. An important factor that impacts a user’s reaction to a warning message is its reliability. If a user does not trust the accuracy or seriousness of a warning message, she is more likely to ignore it, rather than heed it. So part of this study is to develop a warning message mechanism that users can trust. The new mechanism ensures that a user is warned of a security threat not only when she is trying to visit a potentially dangerous website but also when she has visited the page and the page is still loaded in the browser. The warning is displayed unless the user chooses to discard the warning permanently. There are multiple stages of the warning message. In the first three stages, the warning tries to keep the user away from a fake website by utilizing the information of the target website. The final stage comes into picture when the user ignores the earlier stages and moves on visit the phishing website. In this stage, the warning tries to stop the user from entering confidential information into this website. There is an associated assumption that the designed warning mechanism needs to be accepted as a genuine browser warning, like existing warnings displayed by browsers like Mozilla Firefox,

Google Chrome, Safari etc. The objective is to design a warning message that gives better user understanding and thus better user response, with the assumption that user recognizes the warning as a genuine browser warning. Ways of making a user recognize a warning message as a genuine browser warning and not a fake warning message, targeted at the user, is outside the scope of this study.

The remaining of the thesis is divided as follows: Chapter 2 discusses warning literature and previous work done in the development of warning messages, Chapter 3 talks more about using target information in a phishing warning message, its technical feasibility and usability issues, Chapter 4 discusses the experiment conducted and the results and chapter 5 concludes this report with suggestions for future work.

Chapter 2

RELATED WORK: WARNING LITERATURE

Web browsers today show warnings to users in the event of a possible attack. These warnings have been studied by various research groups. Studies have covered all types of browser warnings, like malicious website warning, phishing warnings, SSL/TLS warnings and their various aspects. Research work on warning studied the effectiveness of existing warnings Almuhimedi *et al.* (2014), Vance *et al.* (2014), Egelman *et al.* (2008), Akhawe and Felt (2013), the factors that lead to users ignoring these warnings Almuhimedi *et al.* (2014), Vance *et al.* (2014) and ways to improve these warnings Bravo-Lillo *et al.* (2011), Bravo-Lillo *et al.* (2013), Bravo-Lillo *et al.* (2014), Felt *et al.* (2014), Krol *et al.* (2012). Most studies have concluded that users are the weakest link of the entire security mechanism. Users often try to get past a security warning message, when encountering one, as it comes in the way of the task that is being carried out. Some of the literature and studies that discuss the various aspects of different browser warnings will be discussed here.

2.1 Warning System

A warning is any kind of message that cautions a person from a potentially undesirable situation. Designing an effective warning is a challenging task. Warning literature has identified various factors that are associated with the effectiveness of a warning message. Wogalter proposed various guidelines for an effective warning design Conzola and Wogalter (2001). Some of the important guidelines he identified are the use of short and familiar words, as little words as possible to convey the message, pictorial symbol for attracting attention, plain, familiar and non-fancy font for the

text, etc. He also proposed a conceptual model, *Communications Human Information Processing (C-HIP)* model that divide the warning procedure into the following phases:

- Attention Switch and Maintenance
- Comprehension
- Attitude Beliefs
- Motivation

If a warning message failed, this model became an effective way to determine the cause of the failure of the warning. In a separate study, Egelman et al. examined the effectiveness of active and passive browser warnings by analysing them using the C-HIP model Egelman *et al.* (2008). Their study looked at active warnings of Mozilla Firefox (FF) and Internet Explorer (IE) browsers and passive warnings of IE browser. With the help of this model, they were able to examine the various points of failure of the warnings and also helped them make a comparison study of those warning messages. Passive IE warning was outperformed in every component of the C-HIP model by the active FF warning. The active IE warning was found to outperform passive IE warnings in some of the C-HIP components but were not much different in some others. Overall, this study showed that passive browser warnings were more or less useless, when it came to warning users of any potential threat, and the C-HIP model helped in analysing the weaknesses of these warnings.

2.2 Browser Warning Effectiveness

Various studies have focused on determining the effectiveness of browser warnings. Most studies conclude that internet users mostly ignore browser warnings for a variety

of reasons, some of which will be discussed in this chapter later. Egelman et al. conducted a user study with 60 participants in a lab, where they found that about a fifth of the participants clicked through an active warning message whereas 87% of the participants clicked through passive warnings Egelman *et al.* (2008). In another study conducted to understand the users behavior to browser warning messages, Vance et al. found that users mostly ignored warning messages unless they fell victim to some attack and lost something, like data getting deleted from the hard drive by a virus that got installed after some warning was ignored at the first place Vance *et al.* (2014). It was only after facing such a situation, users were found to have heeded more warning messages. In a 2013 study, Akhawe et al. found that contrary to popular belief, internet users today are much more aware and heedful of warning messages, viz. malicious warnings, phishing warnings as well as SSL warnings Akhawe and Felt (2013). They conducted a study of over 4.5 million users of Google Chrome and Mozilla Firefox users via telemetry and found that with new browsers and new and more refined warning messages, the issue of users ignorance of warning messages is a matter of the past. From the browsing data collected from all these users, the authors calculated the click through rate (CTR, percentage of users that ignore a warning) for different types of warnings from both Firefox and Chrome browsers and found that among Mozilla Firefox users, 7.2% users ignored malware warnings, 9.1% ignored phishing warnings and 33% ignored SSL warnings. Among Google Chrome warnings, these numbers were 23.2%, 18% and 70.2% respectively. These numbers suggest that other than the SSL warnings, browser warnings are not as much ignored by users now as they were earlier. In addition to that, the study found that more technically experienced users tend to be more ignorant of the warnings than other users. However, the causes of this was not studied and has been left open for future research. In 2014, Almuhiemedi et al. conducted a study of malicious warning messages in browsers in

which they collected about 4 million warning impressions via telemetry. The data collected by them showed that the CTR for websites that have not been visited before range between 9.3% and 17.2%, whereas the CTR for websites that have been visited before range from 15.6% to 54.3%. The limitation here was that browsing history could be cleared and thus the numbers above can be inaccurate. However, it only meant that the CTR calculated could be a lower bound on the actual CTR.

These studies indicate that browser warnings in general are not even close to being 100% accurate. For various reasons, users choose to ignore browser warnings and click through them. The following section talks about some of these reasons.

2.3 Warning Ignorance Factors

Various studies have looked into reasons behind users' ignorance of browser warning messages. Of the various reasons found, the most common included users inability to understand the message due to its esoteric nature Vance *et al.* (2014), trust in 'protective technologies' like anti-virus software, Linux operating system, Apple Macs etc. and confusion between different types of warning messages Almuhimedi *et al.* (2014). Also, frequent exposure to warning messages causes habituation and users slowly start ignoring warning messages slowly, over time Bravo-Lillo *et al.* (2013), Bravo-Lillo *et al.* (2014).

In the study conducted by Almuhimedi et al. Almuhimedi *et al.* (2014), it was observed that users ignored browser warnings due to the following primary reasons:

- Trust in 'protective technologies' like anti-virus software, Linux operating systems, Apple Macs etc.
- Confusion of malware warnings with SSL warnings, which many users believe to be mostly false positives.

- Inability to identify if a link from a referrer is actually of a valid website or a fake website, in spite of the warning message showing the URL of the destination.

In addition, it was found that some people trusted reputed websites more than the warnings and thus ignored the warnings, whereas some people would not discard the warnings outright and spend considerable amount of time to understand the warnings if they were shown for highly reputed destination. In the study conducted by Vance et al., it was observed that most participants found the technical language used in the warnings too difficult to understand and some also reported that the warnings lacked clarity and were lacking detailed information Vance *et al.* (2014). It was observed that once users are afraid of possible consequences of ignoring warning messages, their responses improved and they started heeding the warnings more. These observations highlight the importance of keeping warnings simple and easy to understand. The guidelines proposed by Wogalter can help design a very effective warning message Conzola and Wogalter (2001).

2.4 Browser Warning Improvement

There has been a great focus on improving warning message effectiveness. Several studies have focused on new warning designs in an attempt to understand the factors that make a good, effective warning message. Some of these have focused on browser warnings and some others have focused on operating system level security dialogs. Some modified warning designs have been found to be more effective than existing warnings in various aspects, while some were found to be not as effective. Bravo-Lillo et al. conducted a study in which they looked into the relation between a warning design, the users understanding of the warning, users motivation of choosing the safest option for a particular warning and the users final reaction to the warning Bravo-Lillo *et al.* (2011). They studied four existing computer security warnings and

two warnings redesigned for each warning type, categorized under low and high risk situations. Their design was found to be better than existing warnings at improving understanding and motivation in some of the cases, but not all. However, there seemed to be no significant difference in user reaction to warnings in low and high risk scenarios, where the expected reactions were different. In a follow up study, Bravo-Lillo et al. designed a set of modified warnings, termed as ‘attractors’, designed to ‘attract’ user attention to a text field within the warning, and improve their responses Bravo-Lillo *et al.* (2013). Different types of attractors were designed to capture the users attention by requiring user action to get past them. There were five types of ‘inhibitive’ warnings, which were designed to prevent the user from making a potentially hazardous choice until the user took some expected action, depending on the type of the attractor used, or some period of time elapsed after the attractor became visible. Some of the user actions expected included moving the mouse pointer over some part of the warning text, typing in a text box inside the attractor, clicking an OK button over some text in the attractor etc. Three ‘static’ attractors were also designed to attract user attention without the requirement of inhibitive user action. The authors conducted three between-subjects experiments to test the attractors effectiveness with 573 Mechanical Turk workers. The study showed improvement in user reaction to certain attractors, and some insignificant difference in reaction to certain attractors. However, it was shown that with change in design, it was possible to get significant improvement in user reaction to warnings. Following up to this study, Bravo-Lillo et al. conducted another study to improve upon some of the attractors, discussed earlier, to make them more effective in high as well as low habituation scenarios Bravo-Lillo *et al.* (2014). Their experiments in both studies showed that the warnings which were interactive, i.e. needed user interaction with the message text, were very effective in improving user reactions and adherence. However,

this resulted in reduced usability in some of the warnings as users were required to perform actions like selecting text using the mouse, typing in an input box etc.

In another study, Krol et al. conduct a study with 120 participants who were asked to download a PDF file in their own laptop Krol *et al.* (2012). While downloading the file, the users encounter one of two malware warning messages. It was observed that out of 120 participants, 98 ignored the warning and went on to download the PDF. Out of the 22 people who heeded the warning message, 16 were females and 6 were males, which indicated that females could be more wary of security and privacy threats. With an eye tracking mechanism, it was observed that on an average, participants spent about 6 seconds on each warning message, which was enough for them to read the entire message. Based on this and their responses from a post experiment survey, it was found that 107 people were able to understand the warning message. It was also observed, from the survey, that participants with knowledge of computer security and more technical experience were more ignorant of the warning messages. This study gives a conclusive account of how users ignore browser warnings, in spite of understanding them and suggests that users need to be made to understand the seriousness of warning messages and not take them for granted. In a study focused on improving Google Chromes SSL warning effectiveness, Felt et al. incorporated Mozilla Firefox SSL warning design in the Chrome browser Felt *et al.* (2014). This study was based on the results from a previous study by Akhawe et al. Akhawe and Felt (2013) where it was found that Google Chrome SSL warning was ignored by over 70% of the users who encountered them. In comparison, Firefox SSL warning was ignored by just 33% of the users who encountered them. On redesigning the Chrome warning to a 'Mock Firefox' design, the CTR was seen to go down to 56%, which is still a very high number of users ignoring the SSL warning. In an attempt to make SSL warning more effective, Felt et al. worked on redesigning the Chrome SSL

warning by following guidelines found in warning literature. They focused on simple, easy to understand text, without the use of too much technical jargon. They aimed at improving the users comprehension as well as the understanding of the warning message. By collecting feedback from over 7500 participants over Google Consumer Survey (GCS), they found that it was possible to improve users adherence to the warning, i.e. they were able to stop up to 58.3% users from ignoring the warning. However, the warning was not able to improve users understanding of the threat associated with it.

A lot of work has been done to improve the effectiveness of browser warnings and a lot of success has been achieved. But it has become more challenging to make warning messages even more effective, in terms of comprehension as well as understanding. It has been seen that choosing simple and easy to understand message text and an interactive design can be some of the factors that can make warning messages more effective.

2.5 Phishing Warnings

Surprisingly, warning study has not been focused much on phishing warnings. Although there have been a few studies that focused on how users behave to phishing warning messages, Egelman *et al.* (2008), Akhawe and Felt (2013) there has been no study looking into possible ways of improving phishing warning messages. As it was discussed earlier, Egelman *et al.* conducted a comparison study of the effectiveness of active vs. passive phishing warning messages and found that active messages were much more effective than passive warning messages in protecting a user from potentially spoof websites. In their large scale study to calculate the CTR of users of Mozilla Firefox and Google Chrome browsers, Akhawe *et al.* found that about 9% of Firefox users and 18% of Chrome users ignore the phishing warning to proceed to the

spoof website and potentially get confidential information stolen. However, beyond this, there has been no focus on making these warnings more effective. Browser developers have time and again tried to improve phishing warnings, but this has been restricted to the look and feel of the warning only. Modifications have been made to background colors, the warning text, warning images etc. that are used in the warning. But beyond this, there has been no significant design changes in phishing warning messages, since active warnings replaced passive warnings.

PHISHING ATTACK: USING TARGET WEBSITE INFORMATION TO IMPROVE WARNING DESIGN

Earlier it was discussed that to a certain extent, phishing warning messages are ineffective, as they are either unable to stop users from ‘clicking through’ to a fake website, or they are unable to improve the users comprehension of the warning. The comparison of active warnings and passive warnings show that introduction of active warnings brought down the CTR by a huge margin Egelman *et al.* (2008). To bring down this rate further is an even more challenging task. To bring down the CTR to close to zero, there must be significant improvement in the warning design. But other than the introduction of active warnings, there has been no revolutionary change in phishing warning designs. Most of the changes that phishing warnings have undergone after active warnings were introduced are with respect to the look and feel of the warning, like changes pertaining to the warning text, background colors, images used etc. Figures 3.1 and 3.2 show the current phishing warnings for Mozilla Firefox and Google Chrome browsers. In this study, a novel warning mechanism for phishing attacks is proposed, which is designed to be more effective and intended to help users make a safe decision when they are made target to an unexpected phishing attack.

In a phishing attack, a malicious website always masquerades as a genuine website, which is the target website, to fool the unsuspecting user. Typically, when the browser displays a warning message for a potential phishing attack, it informs the user that the requested URL has been reported as a potential fake website, designed to fool unsuspecting internet users. Along with that, there are buttons provided to give users options of going back or ignoring the message and clicking through. Many times, this

does not prove to be good enough as some users eventually click through the warning.

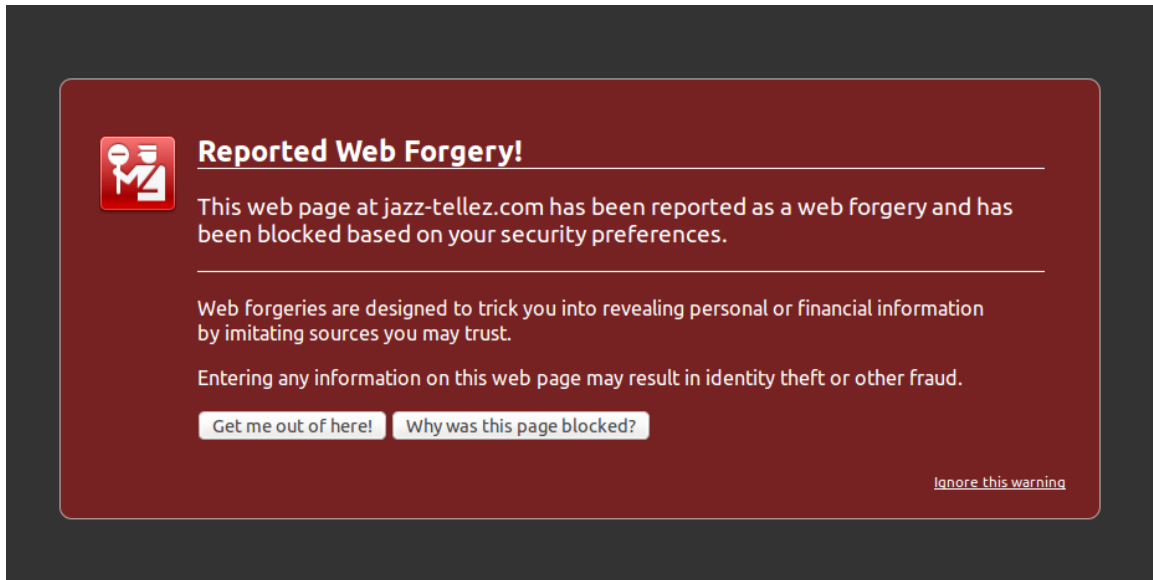


Figure 3.1: Mozilla Firefox 34 Phishing Warning Page

In this proposed design, a multi staged, interactive warning mechanism is proposed, which utilizes the information of the target website to increase the users trust of the warning which can improve the warning effectiveness. But before discussing more details about that, it is important to discuss some technical and usability related issues of using the target website information in a browser warning.

3.1 Technical Feasibility

It was discussed earlier how web browsers check every requested URL against certain blacklists of phishing websites that are maintained by organizations like Google and Phishtank, and decide if a phishing warning needs to be displayed to the user. Google maintains its list of blacklisted websites in a hashed format and the only way

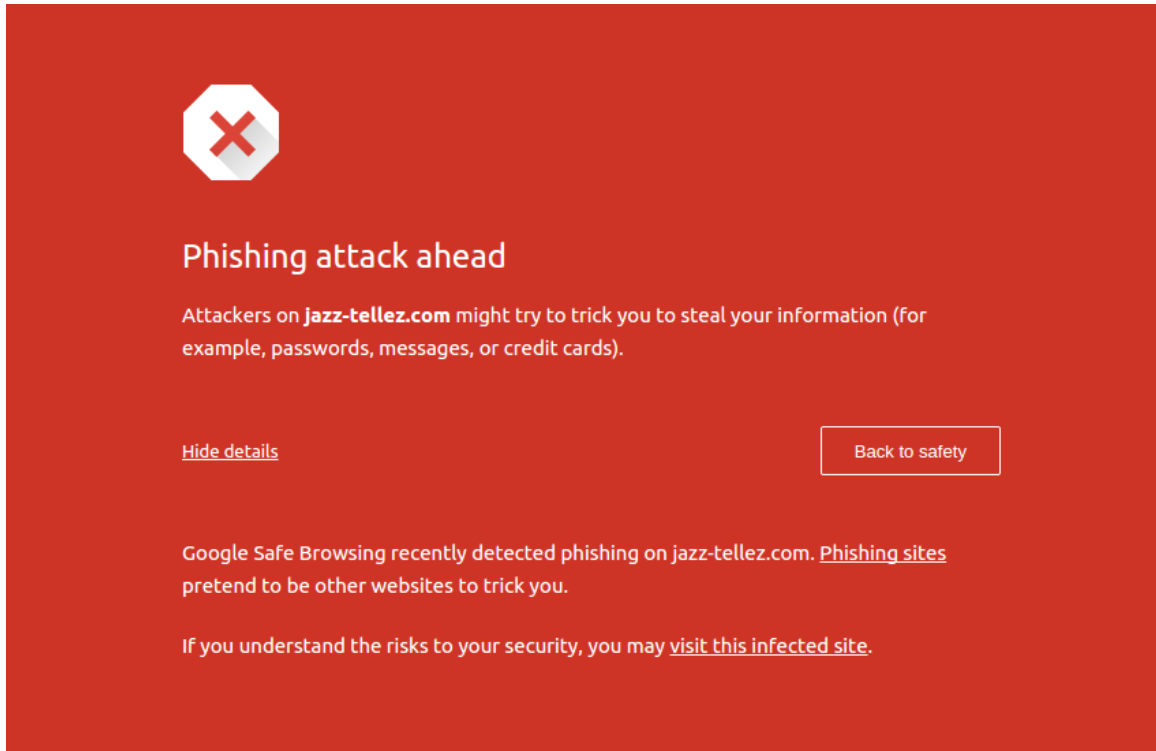


Figure 3.2: Google Chrome 39 Phishing Warning Page

that list could be used was by using its API to check if a particular URL is entered in their list. However, Phishtank maintains a detailed list in plaintext, with the phishing URLs and their corresponding target websites, among other information. This data is available on their website and updated every hour. The target website information can be very instrumental in designing a very effective warning mechanism against phishing warning. This is discussed further below.

3.1.1 Implementation Overview

The goal of this study was to design a phishing warning mechanism, built around the information of the target website of a reported phishing URL. Thus this work does not focus on the detection of a phishing URL but instead relies on external source

for the list of previously reported phishing websites. This study however throws light on certain ways to detect the target information of certain phishing URLs. It is discussed further in section 3.1.2. In the design proposed in this study, Phishtank's data was found to be the most suitable, as they not only provided the reported URLs in plaintext, but they also maintained the corresponding target names for a portion of those URLs. This study introduces an add-on for the Mozilla Firefox browser that can intercept every URL requested by a user and then check this against the blacklist provided by Phishtank. This add-on performs the task of detecting a target website and based on that detection, display the new warning message. Here the mechanism completely relies on Phishtank's data in identifying a phishing URL. If a URL exists in the list, it is considered a web forgery and the add-on proceeds to display a warning message for the requested URL. From the Phishtank data, the add-on reads the target website information for the URL detected as a web forgery and uses that name in the warning message accordingly. If the URL is not found in the list, it is considered harmless and no warning message is displayed.

3.1.2 Technical Details

Phishtank provides a database of thousands of reported phishing URLs, with their corresponding target names. However, it was found in the study that in the Phishtank data, less than 20% of the reported URLs are mapped to respective (known) targets. This number changes minutely as the database gets updated every hour. Table 3.1 gives a breakdown of the URLs whose targets have been verified by Phishtank as of January, 2015. The URLs that are not mapped to any known targets are categorized as 'Others'. It is, therefore, highly probable that many URLs requested by a user will not be mapped to any known target name. For such scenarios, a few measures are proposed to detect the potential target name that the spoof website is mimicking.

Table 3.1: Top 10 Targets Identified by Phishtank

Target	Overall Percentage	Percentage in known targets
Paypal	7.7%	36.78%
Poste Italiane	3.125%	14.9%
AOL	2.44%	11.7%
eBay	1.2%	5.75%
Apple	0.75%	3.57%
Google	0.66%	3.16%
Bradesco	0.48%	2.29%
Allegro	0.41%	1.9%
Capitec Bank	0.31%	1.5%
Orkut, Yahoo	0.245%	1.17%

However, as mentioned earlier, target detection is not the primary objective of this study. The warning message has been designed to be compatible and usable with any target detection mechanism.

The following is a discussion on how the target website detection mechanism proposed in this study works. When a URL is detected as a phishing URL, the add-on searches for possible target names by looking at the data provided by Phishtank. If a target is not identified, then the add-on looks for words similar to the Phishtank identified target names in the URL of the phishing website. In this way, target identification was done for about 25% of the unidentified URLs. Remaining URLs did not contain any word that matched any of the targets identified by Phishtank. Table 3.2 gives a breakdown of the targets that were identified from the URL of the phishing links categorized under *Others* category. However, the accuracy of this could not be calculated due to two problems associated with phishing websites:

Table 3.2: Breakdown of Targets Identified by Add-on

Target	Percentage of Identified URL
Yahoo	19.42%
Google	18.28%
AOL	16%
Windows Live	15.43%
Alibaba	2.86%
Paypal	2.29%
Facebook	2%

- **Brief time frame of existence.**

Phishing websites are live for a very brief period of time. Once some victims fall prey to some attack, these websites are quickly taken down, or they are hosted under different domain names. 500 random URLs were picked from Phishtank's list of reported spoof websites and it was found that over 40% of the websites were no longer available. Table 3.3 gives a breakdown of the top 3 reasons of these URLs not working, which form about 92% of these URLs that do not work.

- **Multiple targets in the same URL**

150 URLs were picked randomly from the phishtank data. Out of these URLs, 89 URLs were found to still work. In 27 of those 89 cases, the URLs targeted multiple websites. The total number of targets from these 89 links was 174. Table 3.4 gives a break down of the some of the major websites these URLs targeted. Because of URLs having multiple targets, the target distribution cannot be ascertained without manually opening all the URLs in the Phishtank

Table 3.3: Unavailable URLs Breakdown

Error Message	Percentage
Web page no longer available	44.26%
Web page did not load	32.79%
Error 404	14.75%

Table 3.4: Breakdown of Targets with Multi Target URLs

Target	Percentage of Identified URL
Google	49.15%
Paypal	10.61%
Guildwars 2	9.67%
Yahoo	7.2%
Apple	6.6%
AOL	5.91%
Dropbox	4.43%

list.

Software security product vendor Kaspersky recently released a report on phishing attacks detected by their anti-phishing component, in Q2, 2014 Lab (2014). As per their record, some of the top phishing targets are shown in Table 3.5. These websites come under Global Portal and Social network sites portal. The report did not provide any detail on websites dealing with online financial transactions.

As mentioned earlier, about 40% of the sample of 500 random URLs were found to be inactive. However, it was also noticed that of the URLs that still worked, almost 100% of them had the name of the target website written on the browser tab title. So the add-on was developed to keep this into account and read the browser tab title

Table 3.5: Top Phishing Targets by Kaspersky

Target	Percentage of Identified URL
Yahoo	30.96%
Google	8.68%
Facebook	8.1%
Live.com	3.72%
Odnoklassniki	2.75%

to get the possible name of the target website. This target information can then be used in the new warning design as discussed in the next section.

3.1.3 Warning Design

The purpose of the designed add-on, that was discussed earlier, is only to intercept every requested URL and check if it has been reported as a phishing website, and if it is, then detect the corresponding target website name. This study proposes a novel warning message design that uses that target website name to warn the users of a potential phishing attack. The design keeps into consideration the fact that the back-end target detection mechanism may or may not be accurate, as it may be bypassed by an attacker, as a result of which, the target detection may be inaccurate. If the detection of the target website is correct, the warning message provides the user the option to navigate to the real and safe target website. Figure 3.3 shows the first warning page that is displayed to the user once the target is identified. The warning message eliminates the possibility of a user proceeding to a spoof website if the target is detected correctly. In case of an incorrect detection of the target website, the warning provides the user a list of three possible targets to choose from. If the actual target matches one of those websites suggested, the warning provides way to directly

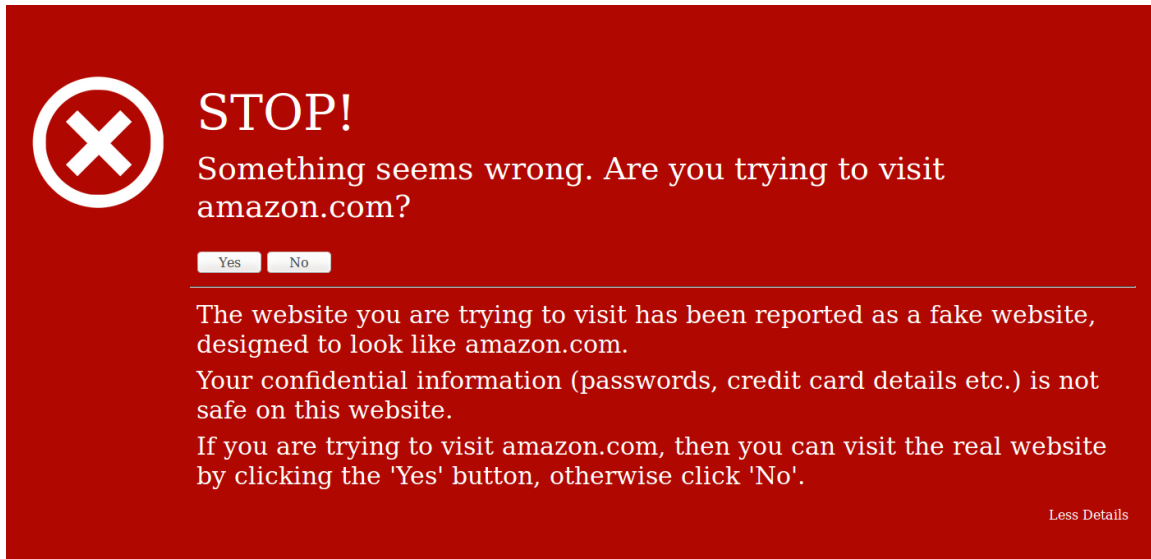


Figure 3.3: Proposed Warning Message - Stage 1

navigate to the corresponding real and safe target. Figure 3.4 shows this scenario. If there is still no match, the warning page gives the user the option to enter as an input, the name of the target website, and redirects to a Google search page of the target name provided by the user. The user can decide to choose the real and safe website from the Google Search page. In this page, the user is also given the option to ignore the input box and proceed to the potentially harmful website if the user does not want to waste time while typing into the input box. This scenario can be seen in Figure 3.5. If the user chooses to proceed to the potentially harmful website, the add-on still keeps track of the users activity on that website. Here the final stage of the warning mechanism comes into action. After ignoring the warning message and proceeding to the potentially harmful website, if the user tries to click on any text box, to enter either user id or password, the add-on will display a small floating passive warning message, to discourage the user from entering any confidential information into the website. The user is given the choice to either leave the page or ignore the warning

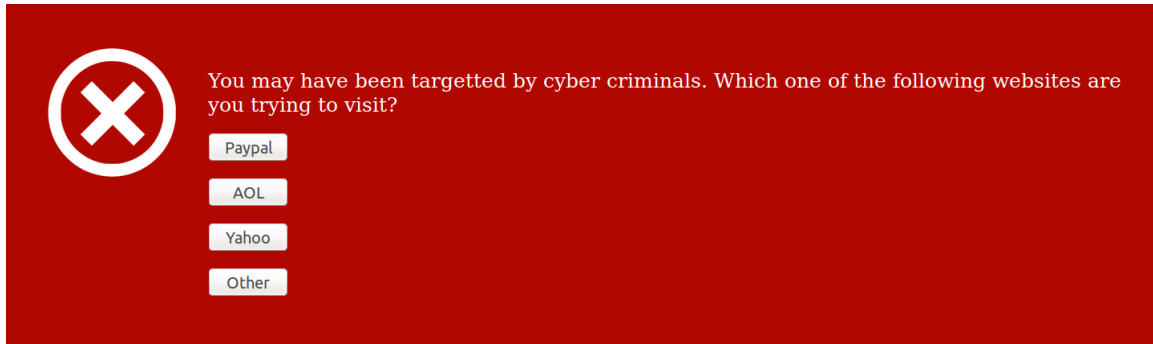


Figure 3.4: Proposed Warning Message - Stage 2

permanently (Figure 3.6). On leaving the page, the tab is closed and the user is protected from a potential theft of confidential information. If the user chooses the latter, then the warning is no more displayed for that URL and the user is on her own after that

3.1.4 Limitation

Our warning design suffers from a couple of limitations. First, the navigation to the real, safe websites of a possible target website is done only to the home page of the website, even if the user was trying to go to a specific page in the website. This is not a major limitation as it is possible that the user was tricked into clicking a fake link to get to some fake page, and a corresponding web-page in the real, safe website may not even exist. The second limitation is that if the user goes on to enter the name of the potential target website name by typing into the input box, then the warning page redirects to a google search page, instead of the website of the organization provided by the user. Future work can focus on addressing these limitations.

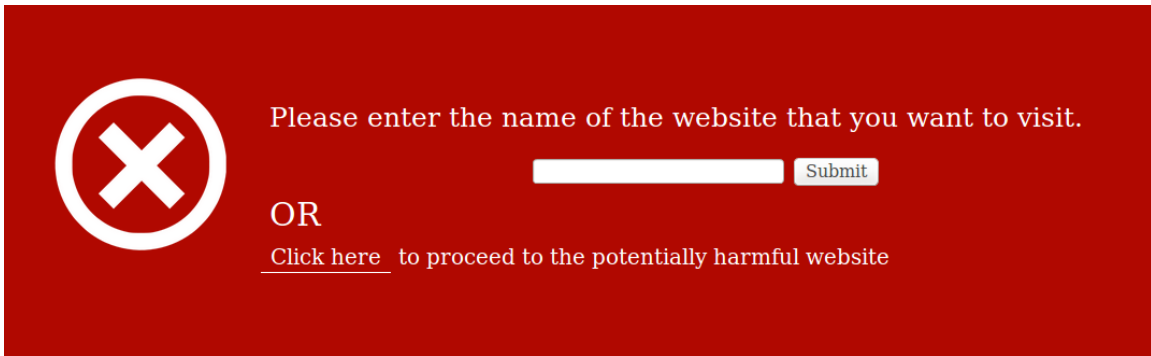


Figure 3.5: Proposed Warning Message - Stage 3

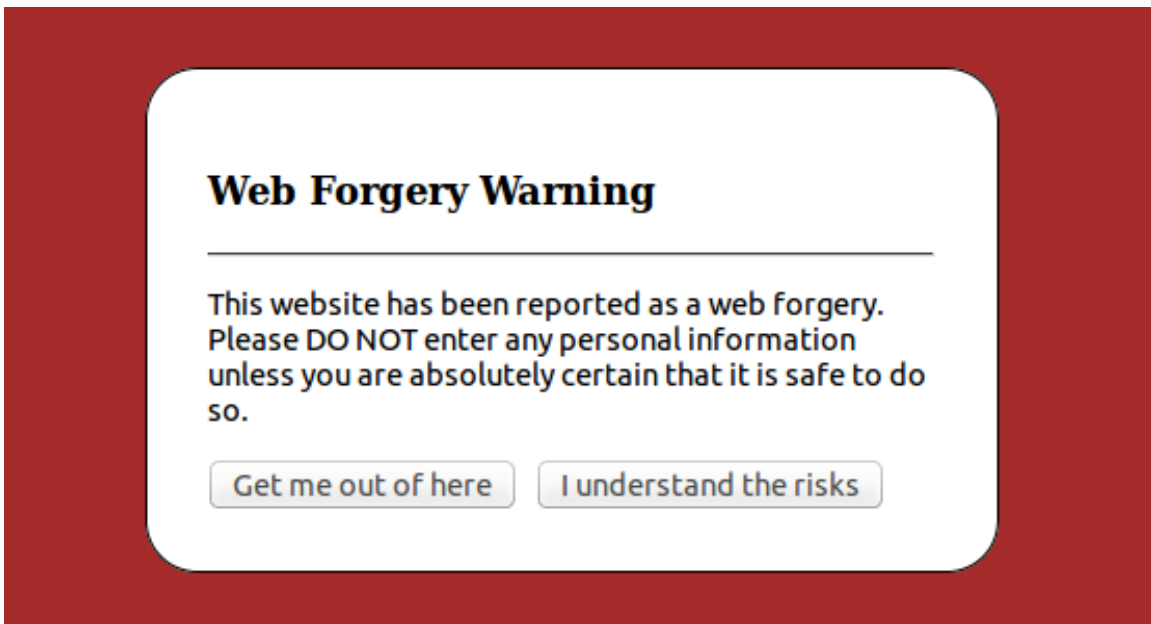


Figure 3.6: Proposed Warning Message - Stage 4

3.2 Usability Issues

While designing a warning message, it is very important to consider the factor of usability. A warning message is expected to be easy to read, easy to understand and should be able to convey the threat associated effectively. In this study, various factors

that go into making the warning message an effective warning while maintaining high usability of the warning are considered.

3.2.1 *Warning Design*

As it was discussed earlier, a target website information is used to inform the user that she is under a potential phishing attack. On identifying a target, the warning provides explanation in simple text about the scenario and asks the user to click on either ‘Yes or ‘No, depending on the accuracy of the target website detection. For a correct detection of the target information, the warning does not provide any option to the user to ignore the warning and carry on with the spoof website. Here, the warning does not decide if the target detection was accurate. It is decided by the user. If the target displayed by the warning message matches the users requested website, then the target detection is considered correct. If the detection is incorrect, the user can click on *No* to get further options. This takes the warning to the next stage, where the user is provided three potential target names that the user might have requested. In choosing the target names in this study, the top three phishing targets identified by Phishtank other than the target originally identified, are used. Here, machine learning algorithms can be used to get better target options. In this study, as discussed earlier, the focus is not on the accuracy of detecting the target warning, but on using the target information in a warning message. In this stage, the options provided to the user are very simple, if the target identification was done correctly, then click on the corresponding button, otherwise click on the *Other* button. The latter option will take the warning to the third stage where the user is given the option to enter the website name in a search box or avoid that and click through to the potentially spoof website. It is understood that asking a user to type in an input box in order to get to a website is not desirable from the usability standpoint, but

to keep a user safe from a phishing website, this is a trade-off. Also, this completely depends on the accuracy of the target detection technology. If target detection is accurate, then the third stage of the warning can be totally avoided. The warning design has been built to work for any type of target detection mechanism.

3.2.2 *User Understanding*

When designing a warning message, one of the most important considerations is the user's understanding of it. It determines how a user is likely to react to the warning on encountering it habitually. If a user does not understand a warning message, then there is a potential risk of her getting habituated to ignoring the message and proceeding with her potentially harmful task. The warning proposed in this study has been designed considering a user's understanding with respect to her desired understanding in mind. Simple, short and easy to understand sentences were used to make user understand the threat associated with the warning message. It is extremely important that a user's understanding of the scenario is correct at the first stage of the warning, as the rest of the stages rely on that. Table 3.6 gives an account of the user's desired response to the first stage of the proposed warning message based on the her perception of the target website vs the real scenario, in the example shown by Figure 3.3. If the user was trying to visit `www.amazon.com`, then for any scenario, her desired action is click on *Yes*, as the browser will then redirect to the real and safe website. If the user was trying to visit some other website, but the target detection mechanism mistakenly detected the target as Amazon, then the user's desired action is *No* for all scenarios. Here, the first two of the 'desired action: click No' scenarios are harmful situations that arise due to failure of target detecting technology. The third 'Click No' situation is harmless, as a harmless website was detected as a phishing website.

Table 3.6: User’s Desired Action Based on Understanding of Warning

	User’s target perception	
Real Scenario	amazon.com	Not amazon.com
Fake website pretending to be amazon.com	Click Yes	Click No
Fake, phishing website not pretending to be amazon.com	Click Yes	Click No
Harmless website, incorrectly detected as amazon.com	Click Yes	Click No

3.2.3 User Trust

Another extremely important factor determining the success or failure of a warning message in keeping a user safe from a phishing website is her trust on the warning message. Once the user is able to correctly understand the warning message, her adherence to it depends on her trust on the warning. The proposed warning mechanism is designed to gain the user’s trust by including the target website information. It is expected that seeing the desired destination website’s name in the warning will make the user treat it not as a generalized warning but a specific warning, meant only for her. This might help in improved trust on the warning mechanism and thus to the adherence of the warning. With habitual encounter of this warning too, the user is expected to display adherence to the warning.

Chapter 4

STUDY AND RESULTS

To understand the effectiveness of the proposed warning message, it has to be run through several tests. People from different backgrounds, different age-groups and different levels of computer security knowledge should be subjected to this warning message and based on their responses and feedback, the warning's effectiveness can be ascertained. As a first step to this process, an online survey was conducted among students of an undergraduate course in the Computer Science department at Arizona State University (ASU). This was approved by the Institutional Review Board (IRB) at ASU with exempt status. Out of the 167 students that were contacted for the study, 27 students participated. In the survey, some user information, like age, gender, browser user, number of hours spent on internet every week etc., were collected first. Of the 27 participants, 5 participants were female and 22 were male. 18 participants (66.67%) users were in the age group 18 - 24 and 6 participants (22.22%) were in the age group 25 - 29. Then the users' feedback on questions related to the proposed warning message and existing warning messages from Mozilla Firefox and Google Chrome was collected. The questions have been listed in Appendix A.

4.1 Hypothesis

The effectiveness of the proposed warning is tested against the phishing warnings of Mozilla Firefox and Google Chrome browser. Three parameters were considered for this: Understanding of the warning, simplicity of the language used and Confidence in landing to a safe website from the warning. There were six hypotheses tests which are as follows:

- Hypothesis 1 (H1): The proposed warning will be easier to understand than the Mozilla Firefox warning.
- Hypothesis 2 (H2): The proposed warning will be easier to understand than the Google Chrome warning.
- Hypothesis 3 (H3): The language used in the proposed warning will be simpler than that used in the Mozilla Firefox warning.
- Hypothesis 4 (H4): The language used in the proposed warning will be simpler than that used in the Google Chrome warning.
- Hypothesis 5 (H5): The proposed warning will generate more confidence in the users of landing in a safe website, than the Mozilla Firefox warning.
- Hypothesis 6 (H6): The proposed warning will generate more confidence in the users of landing in a safe website, than the Mozilla Firefox warning.

All six hypotheses involve testing the first stage of the proposed warning message as the effectiveness of the warning message depended heavily on the user's correct response in the first stage of the warning.

4.2 Results

The responses received from the sample of 27 students did not validate any of the aforementioned six hypotheses. For all six hypotheses, the null hypothesis could not be rejected at 5% significance level. The following were the p-values obtained by conducting a two sample t test for all the six cases:

- H1: $p = 0.8769$
- H2: $p = 0.9777$

- H3: $p = 0.3666$
- H4: $p = 0.751$
- H5: $p = 0.8379$
- H6: $p = 0.945$

Thus it was not possible to reject any of the null hypotheses that stated that the proposed warning would be equally effective as the phishing warnings of Mozilla Firefox or Google Chrome in terms of understanding, simplicity or confidence in landing the user to a safe page.

From the responses of the user, it was observed that when the first page of the warning message was displayed to the participants, 19 users (70.37%) indicated that they would choose the option ‘Close tab by clicking on close tab symbol on the tab’ and 7 users (25.93%) responded that they would choose to click on ‘Yes’. The following were some of the feelings expressed by the participants who chose to close the tab, about the proposed warning design.

- Distrust
- Confusion in understanding

Most users failed to see the proposed warning as a genuine warning message or were confused by it’s functionality. The most common reason for this was found to be ‘lack of familiarity’, which was evident from some comments like:

- *‘ There are never warnings that say “click yes to go to the website you want to go to” unless they’re attacks that really take you to a bad site.’*
- *‘ I haven’t used a browser with this error message. Despite the convenience of clicking ‘Yes’ I am again more inclined to just close the tab and find out more.’*

- ‘ *If I have seen such a warning message before, I’d be more comfortable. But with so many fake warning messages and dialogues out there (as seen in some ransomware, in extreme cases), I’d be suspicious.*’

In scenario B, when the participants were displayed warning stage 2 and subsequently, stage 3, 3 participants (11.11%) reported that they would enter the correct website name and search for the real website, but 24 of them (88.89%) decided to close the tab. This time too, participants expressed distrust in the warning system. However, in this stage, along with unfamiliarity, participants also expressed concerns on seeing multiple warning levels, which was also something that they were not familiar with. This is seen in comments like the following:

- ‘ *Because it kept on changing, I would not feel confident in the error message.*’
- ‘ *Its gone to far nothing I have seen does this.*’
- ‘ *I have never seen a chain of warnings such as this. I would be concerned that there is a (remote) chance that these warnings are illegitimate and are themselves an attempt at something malicious. I would study the email more closely in another tab and probably end up deleting it.*’
- ‘ *To the average user, it might seem like an infection in itself. That’s ALOT of warning boxes to get through.*’

However, when participants were briefed about the warning design towards the end of the study, users showed improved confidence in the warning mechanism. On a scale from 1 to 10 on how safe the warning system would keep the users from a spoof website (Question 18, Section 2, Appendix A), the warning got an average of 7.07, with 12 out of 27 (44.44%) users giving it an 8 or above. Among female participants, this average

score was 7.6, whereas among male participants, the score was 6.95. Interestingly, there was a negative correlation between this score and the users' self reported score on the knowledge of computer security concepts. The Pearson's r value was calculated to be -0.208, i.e. participants who self reported themselves to be more knowledgeable on computer security concepts showed tendency of awarding a lower score to the warning design compared to participants with lower self reported score on computer security concepts. It was also observed that a total of 8 participants (29.63%) reported that the proposed warning design made them feel more secured compared to other warnings seen before (Question 15, Section 2, Appendix A). A similar correlation was observed here too, where Pearson's r value of -0.234 was observed between participants' self reported score on knowledge of computer security concepts and their confidence on the security provided by the warning design. Future work can look deeper into this correlation and find out if the warning design can evoke more confidence in internet users with lesser computer security knowledge.

4.3 Limitations

The user study suffers from the limitation of any survey based user study. It cannot be accurately stated that users, in real life scenario, would react the same way when encountering these warnings as standard browser warnings.

CONCLUSION AND FUTURE WORK

Phishing is a continuously growing threat in the cyber world. With the rapid growth of the Internet user base, phishing attacks are getting more rampant. With a growing attack plane, cyber attackers are constantly targeting unsuspecting internet users, trying to steal their confidential information. Internet browsers provide a mechanism of warning users from a potential phishing attack, but these are not always effective. In spite of providing a warning message, browsers are sometimes unable to stop users from falling prey to these attacks.

5.1 Summary

This study proposes a novel phishing warning design that has not been introduced before. Using the target website information in case of a detected phishing attack can help improve users' reaction to phishing warnings. However, the user study conducted as a first step to test the effectiveness of this new warning design was not completely successful. The study wasn't able to elicit a positive response from the participants. The warning failed to evoke trust in the user, who seemed to consider the warning message itself to be a part of an attack. Lack of familiarity with this type of warning message also seemed to be a reason of users' lack of trust in the warning design, among other reasons. However, participants expressed improved confidence in the design once they were briefed about it, which improved their confidence in it. So, among the particular sample, even though the proposed warning design did not prove to be more effective than phishing warnings of Mozilla Firefox and Google Chrome browsers with respect to understanding of the warning message, simplicity of

language used and confidence in landing on a safe website, it seemed that once users got familiar with the warning design and trusted its genuineness, the warning could be an effective tool in safeguarding them from phishing attacks.

5.2 Future Work

The warning can be designed with more text explaining the results of clicking the ‘Yes’ and ‘No’ buttons in the first page. The warning design can be further tested with people with different background, age group and levels of computer security knowledge. Also, there can be tests involving habituation of users with this warning and checking its effectiveness over extended usage. Future study can involve subjecting users to this warning under real life scenario. Based on the results of this study, later studies can be focussed on making the user aware that the warning message is genuine in the first place, since the ultimate objective is testing the effectiveness of the warning with the user aware of its genuineness, and then studying their responses to the different stages of the warning. Another future work can involve studying different ways of utilizing the target information to develop a warning mechanism, which may differ from the one proposed in this study. One variation can be a warning with more explanation of the ‘Yes’ and ‘No’ buttons, as mentioned earlier.

REFERENCES

- Akhawe, D. and A. P. Felt, “Alice in warningland: A large-scale field study of browser security warning effectiveness.”, in “Usenix Security”, pp. 257–272 (2013).
- Almuhimedi, H., A. P. Felt, R. W. Reeder and S. Consolvo, “Your reputation precedes you: History, reputation, and the chrome malware warning”, in “Symposium on Usable Privacy and Security (SOUPS)”, (2014).
- Bravo-Lillo, C., L. Cranor, S. Komanduri, S. Schechter and M. Sleeper, “Harder to ignore?”, in “Symposium on Usable Privacy and Security (SOUPS)”, (2014).
- Bravo-Lillo, C., L. F. Cranor, J. Downs, S. Komanduri and M. Sleeper, “Improving computer security dialogs”, in “Human-Computer Interaction–INTERACT 2011”, pp. 18–35 (Springer, 2011).
- Bravo-Lillo, C., S. Komanduri, L. F. Cranor, R. W. Reeder, M. Sleeper, J. Downs and S. Schechter, “Your attention please: designing security-decision uis to make genuine risks harder to ignore”, in “Proceedings of the Ninth Symposium on Usable Privacy and Security”, p. 6 (ACM, 2013).
- Conzola, V. C. and M. S. Wogalter, “A communication–human information processing (c–hip) approach to warning effectiveness in the workplace”, *Journal of Risk Research* **4**, 4, 309–322 (2001).
- Dhamija, R., J. D. Tygar and M. Hearst, “Why phishing works”, in “Proceedings of the SIGCHI conference on Human Factors in computing systems”, pp. 581–590 (ACM, 2006).
- Egelman, S., L. F. Cranor and J. Hong, “You’ve been warned: an empirical study of the effectiveness of web browser phishing warnings”, in “Proceedings of the SIGCHI Conference on Human Factors in Computing Systems”, pp. 1065–1074 (ACM, 2008).
- Felt, A. P., R. W. Reeder, H. Almuhimedi and S. Consolvo, “Experimenting at scale with google chrome’s ssl warning”, in “Proceedings of the 32nd annual ACM conference on Human factors in computing systems”, pp. 2667–2670 (ACM, 2014).
- Garfinkel, S. and H. R. Lipford, “Usable security: History, themes, and challenges”, *Synthesis Lectures on Information Security, Privacy, and Trust* **5**, 2, 1–124 (2014).
- Krol, K., M. Moroz and M. A. Sasse, “Don’t work. can’t work? why it’s time to rethink security warnings”, in “risk and security of internet and systems (CRiSIS), 2012 7th International conference on”, pp. 1–8 (IEEE, 2012).
- Lab, K., “Spam and phishing in q2, 2014”, URL https://securelist.com/files/2014/08/Spam-report_Q2-2014_en.pdf (2014).
- McGraw, G., E. Felten and R. MacMichael, *Securing Java: getting down to business with mobile code* (Wiley Computer Pub., 1999).

Symantec, “Symantec blog”, URL <http://www.symantec.com/connect/blogs/don-t-ignore-warnings> (2013).

Vance, A., B. B. Anderson, C. B. Kirwan and D. Eargle, “Using measures of risk perception to predict information security behavior: Insights from electroencephalography (eeg)”, *Journal of the Association for Information Systems* **15**, 10, 679–722 (2014).

APPENDIX A
SURVEY QUESTIONS

Section 1:

1. Age Group?

- 18 - 24
- 25 - 29
- 30 - 34
- 35+

2. Gender?

- Female
- Male
- Decline to Answer

3. What is your current Major?

4. What is your highest level of education?

- High School/GED
- Bachelor's Degree
- Master's Degree

5. Which internet browser do you use?

6. How many hours in a week do you spend surfing the internet?

Section 2:

Scenario A: Consider the following scenario: You receive an email from 'Amazon' with a receipt from the purchase as the image below:

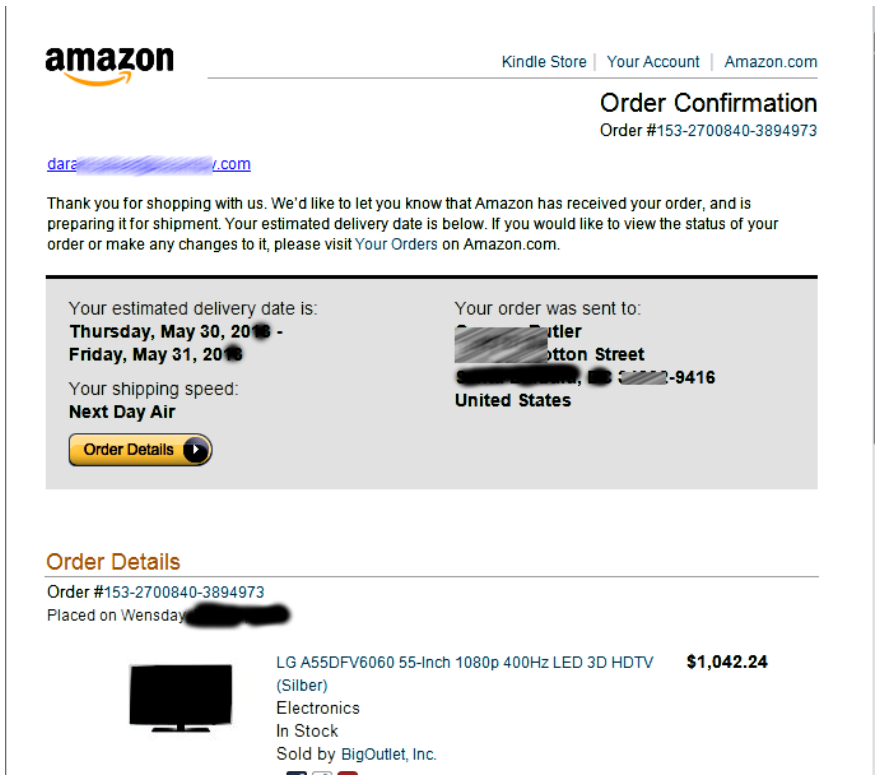


Figure A.1: Appendix- Amazon Receipt

On clicking Your Orders button, you are taken to a genuine warning page by the browser.

Warning 1: Assume that this is a genuine warning message displayed by the browser. Answer following questions accordingly.

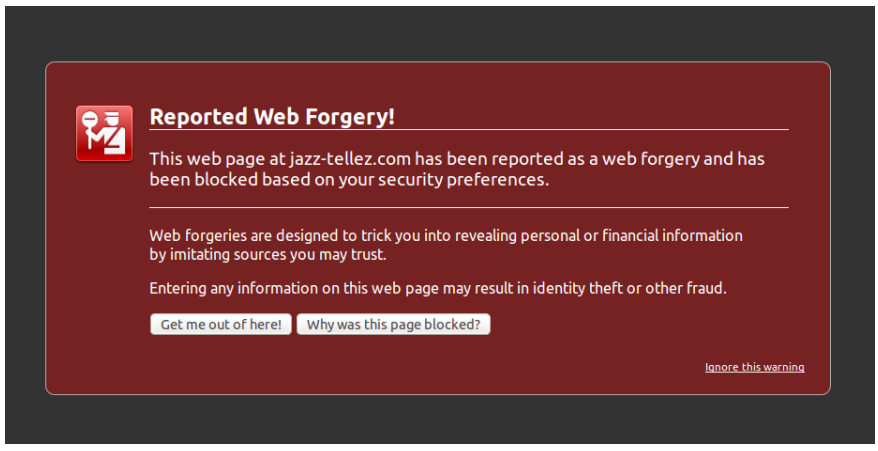


Figure A.2: Appendix- Warning 1

1. What will be your immediate action if you saw Warning 1?

- Click on ‘Why was this page blocked?’ button
- Click on ‘Get me out of here!’ button
- Click on ‘Ignore this warning’ button
- Close the tab by clicking on the close symbol of the tab
- Other

2. Explain in brief.

On a scale of 1 to 10, how would you rate the warning on the following rating?

3.a. Understanding of the warning (1: Least clear - 10: Most clear)

3.b. Simplicity of the language used (1: Least simple - 10: Most simple)

3.c. Confidence in landing on a safe website from the warning (1: Least confident - 10: Most confident)

Warning 2: For Scenario A, assume that this is a genuine warning message displayed by the browser. Answer following questions accordingly.

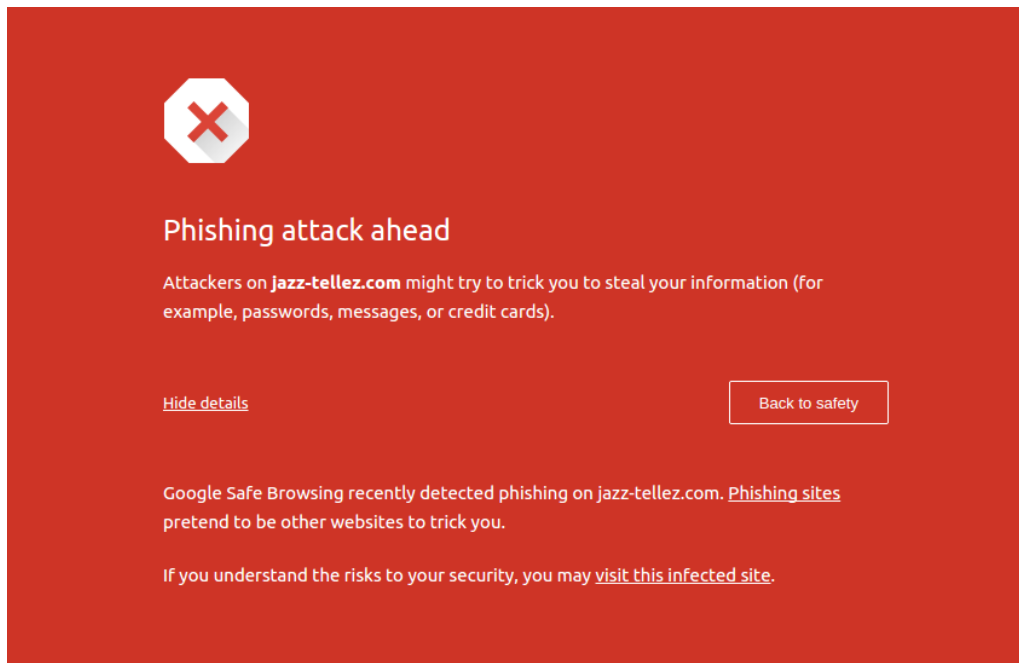


Figure A.3: Appendix- Warning 2

4. What will be your immediate action if you saw Warning 1?
- Click on the 'Back to Safety' button
 - Click on the 'visit this infected site' button
 - Close the tab by clicking on the close symbol on the tab
 - Other

5. Please explain in brief.

On a scale of 1 to 10, how would you rate the warning on the following rating?

6.a. Understanding of the warning (1: Least clear - 10: Most clear)

6.b. Simplicity of the language used (1: Least simple - 10: Most simple)

6.c. Confidence in landing on a safe website from the warning (1: Least confident - 10: Most confident)

Warning 3: For Scenario A, assume that this is a genuine warning message displayed by the browser. Answer following questions accordingly.(Please enlarge image if the text is unreadable.)

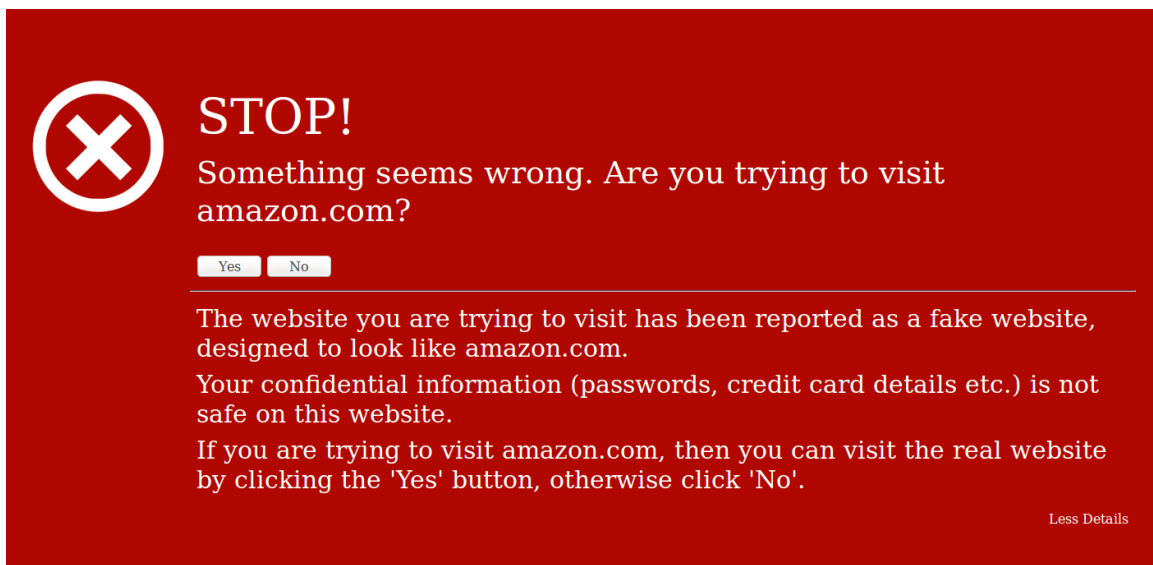


Figure A.4: Appendix- Warning 3

7. What will be your immediate action if you saw Warning 1?
- Click on 'Yes'
 - Click on 'No'
 - Close the tab by clicking on the close symbol on the tab
 - Other

8. Please explain in brief.

On a scale of 1 to 10, how would you rate the warning on the following rating?

9.a. Understanding of the warning (1: Least clear - 10: Most clear)

9.b. Simplicity of the language used (1: Least simple - 10: Most simple)

9.c. Confidence in landing on a safe website from the warning (1: Least confident - 10: Most confident)

Scenario B: You receive an email from Facebook with information of a new login system that is being implemented for more safety and security. To avail the new service, a login link is provided. On clicking the link, you see the following warning message.

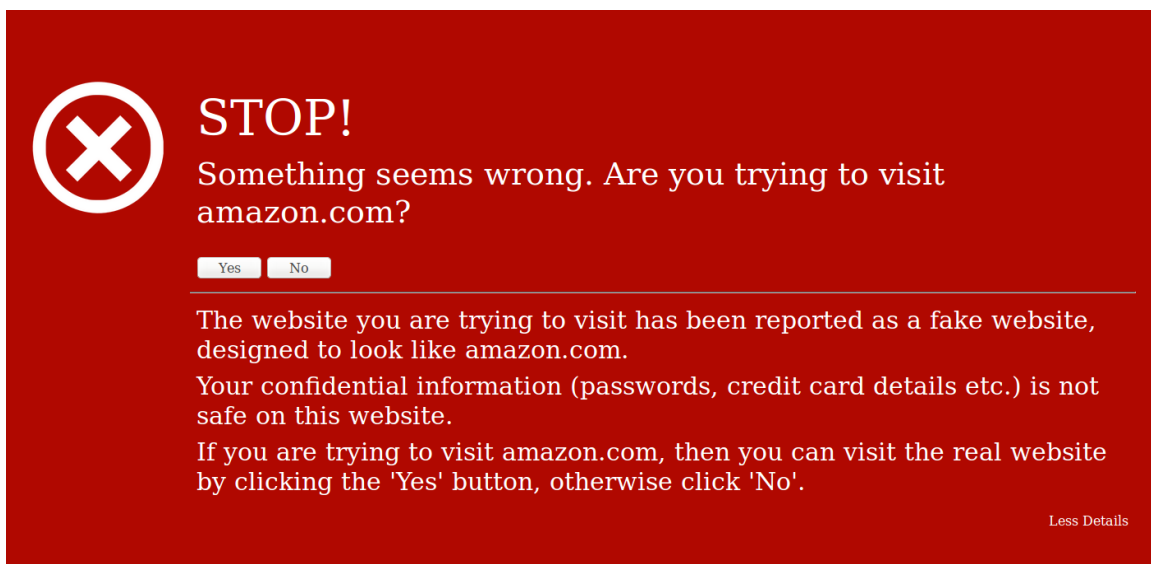


Figure A.5: Appendix- Warning 3

Since you weren't trying to visit amazon.com, you clicked on the No button. Then the warning page changes to the following warning (Warning 3.2):(Please enlarge image if the text is unreadable.)

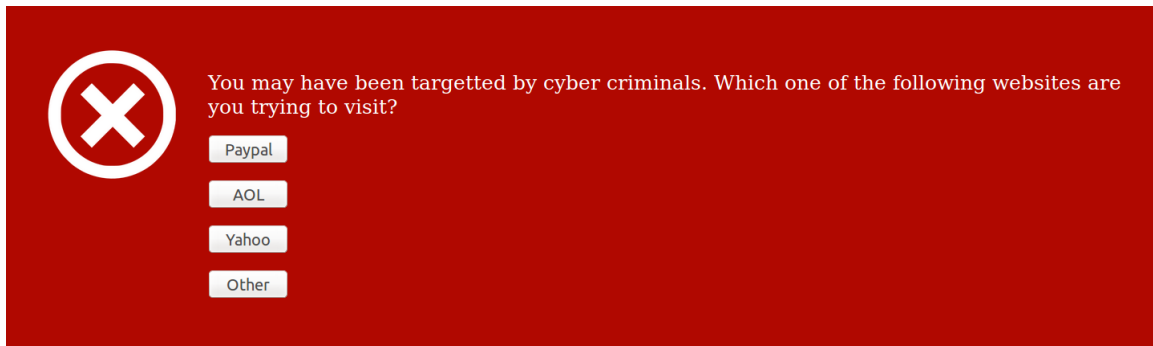


Figure A.6: Appendix- Warning 3_2

In Warning 3_2, shown above, assume that you click on Other since none of the other options take you to www.facebook.com. Then you see the following warning page (Warning 3_3):(Please enlarge image if the text is unreadable)

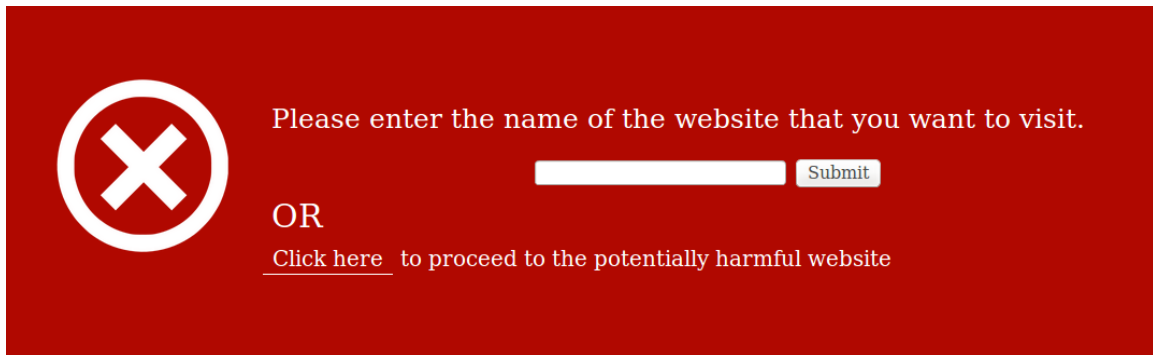


Figure A.7: Appendix- Warning 3_3

10. What will be your reaction when you see this change in the warning message?

- Enter the website name in the text box and click on 'Submit'
- Click on the 'Click here' button
- Close the tab by clicking on the close symbol on the tab
- Other

11. Briefly explain why.

If you chose option #2 for Question #10, please answer Question #12 and #13, otherwise, please go to the next page.

You clicked on ‘Click here’ button for Warning 3.3 in Question 10. Thus you land at the ‘Facebook login page. You proceed to enter your login credentials. But when you click on the user name or password input box, a pop up warning message (Warning 3.4) is displayed as displayed below:

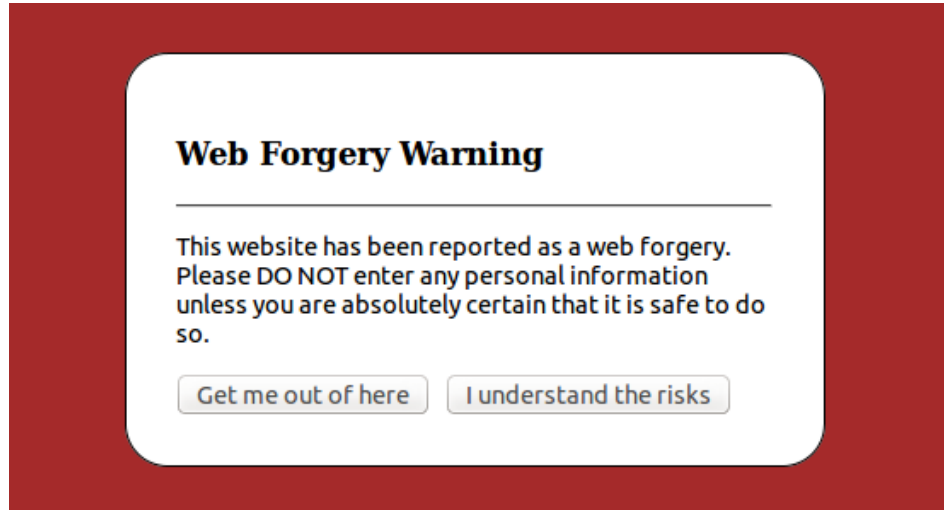


Figure A.8: Appendix- Warning 3.4

12. When you encounter the above warning message, what will be your next step of action?

- Click on ‘Get me out of here’ button
- Click on ‘I understand the risks’ button
- Close the tab by clicking the close symbol of the tab

13. Briefly explain why.

In Scenario B, you encountered a new phishing warning design which is interactive and works in multiple stages. Warning 3 to 3.3 are displayed in an attempt to stop a user from proceeding to a potential fake website by identifying the real website that the user tried to visit and making the user land on the actual website. If the user still manages to proceed to the spoof website, warning 3.4 tries to warn the user once again about the dangers of entering any confidential information in the website.

Please share your thoughts on this warning design by answering these following question:

14. Does the warning system make you believe that you might be under some cyber attack?

- Yes
- No

15. Does this new warning system make you feel more secured compared to the warnings you have seen earlier?

- Yes
- No

16. Did you feel that encountering multiple warning messages had any effect on your judgment about the possible threats?

- Yes
- No

17. Briefly explain why.

18. On a scale from 1 to 10, how would you rate this warning system on keeping you safe from a spoof website? (1: Least effective - 10: Most effective)

19. Please explain if you think this warning system is better/worse than the phishing warnings that you have seen before. (Please answer N/A if you haven't seen a phishing warning before.)

APPENDIX B
INSTITUTIONAL REVIEW BOARD APPROVAL

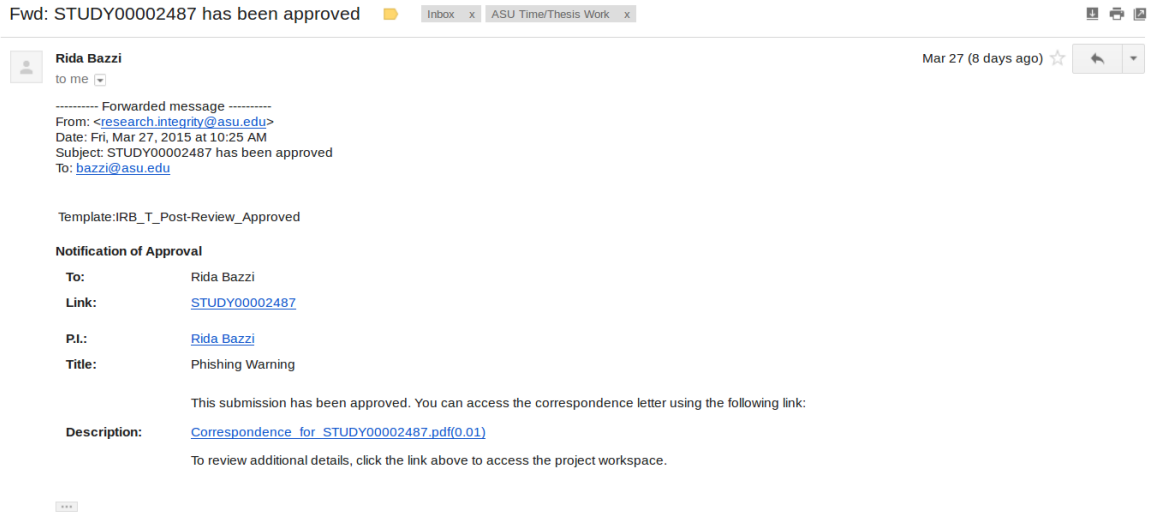


Figure B.1: Appendix- IRB Email Approval

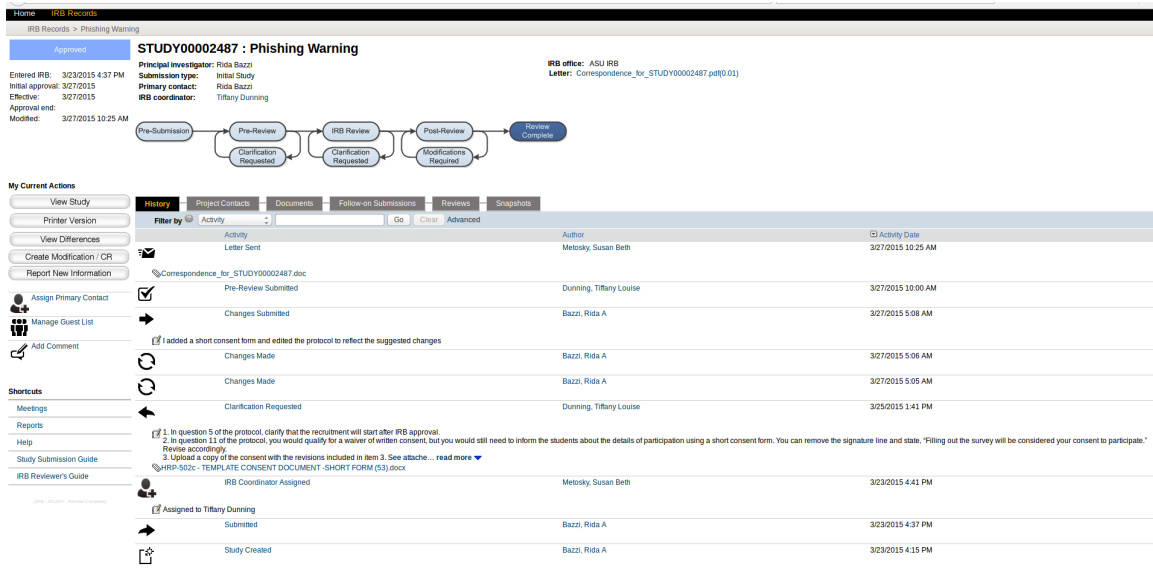


Figure B.2: Appendix- IRB Approval ERA Website

EXEMPTION GRANTED

Rida Bazzi
 CIDSE: Computing, Informatics and Decision Systems Engineering, School of
 480/965-2796
 bazzi@asu.edu

Dear Rida Bazzi:

On 3/27/2015 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Study of effectiveness of using target information in phishing warning messages
Investigator:	Rida Bazzi
IRB ID:	STUDY00002487
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	<ul style="list-style-type: none"> • Bazzi-Security-Study-HRP-503a - TEMPLATE PROTOCOLSOCIAL BEHAVIORAL-2.docx, Category: IRB Protocol; • Bazzi_Recruitment_Form-2.pdf, Category: Recruitment Materials; • RidaBazziCV (1).pdf, Category: Vitaes/resumes of study team; • Bazzi_Study_Questionnaire.pdf, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • Bazzi-Security-Study-HRP-502c%20-%20TEMPLATE%20CONSENT%20DOCUMENT%20-SHORT%20FORM%20(53)-2.pdf, Category: Consent Form;

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (2) Tests, surveys, interviews, or observation on 3/27/2015.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

Sincerely,

IRB Administrator

cc: Satyabrata Sharma

Figure B.3: Appendix- IRB Approval Correspondence