Fighting Internet Fraud: Anti-Phishing Effectiveness for Phishing Websites Dectection.

Alnajim, Abdullah M.

Fighting Internet Fraud: Anti-Phishing Effectiveness for Phishing Websites Detection

Abdullah M. Alnajim

A thesis submitted to the Durham University for the degree of Doctor of Philosophy

Department of Computer Science

Durham University

2009

Abstract

Recently, the Internet has become a very important medium of communication. Many people go online and conduct a wide range of business. They can sell and buy goods, perform different banking activities and even participate in political and social elections by casting a vote online. The parties involved in any transaction never need to meet and a buyer can sometimes be dealing with a fraudulent business that does not actually exist. So, security for conducting businesses online is vital and critical. All security-critical applications (e.g. online banking login pages) that are accessed using the Internet are at the risk of fraud. A common risk comes from so-called Phishing websites, which have become a problem for online banking and e-commerce users. Phishing websites attempt to trick people into revealing their sensitive personal and security information in order for the fraudster to access their accounts. They use websites that look similar to those of legitimate organizations and exploit the end-user's lack of knowledge of web browser clues and security indicators.

This thesis addresses the effectiveness of Phishing website detection. It reviews existing anti-Phishing approaches and then makes the following contributions. First of all, the research in this thesis evaluates the effectiveness of the current most common users' tips for detecting Phishing websites. A novel effectiveness criteria is proposed and used to examine every tip and rank it based on its effectiveness score, thus revealing the most effective tips to enable users to detect Phishing attacks. The most effective tips can then be used by anti-Phishing training approaches. Secondly, this thesis proposes a novel Anti-Phishing Approach that uses Training Intervention for Phishing Websites' Detection (APTIPWD) and shows that it can be easily implemented. Thirdly, the effectiveness of the New Approach (APTIPWD) is evaluated using a set of user experiments showing that it is more effective in helping users distinguish between legitimate and Phishing websites than the Old Approach of sending anti-Phishing tips by email. The experiments also address the issues of the effects of technical ability and Phishing knowledge on Phishing websites' detection. The results of the investigation show that technical ability has no effect whereas Phishing knowledge has a positive effect on Phishing website detection. Thus, there is need to ensure that, regardless their technical ability level (expert or non-expert), the participants do not know about Phishing before they evaluate the effectiveness of a new anti-Phishing approach. This thesis then evaluates the anti-Phishing knowledge retention of the New Approach users and compares it with the knowledge retention of users who are sent anti-Phishing tips by email.

Declaration

The material contained within this thesis has not previously been submitted for a degree at the Durham University or any other university. Parts of the work presented in this thesis have been published as book chapter, journal article and conference proceedings.

Book chapter

1. Alnajim, A. and Munro, M., 2009. An Anti-Phishing Approach for Phishing Websites Detection. *In*: Pichappan, P., ed. *Handbook of Research on Threat Management and Information Security: Models for Countering Attacks, Breaches and Intrusions*. Pennsylvania USA: IGI Global, (To appear).

Journal article

2. Alnajim, A. and Munro, M., 2009. Detecting Phishing Websites: On the Effectiveness of Users' Tips. *Journal of Information Assurance and Security (JIAS)*, ISSN 1554-1010, (To appear).

Conference proceedings

3. Alnajim, A. and Munro, M., 2008. An Evaluation of Users' Tips Effectiveness for Phishing Websites Detection. Proceedings of the third IEEE International Conference on Digital Information Management ICDIM, London, IEEE Press, pp. 63-68.

4. Alnajim, A. and Munro, M., 2009. Effects of Technical Abilities and Phishing Knowledge on Phishing Websites Detection. Proceedings of the IASTED International Conference on Software Engineering (SE 2009), Innsbruck, Austria, ACTA Press, pp. 120-125.

5. Alnajim, A. and Munro, M., 2009. An Evaluation of Users' Anti-Phishing Knowledge Retention. Proceedings of the International Conference on 2009 International Conference on Information Management and Engineering (ICIME 2009), Kuala Lumpur, Malaysia, IEEE Press, pp. 210-214.

6. Alnajim, A. and Munro, M., 2009. An Anti-Phishing Approach that Uses Training Intervention for Phishing Websites Detection. Proceedings of the 6th IEEE International Conference on Information Technology - New Generations (ITNG), Las Vegas, USA, IEEE Press, pp. 405-410.

Acknowledgement

First of all, all praise and thanks go to Allah for all the success in this thesis in particular and in my life in general.

Secondly, my deep thanks to my supervisor Professor Malcolm Munro for his invaluable advice and guidance. Professor Munro gave me all benefits from his wide experience. My thesis would not be completed and succeeded without his great supervision.

Thirdly, my sincere thanks to my great mother for her constant encouragement, support, prays and patience. My thanks to my wife, sisters and brother for their prays.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Introduction

The Internet is a very important medium of communication. Many people go online and conduct a wide range of business. They can send emails, sell and buy goods, transact various banking activities and even participate in political and social elections by casting a vote online. The World Wide Web technologies enable people around the world to participate in commercial activities whenever they wish and wherever they live [Poong et al.06]. There are many successful and widely used e-commercial websites. There are e-marketplace websites such as Amazon[1], and online auction websites such as eBay[2] that offer an online platform where millions of items are exchanged each day. The use of online banking services has been growing at a tremendous rate [Reavley05]. Many banks and financial societies have online.banking platforms. For example, the HSBC bank has nearly 19 million Internet registered users [Hilley05].

Once users go online, they are at risk from online fraud (also known as Internet fraud). Internet fraud is a crime that uses the Internet as the medium to carry out financial frauds [Philippsohn01]. The parties involved in any transaction never need to meet and the user may have no idea whether the goods or services exist. Due to this, the Internet is a good vehicle to defraud people who use it to buy goods or services [ibid]. The application access keys could be stolen. Applications such as electronic commerce, electronic banking, electronic voting and electronic mail are targets for fraudsters.

---

[1] Amazon is a well-known electronic commerce company. Available at: http://www.amazon.com, last access on 4 Feb 2007.

[2] EBay is an online auction website. Available at: http://www.ebay.com, last access on 4 Feb 2007.

Security for conducting businesses online is vital and critical. All security-critical applications (e.g. online banking login page) that are accessed using the Internet are at the risk of Internet fraud. Violations of security in these applications would result in severe consequences, such as financial loss for e-commerce and online banking organizations and for individuals. CyberSource [CyberSource08] has revealed that financial loss due to Internet fraud is huge; in 2007, such losses amounted to $3.6 billion.

1.2. Phishing

Internet fraud has a multiplicity of forms, including Phishing attacks. Phishing has become a serious problem for online banking and e-commerce users [Chandrasekaran et al.06]. It takes the form of an email message or website that tries to trick people into revealing personal security-sensitive information by appearing to be from a legitimate organization but it exploits the end-user's lack of knowledge about web browser clues and security indicators [Dhamija et al.06]. The emails and websites appearing to be from a legitimate organization are known as Phishing emails and Phishing websites respectively. Phishing attacks have increased dramatically. 36,002 unique Phishing URLs were active and 139 brands were hijacked in February 2008 [APWG08].

The Phishing problem arises when a user receives a Phishing email. They may not understand that the link provided may not take them to where they expect. For example, the user's intention may be "go to eBay" but the actual implementation of the hyperlink may be "go to a server in South Korea". This misunderstanding enables Phishing and makes it very hard to defend against. Users gain their understanding of interaction from the presentation or the way it appears on the screen. Some technical details of web pages and email messages are hidden and some of them are not understandable to most users. Thus, the user does not interpret the system clues or is unable to do so.

1.3. Phishing Detection and Prevention

The Phishing problem needs to be mitigated by anti-Phishing approaches. There must be solutions that help in detecting and preventing Phishing attacks. The effectiveness of anti-Phishing approaches must be increased.

There have been some approaches to mitigate Phishing, such as toolbars and anti-Phishing tips. The effectiveness of 24 existing online training materials that teach people how to protect themselves from Phishing attacks have been evaluated [Kumaraguru et al.07b]. However, this research did not consider the effectiveness of each individual tip.

To access and read online training material, users usually need to open new web browsers. Then they go back to their online activity browser to proceed. But this is only likely to happen if users know that there are attacks called Phishing and that there are training materials that help in detecting them. If the users know nothing about Phishing and anti-Phishing training materials, they are unlikely to access them. In fact, few people read anti-Phishing online training materials although they are surprisingly effective when users do read them [Kumaraguru et al.07b]. A novel approach was to design an online game in order to teach users good habits to help them avoid Phishing attacks [Sheng et al.07]. The game presents anti-Phishing information in an enjoyable way. However, the disadvantage of this approach is the same as for other online training materials. Users must know something about Phishing and its dangers before they are likely to access and play the game.

Many commercial institutions, such as Microsoft [Microsoftb], provide a service that periodically sends emails that warn people about Phishing emails and websites and that provide tips to help people detect Phishing websites. However, only subscribed customers receive these emails.

Kumaraguru et al. [Kumaraguru et al.07a] considered training people about Phishing email during their normal use of email. Their aim was to teach people what Phishing clues to look for in emails to make better decisions in identifying Phishing emails. They found that this approach works better than the current practice of publishing or sending anti-Phishing tips by email. However, Kumaraguru et al.'s approach did not consider helping

people with Phishing website-related tips. Phishing sites can be reached via various methods in addition to emails, such as online advertisements and typing their web addresses in a web browser. Helping users in distinguishing between Phishing and legitimate websites during their normal browsing activities is required.

In the process of designing anti-Phishing approaches, user experiments were conducted to evaluate them. Several approaches which used participants recruited on the basis of their technical abilities were evaluated [Downs et al.06, Kumaraguru et al.07a, Kumaraguru et al.07b, Sheng et al.07]. In these studies, participants were classified as 'experts' and 'non-experts' based on pre-study screening questions. Technical ability was judged on whether the participants had changed preferences or settings in their web browser, created a web page, and helped someone fix a computer problem. Participants who said 'no' to at least two of the screening questions were categorized as 'non-experts' and were selected to take part in their experiments. However, no question was asked about Phishing or Internet fraud so it is possible that participants who were considered to be non-experts could know about Phishing and how to detect Phishing attacks before participating in the evaluation experiments. Having participants with Phishing knowledge in advance may provide biased results in evaluation experiments on anti-Phishing approaches. This is because people who know about Phishing before participating in the evaluation experiments may use their prior knowledge rather than the anti-Phishing approaches that are being tested in the evaluation. Downs et al. [Downs et al.07] studied whether there are correlations between some web environment experiences and susceptibility to Phishing. They found that people who correctly answered the knowledge question about the definition of Phishing (i.e. Phishing aware people) were significantly less likely to fail to detect Phishing emails. Low technical users (i.e. non-experts) may be Phishing aware and high technical users (i.e. experts) may be Phishing unaware. Therefore, an investigation on the effects of technical ability and Phishing knowledge on Phishing websites' detection is required. This would clarify whether or not the previous screening questions for recruiting low technical users in evaluating anti-Phishing approaches are beneficial.

In this thesis, problems related to the effectiveness of approaches to Phishing websites detection have been addressed. Firstly, the effectiveness of the most common users' tips for detecting Phishing websites is examined. The effectiveness of each individual tip is assessed

and then the tip is ranked accordingly. The aim is to identify the most effective anti-Phishing tips that users can focus on to detect Phishing attacks.

This thesis also proposes a novel Anti-Phishing Approach that uses Training Intervention for Phishing Websites' Detection (APTIPWD). User experiments were conducted to evaluate this approach. The thesis shows that the approach can be easily implemented.

An investigation that assesses using Phishing knowledge instead of technical ability in the screening questions to recruit participants is also presented. User experiments are conducted to evaluate the effects of technical ability and Phishing knowledge. If the results of the investigation show that there is no effect of technical ability on Phishing website detection, then there is need to make sure that the participants do not know about Phishing regardless of their technical ability level in evaluating a new anti-Phishing approach.

This thesis also assesses the anti-Phishing knowledge retention of users. User experiments are conducted. The knowledge retention of the users of the New Approach (APTIPWD) and the knowledge retention of users sent anti-Phishing tips by email are compared.

1.4. Criteria for Success

In this thesis, the criteria for success are set as follows.

1. An evaluation of the anti-Phishing tips' effectiveness for Phishing websites detection.

   An examination of the effectiveness of the most common users' tips for detecting Phishing websites will be presented. Novel effectiveness criteria will be proposed and used to examine each single tip and to rank them based on their effectiveness scores.

2. Development of a more effective anti-Phishing approach and its evaluation.

   This thesis will propose a more effective approach that resolves some issues identified in previous approaches. The New Approach will be evaluated and the results will be discussed.

3. Success to identify factors that influence users decisions against Phishing websites.
   This criterion is divided into two sub-criterions. They are as follows:

   3.1. Effect of technical ability on Phishing websites detection.

   The effects of the technical ability of users on Phishing website detection will be discussed. User experiments will be conducted to evaluate the effects of technical ability and the results will be analyzed.

   3.2. Effect of Phishing knowledge on Phishing websites detection.

   The effects of the Phishing knowledge of users on Phishing website detection will be evaluated. User experiments will be conducted and the results will be analyzed.

4. An evaluation of the anti-Phishing knowledge retention when using the New Approach.
   This thesis will evaluate the anti-Phishing knowledge retention of users of the New Approach. User experiments will be conducted and the results will be analyzed.

5. Comparisons with other related studies.
   The work in this thesis will be compared with the relevant work of others. Discussions on the similarities and differences will be presented.

6. A proof of concept implementation.
   A prototype proof of concept will be presented in order to demonstrate that the New Approach is implementable and viable.

1.5. Thesis Overview

This thesis is structured as follows. Chapter 2 gives an overview of Internet fraud in general. Internet fraud is defined, and types and examples are discussed. The chapter identifies some web applications that are suffering from Internet fraud and gives some statistics for its impact. It also presents existing techniques and strategies to detect and prevent Internet fraud.

Chapter 3 discusses Phishing attacks. The problem is defined and some real examples are discussed. The impact of Phishing attacks is presented with statistics that reveal its trends. The chapter reviews existing approaches in detecting and preventing Phishing emails and websites and discusses their limitations.

Chapter 4 begins with an overview of training and discusses its definition and methodologies. It goes on to present an overview of embedded training, discusses its advantages and provides examples. Finally, the chapter looks at people's retention of the knowledge obtained from training and the factors that affect the retention rate.

Chapter 5 presents an overview of the experimental designs and statistical analysis used in this thesis. It shows how the research question is translated into a hypothesis and the steps to performing an experiment and testing the hypothesis. The chapter concludes with an overview of common statistical analysis methods that are used in this thesis.

Chapter 6 examines the effectiveness of the most common users' tips for detecting Phishing websites. A set of novel effectiveness criteria is proposed and used to examine each single tip and rank it based on its effectiveness score. An attempt is made to find the best anti-Phishing tips that users can focus on to detect Phishing attacks by themselves.

Chapter 7 proposes a novel anti-Phishing approach that uses training intervention for Phishing websites' detection (APTIPWD). The chapter also presents a prototype proof of concept implementation of the proposed approach. The chapter shows the design and then the implementation of the prototype. The aim of the implementation is to validate whether the New Approach is doable and viable.

Chapter 8 presents the design of the evaluation experiments and the research hypotheses. Details are provided about the way in which the experiments' participants were recruited and about their demographic profile. The chapter presents the effectiveness ratios that are used in evaluating the hypotheses. It also shows comparisons between real Phishing attacks and Phishing experiments in order to decide what should be simulated in the experiments. The story board of the experiments is also presented.

Chapter 9 discusses the evaluation of the research hypotheses. The hypotheses are classified into four research themes, which are evaluating the New Approach, the effect of high and low technical abilities on Phishing detection, the effect of Phishing awareness and Phishing unawareness on Phishing detection and anti-Phishing knowledge retention.

Chapter 10 compares the work in this thesis with related anti-Phishing approaches. It includes a discussion on the similarities and differences between the evaluations in this thesis and the work of others. Issues such as participants' recruitment, scenarios, emails and websites, anti-Phishing tips used, results and implementation are discussed.

Chapter 11 presents the conclusion of this thesis, summarizes its original work and identifies directions for future research.

1.6. Assumptions

In this thesis, there is an assumption that Phishing attacks do not use either software to change the host files in users' operating systems or any malicious software, such as a virus, worm or Trojan horse, that runs in users' operating systems. These are called 'Pharming' and 'Malware' and are different from Phishing. Phishing is a deceptive attack which aims to take advantage of the way humans interact with computers or interpret messages rather than taking advantage of the technical system vulnerabilities [Downs et al.06].

1.7. Summary

In this chapter, an introductory overview of Internet fraud and Phishing was presented and the problem of Phishing was briefly discussed. The thesis's original work and its criteria for success were given. Finally, the structure of the thesis was shown.

2. Internet Fraud

2.1. Introduction

The aim of this chapter is to give an overview of Internet fraud in general. The Internet fraud definition, types and examples will be discussed. The chapter identifies some web applications that are suffering from Internet fraud as well as some statistics for Internet fraud impact. It also presents some existing techniques and strategies to detect and prevent Internet fraud.

2.2. Definition

Fraud is defined as '*an act or instance of deception, an artifice by which the right or interest of another is injured, a dishonest trick or stratagem*' [OED]. Fraud can be committed using variety of methods. In recent times, the Internet has been a suitable method for committing fraud because the Internet allows hiding real identification of people who deal with it. Therefore, fraudsters use the Internet in order to appear anonymous.

Once users go online, they are at risk from Internet fraud. Internet fraud is defined by Philippsohn [Philippsohn01] as any crime that uses the Internet as the medium to exercise the ability to carry out financial frauds. In addition to this, Internet fraud is sometimes called 'Internet Scams' [CAB06]. The parties involved in any transaction never need to meet and the user may have no idea whether the goods or services exist. Due to this, the Internet is a good vehicle to defraud the users who would like to buy goods or services using it [Philippsohn01].

2.3. Internet Fraud Types

2.3.1. *Types*

Internet fraud has a multiplicity of types. In the literature, there is no exact number or fixed list of these types. Below is a simple taxonomy described for these types:

- **Rogue traders:**

Rogue Internet traders are untrustworthy or dishonest merchants who sell goods or services using the Internet [CAB06]. The most common fraud cases that dishonest merchants commit when selling something online are [CAB06]:

- Merchant advertises goods that do not exist.
- Merchant makes untrue statements about the things they are selling.
- Merchant sells dangerous goods.
- Merchant does not tell about import or transport costs.
- Merchant sends different goods to the ones they advertised.
- Merchant does not deliver on time.
- Merchant does not deliver at all.

Few of the listed cases occur in the Internet auction fraud. The Internet Crime Complaint Center (IC3) [IC3] states that '*auction fraud involves fraud attributable to the misrepresentation of a product advertised for sale through an Internet auction site or the non-delivery of products purchased through an Internet auction site*'.

- **Credit card fraud:**

Credit card fraud is where an unauthorized person uses a credit or debit card to obtain money or purchase merchandise [IC3]. The fraudsters make online purchases with the credit card details of other people which is known as Card-Not-Present (CNP). CNP fraud is a credit card fraud that is committed over the Internet, mail, fax or phone without the need to present the card physically [APACSc]. According to APACS [APACSa] '*the anonymity of CNP transactions allows fraudsters to disguise their true identity. They may use fictitious personal details in conjunction with fraudulently obtained card details to make illegal purchases*'. For most of CNP cases, credit or debit card numbers are stolen from unsecured

websites, taken from discarded receipts or obtained in an identity theft scheme without the cardholder's knowledge [APACSa, IC3].

- **Lottery fraud:**

Lottery fraud has been very common on the Internet even though the victim may never have participated in a lottery. The victims receive an email that says they are the winners of an International lottery. Next, the victims are told that they have to send money to claim the prize or has to ring a premium rate number which is very expensive [CAB06, IC3].

- **Pharming:**

Pharming is defined by APWG as a web security attack that happens when a user types in an address and the browser they use redirects them to a fraudulent website without their knowledge [APWG07a]. Pharming can be conducted by exploiting vulnerability in Domain Name Server (DNS)[3] and changing the content of the directory, which contains the domain, to IP directory [Jammalamadaka et al.05].

- **Phishing:**

Recently, Phishing scams have become one of the serious problems encountering end-users in the Internet world [Chandrasekaran et al.06]. Phishing is an attack that exploits the end-user lack of knowledge in terms of web browser clues and security indicators and uses similar looking emails and websites for legitimate organizations to trick people in order to reveal sensitive information [Dhamija et al.06]. Due to the fact that the main focus of this thesis is Phishing, there will be more detailed description of Phishing in Chapter 3.

2.3.2. Summary

There is no fixed list of the Internet fraud types. A simple taxonomy of the types was described. Parties involved in a transaction on the Internet may commit fraud. In the Internet auctions, merchants can defraud customers. The Internet users also receive fraudulent

[3] DNS stands for Domain Name System. The DNS main task is mapping symbolic host names to their IP addresses [Friedlander et al.07].

emails discussed as 'lottery fraud'. Pharming and Phishing also are used to steal users' sensitive information such as credit card details.

2.4. Internet Applications that are Suffering from Internet Fraud

All security-critical applications (e.g. online banking login page) that are accessed using the Internet are at the risk of Internet fraud. The reason is that the application access keys could be stolen. Applications such as electronic commerce, electronic banking, electronic voting, electronic mail and so forth might be targets for fraudsters. Due to their financial losses, electronic commerce and online banking will be briefly presented.

2.4.1. Electronic Commerce

Gatautis and Neverauskas [GatautisNeverauskas05] described electronic commerce as *'form of trading relations, in which interrelated parties interact in electronic way, using information technologies'*. The World Wide Web technologies enable people around the world to participate in commercial activities whenever and wherever without any boundaries [Poong et al.06].

There are many successful and widely used e-commercial websites. There are e-marketplace websites such as Amazon[4]. Also, there are Internet auction websites such as eBay[5] which offers an online platform where millions of items are exchanged each day.

An e-commerce transaction involves some steps. A customer browses a commercial website, selects goods and then checks out. Then, the customer reaches the payment process where they need to provide the payment page with valid payment card details. After that, the

[4] Amazon is a well-known electronic commerce company. Available at: http://www.amazon.com, last access on 4 Feb 2007.
[5] EBay is an online auction website. Available at: http://www.ebay.com, last access on 4 Feb 2007.

merchant delivers the goods to the customer's physical address or the customer's email when the goods are digital (e.g. e-tickets and music).

The payment process in the e-commerce transaction is one of the most important success factors in electronic commerce [Juang07]. However, the payment process is likely to have fraud possibility. One possible fraud case is a fraudster uses stolen credit card details in a payment process. This case can be considered CNP fraud case. Also, a dishonest merchant can commit fraud in the payment process. The merchant can double bill the customer, or can use the customer payment details in another payment process. Moreover, the merchant can pass on the customer payment details to criminals [DaraGundemoni06].

In the case of digital goods, the merchants can commit fraud by not delivering the goods (e.g. piece of music) to the customers' email addresses after they receive their money. The delivery of the digital goods is difficult to verify since there is no signature required when the goods are delivered as used now [Alfuraih02]. The signature is required when hard-goods[6] are delivered to a physical address.

2.4.2. Online Banking

Online banking (also known as Internet banking) is a term described by Aladwani [Aladwani01] as carrying out most banking services such as accessing bank accounts, balance reporting, money transfers and bill-payment electronically using the Internet. The use of online banking services has been growing at a tremendous rate [Reavley05]. Claessens et al. [Claessens et al.02] point out that online banking systems give everybody the chance to access their banking details and do banking activities easily.

Today, many banks and financial societies have their online banking platforms. For example, the HSBC bank has nearly 19 million Internet registered users [Hilley05]. Because

[6] Hard-goods include all tangible products that require delivery to a physical address if purchased, such as laptops or clothes [Alfuraih02].

of wide usage of online banking, more than 130 million Europeans were expected to conduct their banking transactions online in 2007 [Reavley05].

Any online banking transaction involves two parties which are the customer and the bank. In contrast, e-commerce transaction requires one additional party which is the merchant. In e-commerce there are more possibilities for fraud as both the customer and merchant can commit fraud.

2.5. Impact and Statistics

The following presents some figures of the negative impact of Internet fraud on companies and financial market in the last few years. According to McKenna [McKenna05], a survey revealed that:

- 90% of the 200 companies participated in the survey suffered from unauthorized penetration of company systems.
- 89% suffered from theft of information.

The Internet fraud influence hits everywhere. CyberSource [CyberSource08] has revealed that the total financial losses from Internet payment fraud alone in the United States and Canada have steadily increased in the period between 2004 and 2007 as e-commerce has continued to grow approximately 20% each year. In addition, online theft costs $1 trillion a year and the number of Internet fraud attacks is increasing sharply and too many people do not know how to protect themselves [Weber09].

The fear of Internet fraud also drives the Internet shoppers away from practicing e-commerce. More than half of the adult population in the UK does not shop online because they do not know how to use a computer, they prefer shopping on the high street or they do not have an Internet access [CyberSource09]. But worryingly, 41% of those who do not shop online said it is because of the fears of Internet fraud [ibid].

2.6. Mitigating Internet Fraud

Due to the fact that there have been fraud attempts (attacks), it is normal to have protection attempts (defenses) against them introduced by industry and academic researchers alike. Some technologies used in mitigating Internet fraud as well as efforts for increasing user anti-fraud awareness are discussed in this section.

2.6.1. Technologies Used

In e-commerce, the main transaction's stockholders, customer, merchant and the card issuing bank, need to make sure that each one is satisfied and authenticated [Cook02]. Merchants need to be reassured that the customers they do business with really are legitimate. Customers need to be reassured that their card details are not being used by unauthorized persons to make purchases on the Internet in their name. Also, the card issuers need to know that they are not involved in a fraud loss [ibid].

Once customers have completed their purchase on a merchant's Internet payment page, their card data is transferred directly to the card issuer. The problem here is that if the card issuer considers the transaction is fraudulent, the merchant, who is held responsible for not verifying the cardholder's identity, is likely to lose the income from the sale in addition to the value of the products sold [ibid]. The merchant then will dispute the transaction and claim their money back. This process is known as 'chargeback' [ibid].

To solve this problem, some technologies such as MasterCard's '*Secure Code*', Visa's '*Verified by Visa*', Address Verification Service and Card Code Value have been introduced and aimed to verify the card information and, in turn, to authenticate its user.

- **MasterCard's SecureCode and Visa's Verified by Visa (VbV):**

MasterCard SecureCode and *Verified by Visa (VbV)* services are similar and based on the 3D Secure Protocol [APACSa]. MasterCard's solution is called Secure Payment Application (SPA) [Cook02]. In *MasterCard SecureCode*, cardholder registers to

MasterCard and then downloads and installs a browser plug-in or electronic wallet. After that, if a customer starts purchasing process at a MasterCard's participating Internet retailer, a secure window will appear requesting the customer's SecureCode pin number. Then the customer enters their pin number in the window. In seconds, the transaction will be authenticated and confirmed. Then, the purchase can be completed [MasterCard].

The *Verified by Visa (VbV)* case is explained as follows. Merchants Plug-In (MPI) software is installed on the Merchants' systems [Visa05]. Then, when a registered cardholder executing a transaction reaches the check-out page and clicks the 'buy' button, a VbV session is automatically initiated. Customers will know that they are on a secure website, since it will carry Visa's VbV symbol. The e-shopper will complete the payment page normally, submit their card details and then the system will check if the card issuer is participating in the VbV scheme. Shoppers who are registered with their card issuer will then be presented with a pop-up window and asked to enter a PIN number to prove their identity [Cook02].

- **Address Verification Service (AVS) and Card Security Code (CSC):**

The banking industry introduced AVS/CSC services in 2001 to help merchants in preventing CNP fraud [APACSa]. Both AVS and CSC are designed to make it difficult for fraudsters because they require more knowledge than just a card number and expiry date [LogicGr]. As Figure 1 illustrates, CSC code is the last three digits located on the signature strip on the back of the card. CSC ensures that the card is with the customer while they are making a transaction on the web. While, AVS allows the merchant to confirm the numbers in the billing address of a cardholder with the card issuer database [APACSa]. Therefore, it is less likely that fraudsters will be able to provide the genuine cardholder's address whereas they may be able to provide a CSC with a lost or stolen card [APACSa].

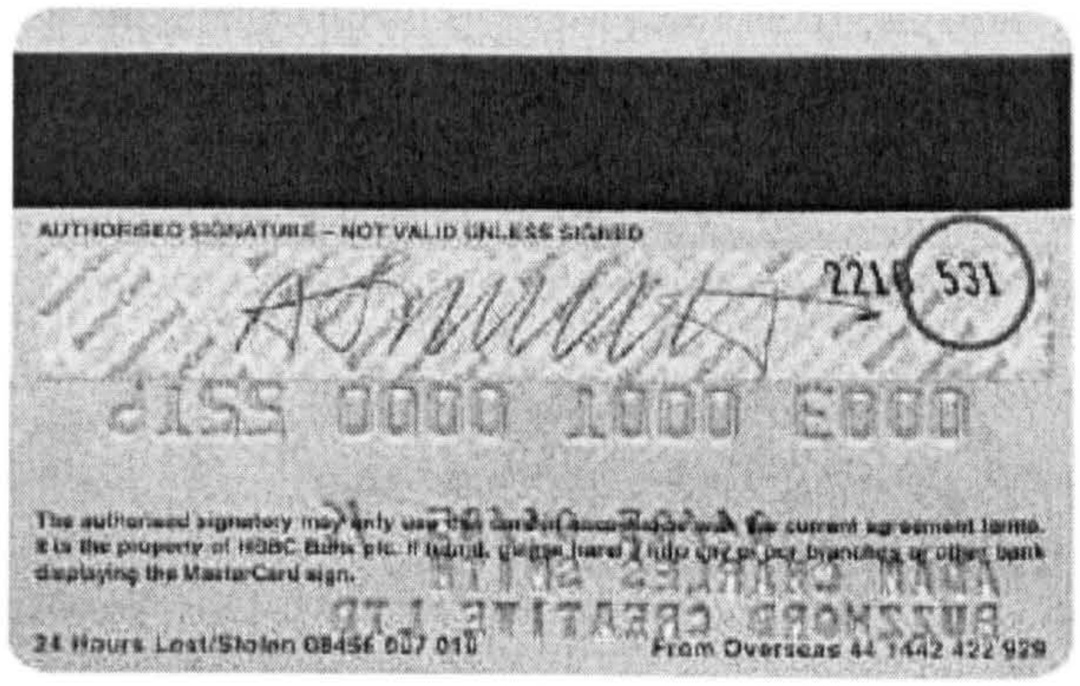


Figure 1: Card security code (CSC) [APACSb]

2.6.2. Consumer Training

Due to the fact that no complete solution mitigates Internet fraud, both technical and educational solutions are required methods of preventing the fraud [Symantec].

Financial and commercial, private and government institutions have developed Internet fraud awareness websites in which provide Internet fraud prevention training materials. For example, there are banks and financial institutions such as PayPal, HSBC and Citibank. Also, there are some popular electronic commercial websites such as Amazon and eBay. Some anti-fraud organizations such as Anti-Phishing Working Group (APWG)[7], Bank Safe Online[8] website and Card Watch[9] website which are run by APACS, the UK payments association and Get Safe Online[10] website (See Figure 2) which is sponsored by the UK government jointly with some organizations such as eBay, BT[11], Microsoft, HSBC and Securetrading. In addition to this, as Figure 3 illustrates, credit card companies such as Visa[12] and MasterCard[13] have been doing work in the field of 'Consumer Education'. Thus,

---

[7] Anti-Phishing Working Group. Available at: http://www.antiphishing.org, last access on 4 Feb 2007.
[8] Bank Safe Online. Available at: http://www.banksafeonline.org.uk, last access on 4 Feb 2007.
[9] Card Watch. Available at: http://www.cardwatch.org.uk, last access on 4 Feb 2007.
[10] Get Safe Online. Available at: http://www.getsafeonline.org, last access on 4 Feb 2007.
[11] BT is the British Telecommunications company.
[12] Visa. Available at: http://www.visa.com, last access on 4 Feb 2007.
[13] MasterCard. Available at: http://www.mastercard.com, last access on 4 Feb 2007.

these institutions are interested in making the human element (end-user) good enough in detecting and preventing Internet fraud.

The online training materials provided by the institutions and organizations are typically written as textual lists. The tips provided by these institutions are for variety of purposes. They are for users doing online shopping, online auction, online banking, receiving emails and so forth. Some examples of the anti-fraud textual tips are worth mentioning.

Figure 2: Get Safe Online website

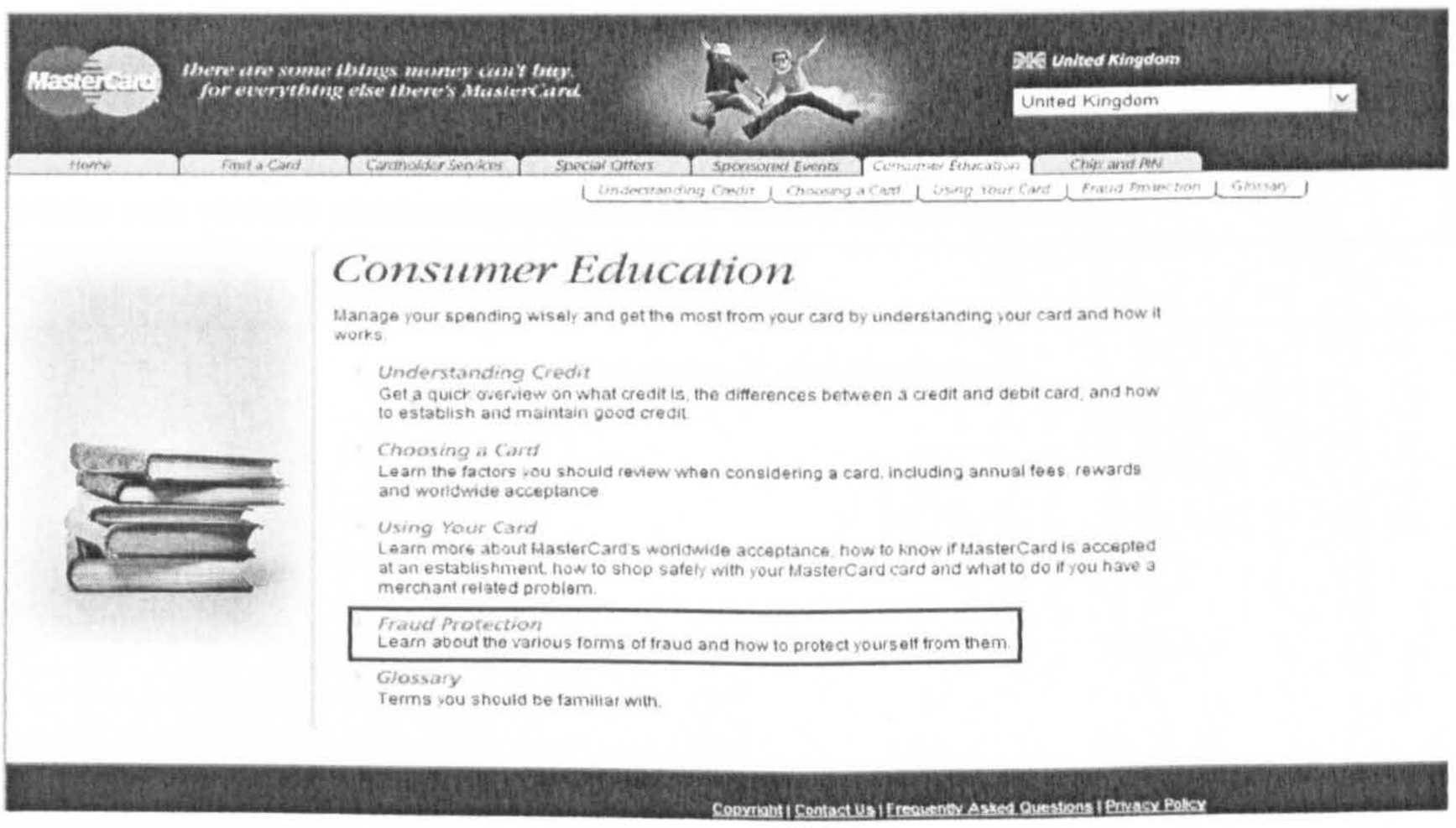


Figure 3: Consumer education section at MasterCard website

- **Anti-fraud textual tips**

The anti-fraud user's tips are different based on the task they intend to do. Examples of the tips are shown in Table 1.

#	*Tip*
1	When creating a password, it should not be a word at all. It can be a combination of letters, numbers and keyboard symbols.
2	Password should contain a mix of upper and lower case letters, numbers and keyboard symbols.
3	Be different by avoiding using the same password for different services.
4	If you cannot avoid using a public computer, after you are done, log out of all websites, clear the browser's cache and history, and close the browser.
5	Delete suspicious emails with attachments and never open the attachments.
6	Do not download attachments: we will never send you an attachment or software update to install on your computer.
7	Never open an email attachment that contains a file ending with .exe, .pif, .vbs as these are commonly used with viruses.
8	Do not send your credit card number to anyone in an unsecured email.
9	Only access your personal financial information from a computer you trust. Avoid using shared computers such as those in Internet cafes.
10	Make sure you are on a secure connection when entering sensitive information. Secure Web pages will have the text https: (note the "s") instead of http:
11	Before using the website, check out the security/encryption software it uses.
12	Look for Third-Party Merchant Reviews. Many news websites offer reviews of shopping sites. These resources can be a great place to start your online shopping searches.
13	Make sure you are purchasing merchandise from a reputable source.
14	Do not judge a person/company by their web site.

Table 1: Examples of anti-fraud tips[14]

14 The tips are cited from PayPal, Amazon, eBay, HSBC, Card Watch and Get Safe Online.

2.7. Summary

This chapter has presented an overview of Internet fraud that is a major problem to e-commerce and online banking. It has defined Internet fraud, described its different types and presented existing solutions to combat fraud. With regards to the potentiality to have fraud incidents, it is clear that a single online banking transaction involves two parties which are the customer and the bank. In contrast, e-commerce transaction involves one more party which is the merchant. Due to the nature of the parties, e-commerce is more likely to have fraud incidents since both the customer and merchant can possibly commit fraud.

This chapter showed that Phishing attacks are types of Internet fraud. Any method that helps in reducing and mitigating Phishing attacks will help in reducing Internet fraud.

3. Phishing

3.1. Introduction

Gullibility is the quality of being gullible. Gullible is defined as '*too willing to believe or accept what other people tell you and therefore easily tricked*' [Hornby00] whereas trust is defined as believing others in the absence of hard and clear evidence to disbelieve [Rotter80]. The question 'whether trust means gullibility or not' is widely discussed by many psychologists. People who are considered trustful are also considered to be naive and gullible in the conception of trust [Yamagishi et al.99]. It is commonly believed that those people who tend to trust others without hard evidence are easy victims to fraudsters in the social jungle [ibid].

Phishing attacks are committed by fraudsters. In this chapter, Phishing attacks are considered. The problem definition and some of examples are discussed. The impact of Phishing attacks is presented with some statistics that reveal its trends. The chapter presents the existing research in suitability to Phishing risks. It also reviews the existing approaches in detecting and preventing Phishing emails and websites. The chapter concludes with a discussion on limitations of anti-Phishing approaches.

3.2. Problem Definition

Recently, Phishing attacks have become a serious problem for end-users, financial and commercial websites alike [Chandrasekaran et al.06]. Phishing is an attack that exploits the end-user lack of knowledge in terms of web browser clues and security indicators and uses

similar looking emails and websites for legitimate organizations to trick people in order to reveal sensitive information [Dhamija et al.06]. The similar looking emails and websites for legitimate organizations are known as Phishing emails and Phishing websites respectively. Phishing is aimed to take advantage of the way humans interact with computers or interpret messages rather than taking advantage of the technical system vulnerabilities [Downs et al.06]. Orgill et al. [Orgill et al.04] point out that Phishing uses human emotion and manipulation to trick the victim into giving out important information.

Phishing is about other parties attempting to gain personal information such as bank details and passwords. As the Internet has become a vital medium of communication, Phishing can be performed in different ways. They are as follows:

1. email-to-email: this happens when someone receives an email asking for sensitive information to be replied to the sender email or sent to another email.
2. email-to-website: this happens when someone receives an email with embedded web address that leads to a Phishing website.
3. website-to-website: this happens when a Phishing website is reached by clicking on an online advert or through a search engine.
4. browser-to-website: this happens when someone misspelled a web address of a legitimate website on a browser and then goes to a Phishing website that has a similar address.

Wu [Wu06] explains the human interaction with Phishing attacks are as follows. When a user receives a Phishing email, they may not understand that the link provided may not take them to where they expect. For example, the user's intention may be "go to eBay" but the actual implementation of the hyperlink may be "go to a server in South Korea". This misunderstanding enables Phishing and makes it very hard to defend against. Wu called this the "semantic gap" between the user's understanding and the system model (See Figure 4). Users gain their understanding of interaction from the presentation or the way it appears on the screen. Some technical details of web pages and email messages are hidden and some of them are not understandable to most users. Thus, the user does not interpret the system clues or is unable to do so. On the other hand, email clients and web browsers follow the coded instructions and are unable to check the user's intentions. Therefore, without awareness of both models, neither the user nor the computer is able to bridge the semantic gap in Phishing attacks.

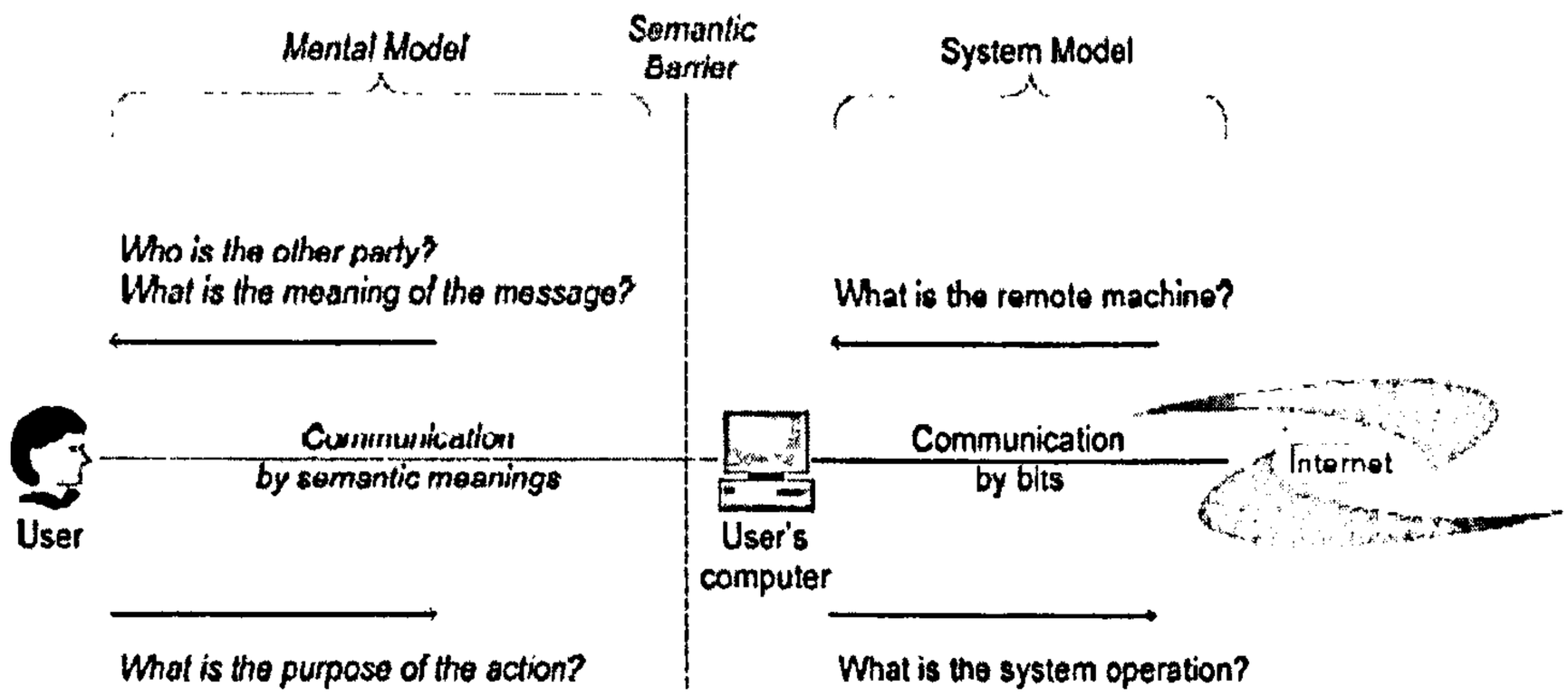


Figure 4: There is a semantic barrier between a user and his computer [Wu06]

The success of a Phishing attack lies in the fraudster's ability to craft the attack in a manner that a naïve user is unable to differentiate between legitimate and fraudulent websites and messages [Chandrasekaran et al.06].

3.3. Clues of Recent Phishing Scams

Chou et al. [Chou et al.04] described common characteristic of Phishing websites and referred to them as Phishing clues. They are as follows:

- *Logos*. The Phishing website uses logos found on the legitimate website to imitate its appearance.
- *Suspicious URLs*. Phishing websites are located on servers that have no relation with the legitimate website. The Phishing website's URL may contain the legitimate website's URL as a substring (http: //www.ebaymode.com), or may be similar to the legitimate URL (http://www.paypa1.com) in which the '*l*' letter in PayPal is substituted with number '1'. IP addresses are sometimes used to mask the host name (http://25255255255/top.htm). Others use @ marks to make host names difficult to understand (http://ebay.com:top@25525525525 5/top.html) or contain suspicious usernames in their URLs (http://middleman/http://www.ebay.com)

- *User input*. Phishing websites typically contain pages for the user to enter sensitive information, such as password, social security number and so on.
- *Short lived*. Most Phishing websites are available for only a few hours or days – just enough time for the attacker to defraud a high enough number of users.
- *Copies*. Attackers copy html from the legitimate websites and make minimal changes.
- *Sloppiness or lack of familiarity with English*. Many Phishing pages have misspellings, grammatical errors, and inconsistencies.
- *HTTPS is uncommon*. Most Phishing websites do not use https[15] even if the legitimate website does. This simplifies recognizing the Phishing website.

These characteristics are not exhaustive and an extended set is shown in Section 6.2.3 in Chapter 6.

3.4. Examples of Phishing

Symantec [Symantec04] presents a typical example of a Phishing attack as shown in Figure 5. There are many real Phishing examples collected and archived by the Anti-Phishing Working Group (APWG). One example on APWG is an attack against eBay customers that was first reported on 18th April 2005 [APWG07a]. The attack goal was to get victim's eBay and PayPal username/password, credit card information, bank account information and so forth. The email, as Figure 6 shows, sent to the customers was well designed and convincing.

[15] Https is a secured http. It uses SSL (Secure Sockets Layer) which is implemented in most commercial web servers [Hassler01, p. 269].

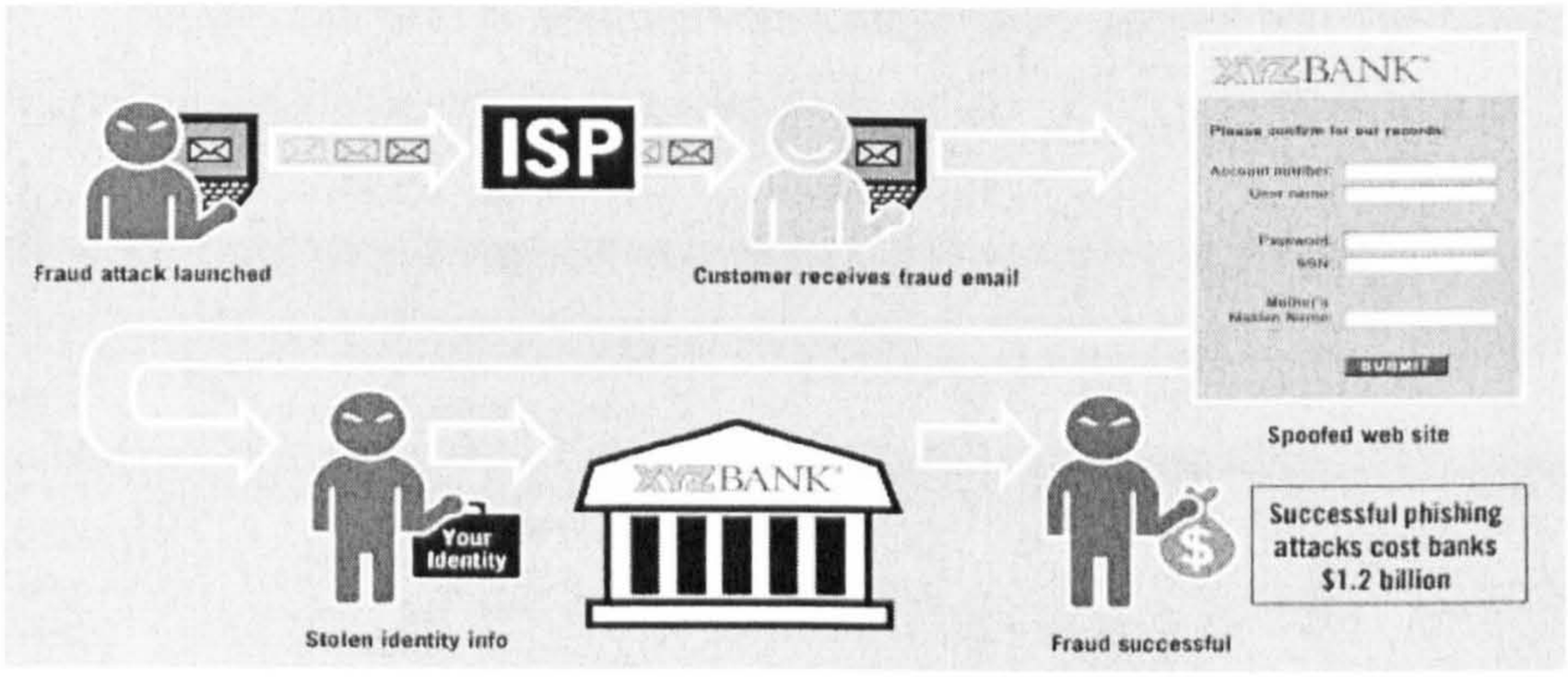
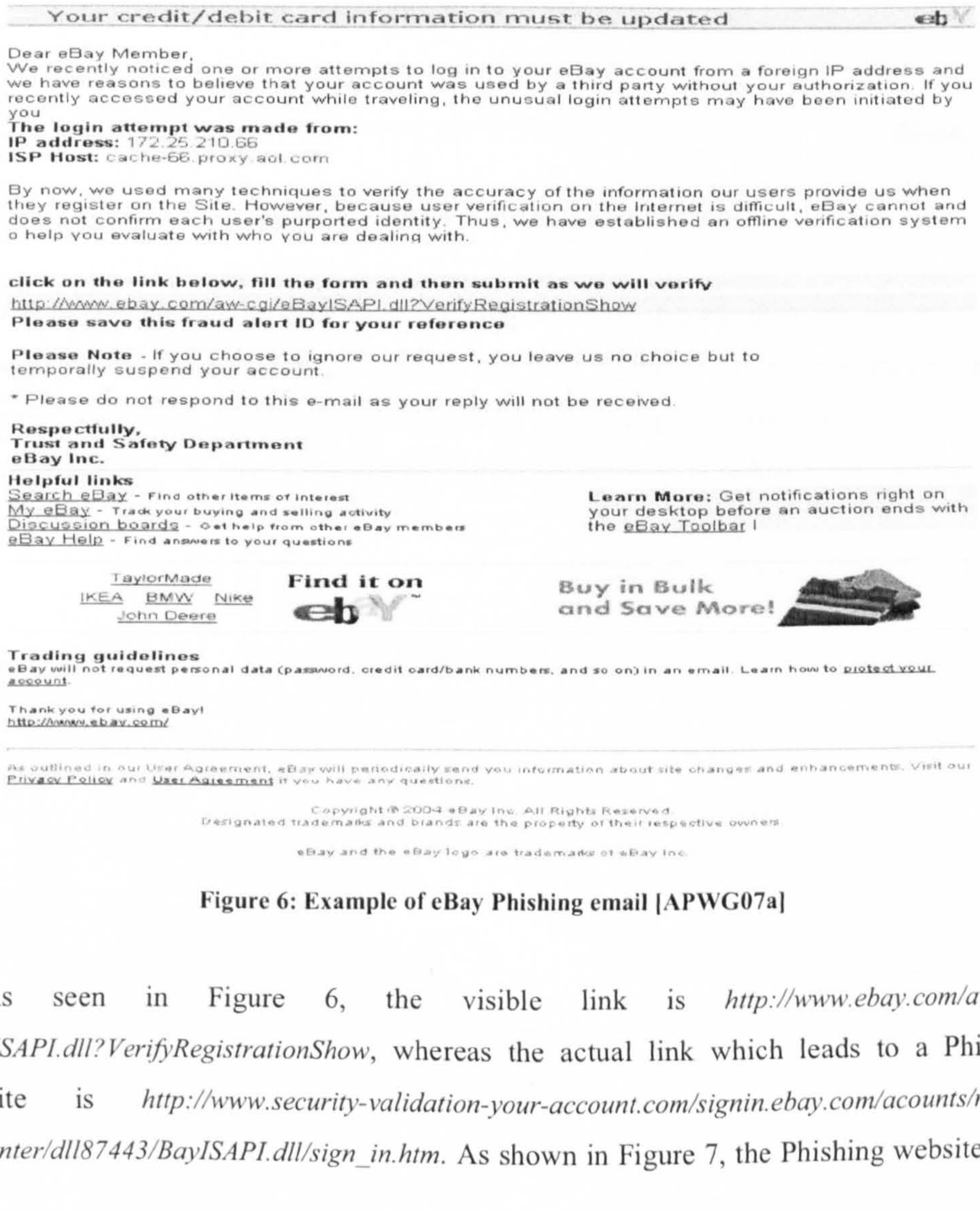


Figure 5: A typical scenario of Phishing



Your credit/debit card information must be updated

Dear eBay Member,
We recently noticed one or more attempts to log in to your eBay account from a foreign IP address and we have reasons to believe that your account was used by a third party without your authorization. If you recently accessed your account while traveling, the unusual login attempts may have been initiated by you
The login attempt was made from:
IP address: 172.25.210.66
ISP Host: cache-66.proxy.aol.com

By now, we used many techniques to verify the accuracy of the information our users provide us when they register on the Site. However, because user verification on the Internet is difficult, eBay cannot and does not confirm each user's purported identity. Thus, we have established an offline verification system o help you evaluate with who you are dealing with.

click on the link below, fill the form and then submit as we will verify
http://www.ebay.com/aw-cgi/eBayISAPI.dll?VerifyRegistrationShow
Please save this fraud alert ID for your reference

Please Note - If you choose to ignore our request, you leave us no choice but to temporally suspend your account.

* Please do not respond to this e-mail as your reply will not be received.

Respectfully,
Trust and Safety Department
eBay Inc.

Helpful links
Search eBay - Find other items of interest
My eBay - Track your buying and selling activity
Discussion boards - Get help from other eBay members
eBay Help - Find answers to your questions

Learn More: Get notifications right on your desktop before an auction ends with the eBay Toolbar !

TaylorMade
IKEA BMW Nike
John Deere

Find it on ebay

Buy in Bulk and Save More!

Trading guidelines
eBay will not request personal data (password, credit card/bank numbers, and so on) in an email. Learn how to protect your account.

Thank you for using eBay!
http://www.ebay.com/

As outlined in our User Agreement, eBay will periodically send you information about site changes and enhancements. Visit our Privacy Policy and User Agreement if you have any questions.

Figure 6: Example of eBay Phishing email [APWG07a]

As seen in Figure 6, the visible link is *http://www.ebay.com/aw-cgi/eBayISAPI.dll?VerifyRegistrationShow*, whereas the actual link which leads to a Phishing website is *http://www.security-validation-your-account.com/signin.ebay.com/acounts/memb/avncenter/dll87443/BayISAPI.dll/sign_in.htm*. As shown in Figure 7, the Phishing website does

not use a technical trick to hide its actual web address in the address bar. The lack of https session protection is also a clue for this scam.

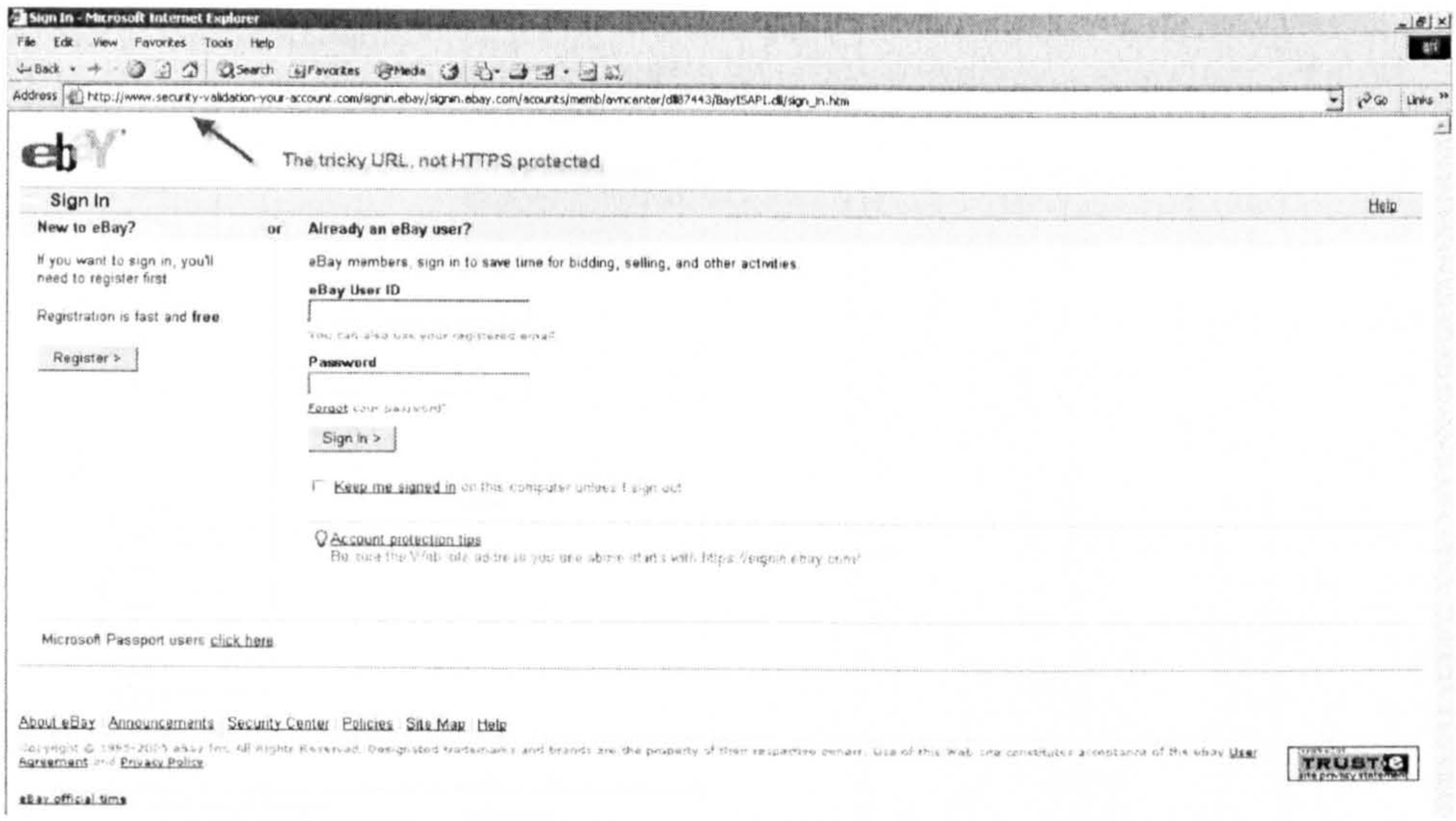


Figure 7: Example of eBay Phishing website.

When the "Sign In" button, shown in Figure 7, is clicked, then the second page is presented which requests personal information (See Figure 8). In order to make the trick more legitimate, none of the information will be checked against errors by the website. The next page is a 'logout' page (See Figure 9) then the browser is redirected to the legitimate eBay homepage.

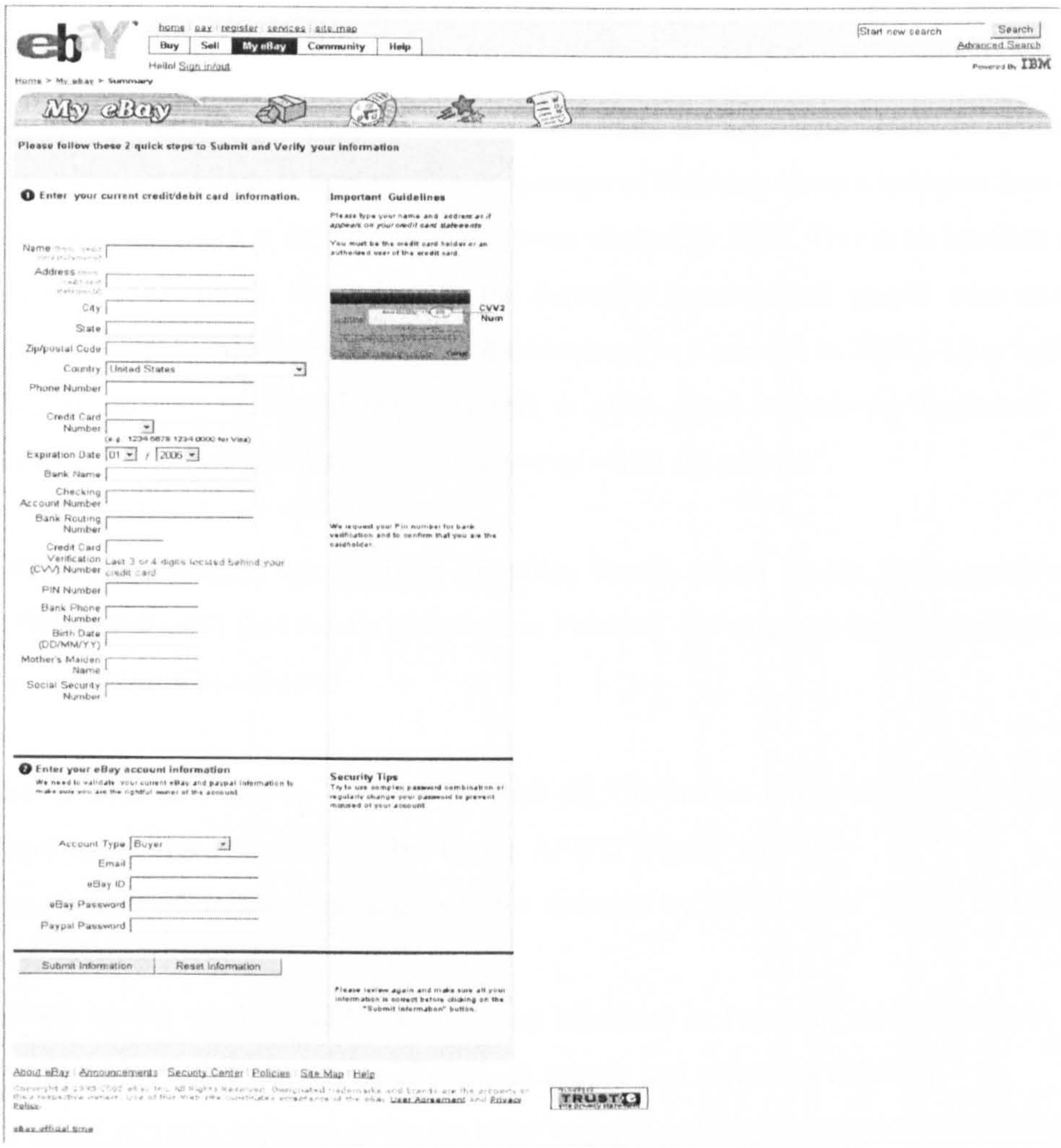


Figure 8: Example of eBay Phishing website (personal information)

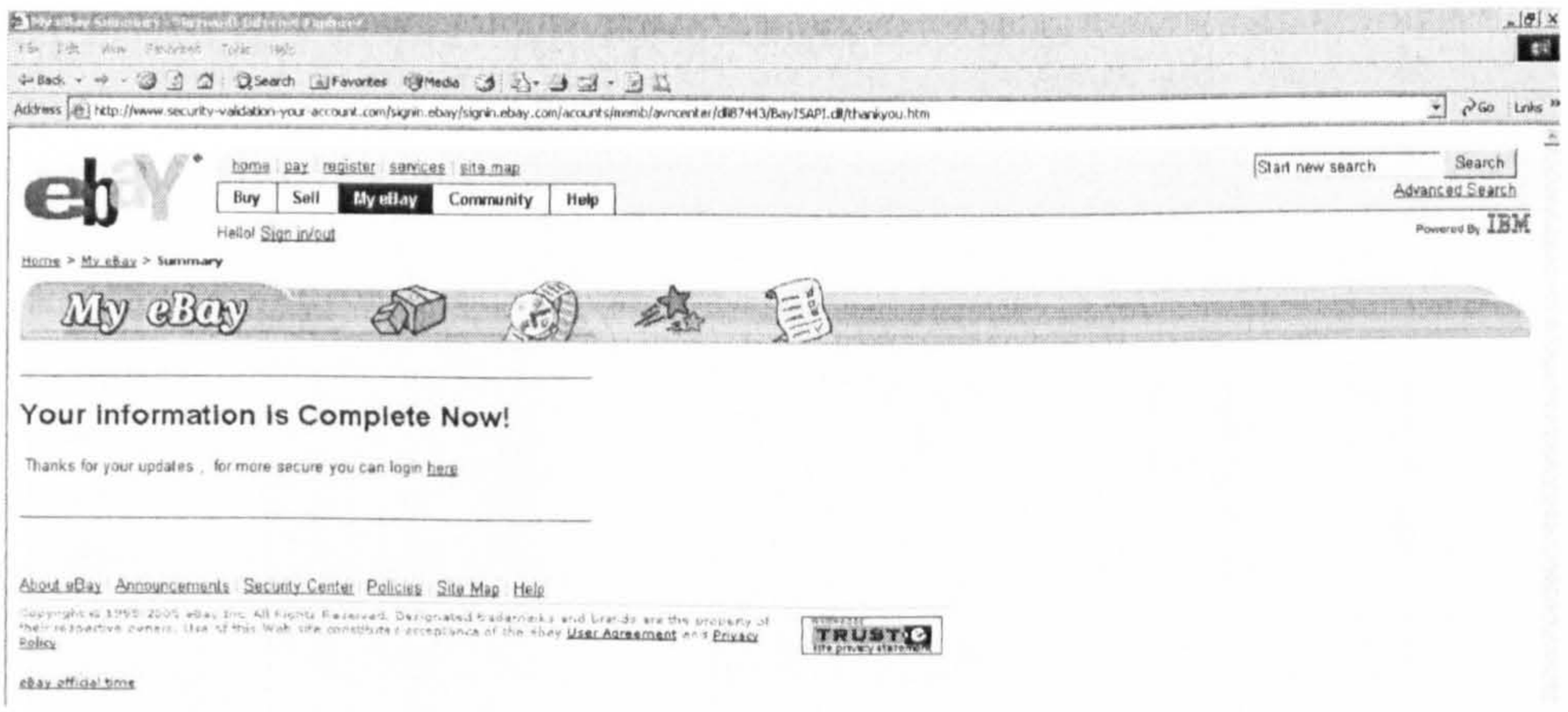


Figure 9: Example of eBay Phishing website (logout page)

3.5. Impact and Statistics

Gartner conducted a survey of 3,985 individuals in September 2008 to determine consumer Phishing trends [Litan09]. The percentage of Phishing victims is higher than ever. 5 million Internet users in the United States were victims in 2008. This is an increase of 40 percent over 2007 [ibid]. According to the survey, 4.3 percent of people who received Phishing emails lost money from the attack (compared to 3 percent in 2005). Litan believed that '*a four percent successful response rate is quite good, considering legitimate mass email marketing campaigns have a success rate of about 1.5 percent*'.

APWG has produced the Phishing Activities Trends report for the first quarter of the year 2008 [APWG08] that details statistics on Phishing activities. A summary of the main figures is presented as follows:

- The number of Phishing websites reached 30,716 unique URLs reported in February which is the high recorded number by the APWG (Figure 10).
- The number of unique Phishing websites detected by APWG was 36,002 in February 2008 (Figure 11).
- APWG saw a total of 139 brands being hijacked in February with numerous non-traditional websites such as social network portals and gambling websites.
- Financial services continue to be the most targeted industry sector at 92.4-94.2% of all attacks in over the quarter (Figure 12). Also, more international banks and brands were spoofed.
- The longest time online for a Phishing website is **31** days.

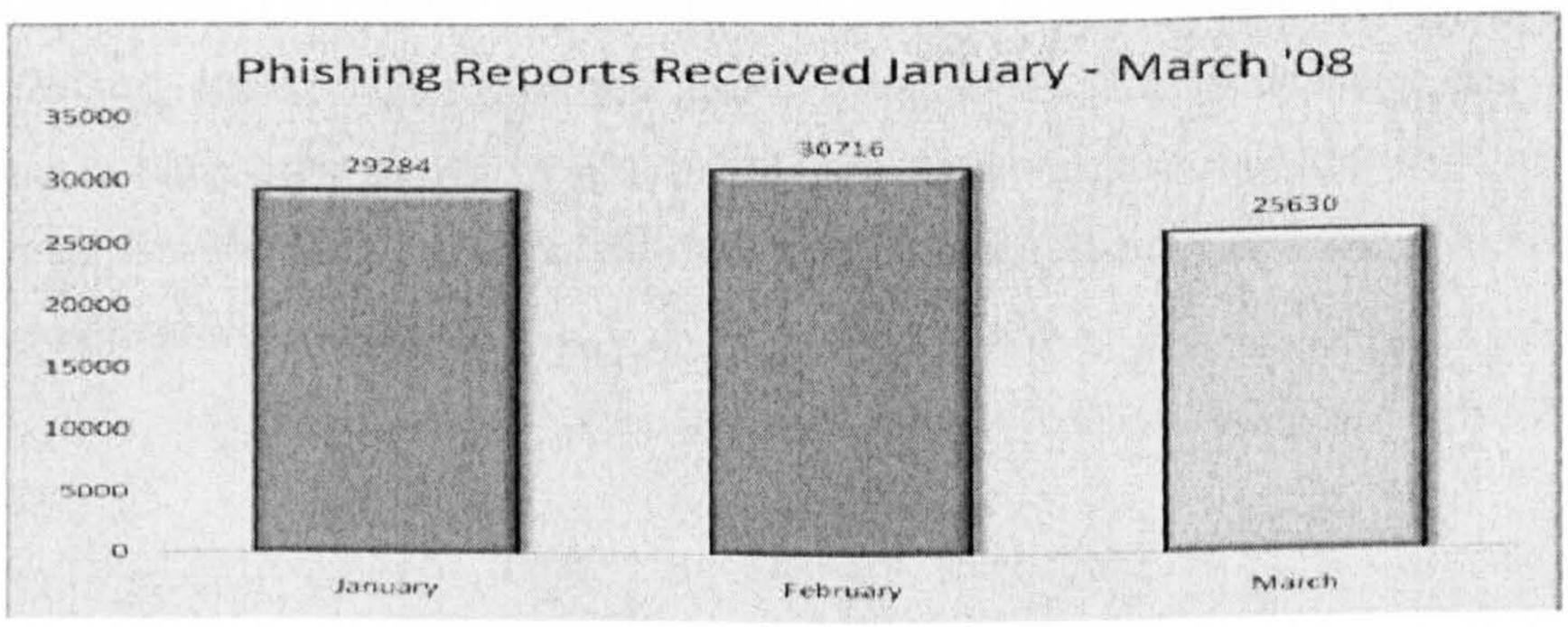


Figure 10: Phishing reports detected in the period January 2008 – March 2008

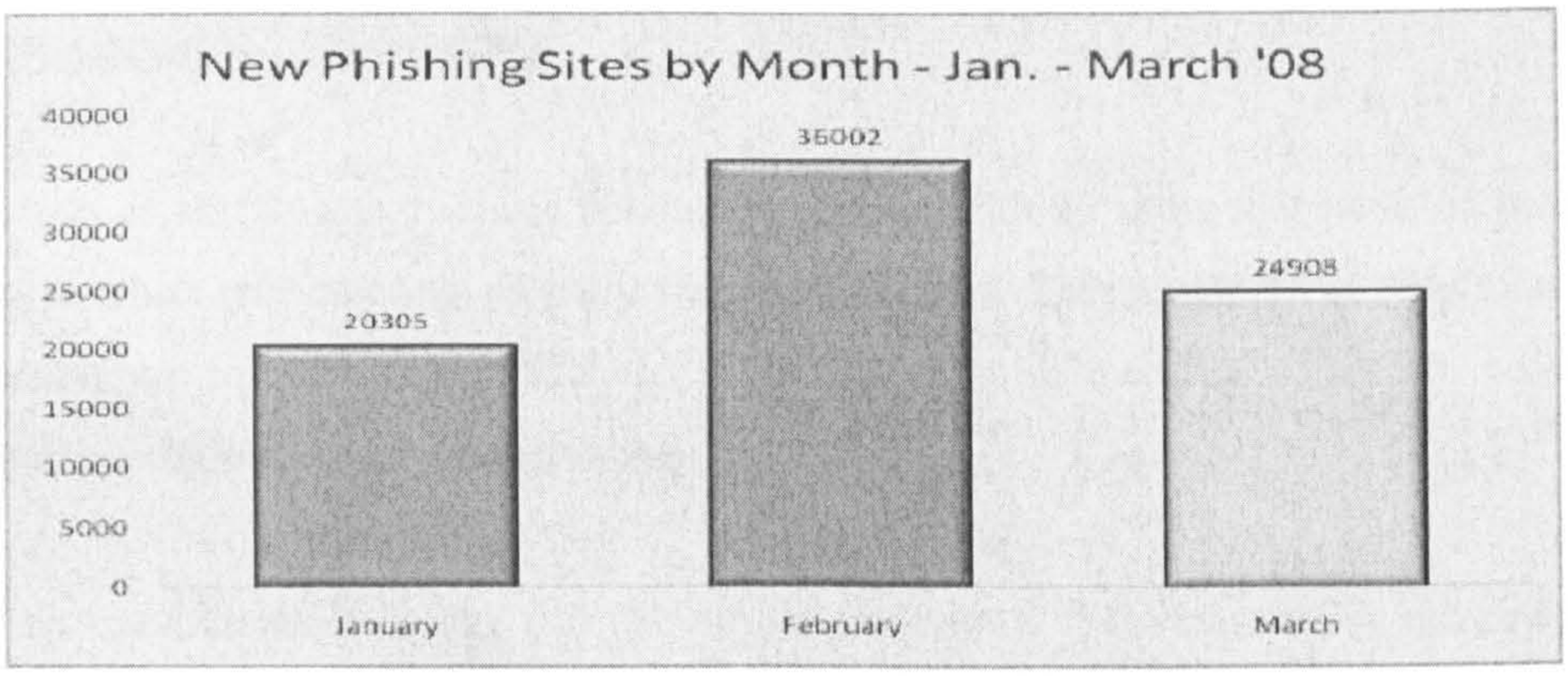


Figure 11: New Phishing websites detected in the period January 2008 – March 2008

	January	February	March
Financial Services	92.4%	94.2%	92.9%
Retail	1.5%	1.4%	1.4%
ISPs	3.8%	2.2%	1.4%
Government and Others	2.3%	2.2%	4.3%

Figure 12: Most targeted industrial sectors in the period January 2008 – March 2008

The most targeted customers by phishing are PayPal's and eBay's customers. This is because they are such a huge demographic – 123 million customers at the end of 2006 and also because PayPal is designed for transfering money around easily and thus makes it a target for fraudsters [Young07].

Phishing attacks also have a big negative impact on consumer trust about online banking and e-commerce. The Gartner group conducted a survey in 2005 [Litan05] which estimated that Phishing attacks will reduce the growth rates in the U.S. e-commerce and online banking in three years period by 1% to 3%. The survey also showed that 80% of users confirmed that Phishing has impacted their trust in email from companies or individuals they do not know personally.

3.6. Phishing Susceptibility

Downs et al. [Downs et al.06] conducted a study with 20 users that revealed peoples' strategies when encountering possibly suspicious emails. They explored several strategies that people use in evaluating emails and in making sense of warnings offered by browsers attempting to help users navigate the web.

Their research methodology was one-on-one interviews. Participants were informed that the interview was about "your computer use" and "how people make decisions while using their email and visiting websites". The interview protocol had two parts. The first part is the *email and web role play* in which participants read and responded to a set of emails. The second part was the *security and trust decisions* in which participants described their concepts related to their trust on the Internet and their awareness of online security measures.

Downs et al. selected their participants based on criteria they created to filter only those who were considered 'non-experts' in terms of computer technical ability. Their criteria for filtering their participants was that people who answered "no" to two or more of the following screening questions were included in the study:

1. whether they had ever changed preferences or settings in their web browser,
2. whether they had ever created a web page, and
3. whether they had ever helped someone fix a computer problem.

Downs et al. found that all participants had noticed different clues that they might use to decide whether an email or website was trustworthy such as false addresses in the "from" line, absence of lock icon and broken images on a website. In contrast, they did not necessarily interpret these clues correctly. For example, many of them did not know that a lock in the content of a web page was not the same thing as a lock in the browser's chrome[16]. In addition, many participants thought that the existence of broken images was a

[16] The borders of a web page window, which include the window frames, menus, toolbars, address bar and status bar [Dhamija et al.06].

problem with their computer rather than an indication about the source of the website. Fewer participants noticed clues in URLs.

Participants used different strategies to determine about the trustworthiness of email. One of theses strategies was where participants looked for emails that appeared to be for them personally. Another strategy was that participants would reply to companies that they did business with. The third strategy was where participants thought that reputable companies will send emails. Participants mostly focused on interpreting the text of the email rather than any clue in email headers or links included in the email. None of these strategies appeared to be particularly effective in helping these naive users avoid falling for scams. Participants' experience with very particular attacks seemed to be the best clues for spotting similar ones. However, this clue could not be applied to unfamiliar attacks.

Downs et al. [Downs et al.07] conducted further research in order to find out whether there are correlations between web environment experiences and the susceptibility to Phishing. This study reported a survey of 232 computer users. The survey included sections such as a URL evaluation where respondents identified features of URLs, an email role play where respondents responded to screenshots of emails and websites, past experience with websites, ratings of potential negative consequences of Phishing and a knowledge section where respondents interpreted the meaning of lock icons. They found the following:

1. Those who properly answered the knowledge question about the definition of Phishing were significantly less likely to fall to detect Phishing emails (Behavior is correlated with Phishing knowledge),
2. Knowledge about other computer risks and concepts such as cookies, spyware, or viruses was unrelated to clicking on the Phishing link (Behavior is not correlated with computer risks knowledge),
3. Participants, who correctly answered that non-chrome lock images were not the same thing as the standard lock image in chrome (See Figure 13), were less likely to fall to detect Phishing emails (Behavior is correlated with browser-security-lock knowledge),

Figure 13: Standard lock image in window chrome

4. Participants who had experience with Phishing websites were less likely than others to click on the Phishing links. Similar results were found for visiting the Phishing website and entering information there (Behavior is correlated with having experienced Phishing),
5. Participants who recognized from the URL (Table 2 shows the URLs used in the research) that the website was untrustworthy or not secure were less likely to fall to detect Phishing than others (Behavior is correlated with URL parsing knowledge) and

URLS evaluated
http://cgi.ebay.com/ws/eBayISAPI.dll?ViewItem&item=660037851
http://antwrp.gsfc.nasa.gov/apod/astropix.html
http://www.payaccount.me.uk/cgi-bin/webscr.html?cmd=_login-run
http://www.ebay.me.uk/cgi-bin/webscr.html?cmd=_login-run

Table 2: The URLs evaluated used in Downs et al.'s research [Downs et al.07]

6. Participants perceived negative consequences were unrelated to any of the behaviors relating to falling for Phishing email (Behavior is not correlated with perceiving negative consequences for Phishing).

Downs et al. concluded that deeper understanding of the web environment is associated with less vulnerability to Phishing attacks.

3.7. Solutions

There have been solutions to mitigate and reduce the risk of Phishing scams. These solutions are technical and educational.

3.7.1. Technical

There have been technical solutions to mitigate the problem of Phishing. Anti-Phishing email filters to detect and delete emails automatically at the email server. However, there is

a risk of mistakenly blocking legitimate email if the filter is configured to be sufficiently sensitive to detecting Phishing email [Emigh05]. Security toolbars have been used to prevent Phishing websites such as SpoofStick, TrustBar and SpoofGuard as Figure 14 shows. The anti-Phishing toolbars are web browser plug-ins that either detect and prevent users from reaching Phishing website or warn users when they reach a suspected Phishing website. The web browser "Internet Explorer 7" has an anti-Phishing toolbar called 'Microsoft Phishing Filter' [Microsofta]. Microsoft [Microsofta] states that '*Phishing Filter checks the sites you visit against an up-to-the-hour, dynamic list of reported Phishing sites. If it finds a match, Phishing Filter will show you a red warning notifying you that the site has been blocked for your safety*'.

Figure 14: Existing security toolbars [Wu et al 06]

Cranor et al. [Cranor et al.06a] examined the effectiveness of 10 popular anti-Phishing toolbars and found that they had many limitations. SpoofGuard was very good at identifying fraudulent websites, but it also incorrectly identified many legitimate websites as fraudulent (FP)[17]. EarthLink, Google, Netcraft, Cloudmark, and Internet Explorer 7 identified most Phishing websites correctly and had few false positives, but they still missed more than 15% of Phishing websites.

Anti-Phishing tools use two major methods for detecting Phishing websites. The first one is to use heuristics such as checking the host name and checking the URL for common spoofing techniques. The heuristics approach is not 100% accurate since it produces low

[17] A false positive takes place when a legitimate website is mistakenly judged as a Phishing website [Zhang et al.07].

false negatives (FN)[18], which implies they do not catch all Phishing websites, and high false positives (FP) [Zhang et al.07]. The second method is to use a blacklist that lists Phishing URLs verified by paid experts. When experts check a reported URL and decide that it is Phishing URL due to some Phishing clues, they add the URL to their blacklist. Blacklists have a high level of accuracy [ibid]. However, a reported website is not blacklisted until it is verified. Therefore, blacklists require verification and updates by humans. One problem here is that verification and updates consume a great deal of resources, especially time. Another problem is that unlisted and unreported Phishing URLs bypass blacklists and reaches their goal. These limitations significantly complicate the process of compiling a blacklist which then can reduce blacklists' effectiveness [ibid].

Wu et al. [Wu et al.06] carried out two experiments using three security toolbars and other browser security indicators and they found them all ineffective at preventing Phishing attacks. They also found that many subjects failed to look at the toolbars and few others noticed the suspicious signs coming from the indicators but they either did not know how to interpret the signs or they improperly explained them. In addition, they concluded that many users do not understand the Phishing attacks and do not know good practices for staying safe online.

Dhamija et al. [Dhamija et al.06] carried out research on how Phishing works. Their findings are:

- Good Phishing websites fooled 90% of participants.
- Many subjects lacked knowledge of how computer systems worked and did not understand security systems and indicators. For example, some subjects do not understand the domain name's syntax meaning and can not distinguish the deference between legitimate and fraudulent URLs (e.g. they may understand that www.ebay-members-security.com is related to www.ebay.com).
- Existing anti-Phishing browsing cues are ineffective. 23% of participants in their study did not look at the address bar, status bar, or the security indicators.

[18] A false negative takes place when a Phishing website is mistakenly judged as a legitimate website [Zhang et al.07].

- Some visual deception attacks (e.g. copying images of browser chrome or the SSL[19] indicators in the address bar or status bar) can fool even the most sophisticated users because they sometimes look like authentic indicators.

3.7.2. Training

3.7.2.1. Importance

Anti-Phishing training for end-users is complementary to any proposed technical solution. Robila and Ragucci [RobilaRagucci06] suggest that while technical improvements continue to stop the attacks, end-user training is a key component in Phishing attacks mitigation. Symantec [Symantec04] believes that '*customer education is central to helping consumers change their behavior to prevent online fraud*". Security training and awareness programs have done a good job in mitigating the risk of Phishing [Dodge et al.07].

Anti-Phishing training will make the end-user aware and an effective barrier against Phishing attempts. Furthermore, training end-users on how to detect and prevent Phishing is a strongly recommended practice. Orgill et al. [Orgill et al.04] point out user training is an important part of mitigation against Phishing attacks on information systems. Robila and Ragucci [RobilaRagucci06] believe that Phishing attacks have an extremely high success rate since they most likely appeal to the user's emotions. Accordingly, anti-Phishing training will continue to be considered and improved.

3.7.2.2. Approaches

The most basic approach is publishing guidelines for the Internet users to follow when they go online. Theses guidelines are referred as users' tips. Many financial and commercial, private and government institutions (e.g. eBay, PayPal, Amazon and HSBC) have provided anti-Phishing training tips for the Internet Users. All the information used in

---

19 Secure Sockets Layer (SSL) is a secure communications protocol [Oppliger00, p. 132].

the training approaches is based on the users' tips. There are many different tips to use. The first type is anti-Phishing tips for detecting Phishing emails. Secondly, anti-Phishing tips for detecting Phishing websites. Table 3 shows some of the anti-Phishing practices provided by APWG [APWG07b]. The aim of the tips is to train users to look for Phishing clues located in emails and websites to enable them to make better decisions in distinguishing Phishing emails and websites. Users usually need to open new web browsers and access online material published by institutions to read and then go back to their online activity browser to proceed. This scenario happens in the case that users know that there are Internet fraud attacks called Phishing and there are training materials for detecting and preventing them.

#	*Tip*
1	Phishers typically include upsetting or exciting (but false) statements in their emails to get people to react immediately.
2	Phishers typically ask for information such as usernames, passwords, credit card numbers, social security numbers, date of birth, etc.
3	Phishers emails are NOT personalized, but they can be. Valid messages from your bank or e-commerce company generally are personalized, but always call to check if you are unsure.
4	Don't use the links in an email, instant message, or chat to get to any web page if you suspect the message might not be authentic or you don't know the sender or user's handle. Instead, call the company on the telephone, or log onto the website directly by typing in the Web address in your browser.
5	Always ensure that you're using a secure website when submitting credit card or other sensitive information via your Web browser
6	Phishers are now able to 'spoof' or forge BOTH the "https://" that you normally see when you're on a secure Web server AND a legitimate-looking address. You may even see both in the link of a scam email. Make it a habit to enter the address of any banking, shopping, auction, or financial transaction website yourself and not depend on displayed links.
7	Phishers may forge the yellow lock you would normally see near the bottom of your screen on a secure website. The lock has usually been considered as another indicator of a 'safe' website. The lock, when double-clicked, displays the security certificate for the website. If you get any warnings displayed that the address of the website you have displayed does NOT match the certificate, do not continue.
8	Remember not all scam websites try to show the "https://" and/or the security lock. Get in the habit of looking at the address line, too. Were you directed to PayPal? Does the address line display something different like "http://www.gotscammed.com/paypal/login.htm?" Be aware of where you are going.

Table 3: Examples of anti-Phishing tips [APWG07b]

Kumaraguru et al.'s [Kumaraguru et al.07b] tested the effectiveness of 24 existing online training materials that teach people how to protect themselves from Phishing attacks. They collected online anti-Phishing materials such as eBay's tutorial on spoofed emails,

Microsoft's security tutorial on Phishing, Phishing E-card from the U.S. Federal Trade Commission and tutorial from MySecureCyberspace. They had two groups. They recruited 14 participants for each groups, for a total of 28 participants. To recruit participants, they filtered participants with respect to their computer technical ability. They used the same criteria that Downs et al. [Downs et al.06] used and mentioned earlier in Section 3.6. They aimed to recruit only participants who were considered "non-experts'.

Their participants spent approximately 15 minutes in reading anti-Phishing training materials and then showed good improvements in their ability to identify Phishing websites when compared to a control group. They found that '*these training materials are surprisingly effective when users actually read them*'. Then, they provided some recommendations on how to improve training materials based on these principles.

Robila and Ragucci [RobilaRagucci06] proposed a new technique for training users by combining class discussions and Phishing IQ tests. They included Phishing topics in an Introduction to Computing course aimed at students studying a non-computer science subject. Robila and Ragucci have built a training tool for users that uses Phishing IQ tests. The tests included displaying both legitimate and fraudulent emails to users and having them identify the Phishing attempts from the legitimate emails. Then, the tool gives a score for each user and feedback. Robila and Ragucci concluded that '*class assessment indicates an increased level of awareness and better recognition of attacks*'.

Anandpara et al. [Anandpara et al.08] argue that Phishing training using IQ tests seems to affect the users' judgment and then increase their fear because it makes them suspicious rather than improving their ability to recognize Phishing from legitimate email. They conducted a study where 40 subjects were asked to answer a selection of questions from existing Phishing IQ tests. They excluded subjects who have unusual knowledge about computer science or security. They also included subjects who either use or would consider using online shopping, banking or bill paying. Their experiment was divided into three parts. The first part was that they gave subjects a short IQ test which contained five different emails and asked them to identify Phishing emails. Then, the second part was that the subjects were asked to read existing Phishing training. The third part was that subjects were asked to take a second Phishing IQ test, with the same design as the first one, but with

different emails. As a result of their study, Anandpara et al. found that '*the number of times a subject labels an example as Phishing does not depend on the number that actually are Phishing*'.

Many commercial institutions, such as Microsoft (See Figure 15), periodically send email security information to help their customers in protecting their online security [Microsoftb]. This email provides practical security tips, useful resources and links, and a forum to ask security-related questions.

Microsoft states that the email is a suitable way for customers to stay up to date on the latest issues and events with:

- Security tips including anti-Phishing tips.
- Security critical updates.
- Answers to frequently asked questions (FAQs) on security topics.
- Information about security trials and downloads.
- Tips from security team for home users.

These emails are known as 'anti-Phishing emails' if they include anti-Phishing tips. These emails are usually sent in text and HTML formats. Customers who are interested in receiving these emails need to subscribe with the commercial institutions (i.e. anti-Phishing emails providers) in order to be included in receiving them.

Figure 15: An example of anti-Phishing email

Kumaraguru et al. [Kumaraguru et al.07a] also provided an approach which focuses on teaching people about the risks of Phishing and training them to recognize and avoid email-related Phishing attacks. They developed an embedded training approach that teaches people how to protect themselves from email-related Phishing attacks during their normal use of email. Kumaraguru et al. point out that their approach based on periodically sending users fake Phishing emails that are actually from their system rather than from a fraudster. If users fall for a fake email and clicks on a link, an intervention is displayed that provides feedback about what simple steps users could follow to protect themselves. Kumaraguru et al. [Kumaraguru et al.07a] designed two email-related training interventions. The first intervention used multimedia (screenshot of Phishing email). Figure 16 illustrate the screenshot based intervention used in the approach. Also, the researchers used comic strip format to design the second training intervention. The comic strip design gave a short story explaining how fraudsters work and how the user could do simple things to avoid Phishing emails.

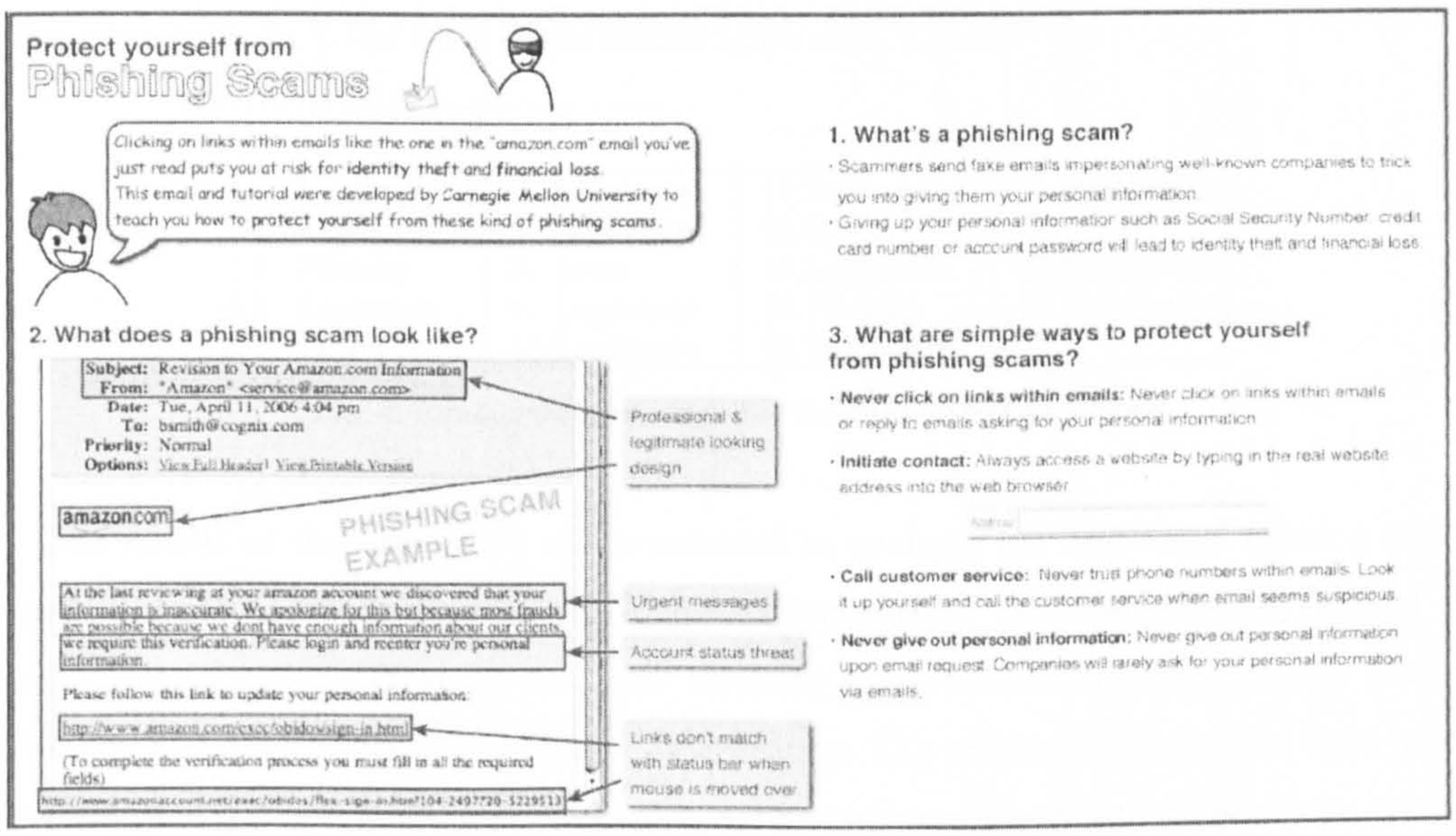


Figure 16: An intervention includes text with an annotated image of the training email approach

They compared their two training interventions (screenshot and comic strips) with the current practice of sending anti-Phishing tips. They evaluated their approach by conducting a user study using three groups. Two groups had screenshot and comic strips interventions. The third group had anti-Phishing email. There were 10 participants in each group for a total

of 30 participants. They included participants in their evaluation experiments with little technical knowledge. They recruited only 'non-experts' in terms of their computer technical ability. They used the same criteria used in Downs et al.'s [Downs et al.06] and Kumaraguru et al.'s [Kumaraguru et al.07b] studies which were mentioned earlier in Section 3.6 and this section respectively.

Participants played the role of an imaginary person called 'Bobby Smith'. Participants were not told that the experiments were about Phishing. However, they were told that the study investigated "how people effectively manage and use emails." They were told that they should interact with their email the way they would normally do in their real life. Each participant was shown 19 email messages. The emails arranged in a predefined order. Nine messages were emails that Bobby Smith received from his work, friends and family. Two emails were legitimate emails from organizations with which Bobby Smith had an account such as Amazon and Paypal. Two spam emails, four Phishing emails, and two training emails (anti-Phishing email or embedded training interventions). Table 4 shows the email arrangement shown to the users in the Kumaraguru et al.'s study.

1. Legitimate	6. Legitimate	11. Intervention	16. Phishing
2. Legitimate	7. Legitimate	12. Spam	17. Phishing
3. Phishing	8. Spam	13. Legitimate	18. Legitimate
4. Legitimate	9. Legitimate	14. Phishing	19. Legitimate
5. Intervention	10. Legitimate	15. Legitimate	

Table 4: Email arrangement in the Kumaraguru et al.'s study

The results of the user study that conducted to evaluate the embedded training email system shows that both training interventions (screenshot and comic strips) helped in teaching people about Phishing and how to avoid email-related Phishing attacks. Comic strip intervention was the most effective intervention. The training interventions were more effective than the current practice of sending online training materials to users.

In August, 2008, the APWG and Carnegie Mellon CyLab launched the "*Phishing Education Landing Page Program*" [PEI08]. The program's idea is simple. It redirects users who have clicked on links in Phishing email or otherwise to training materials that explains that they have just fallen for a Phishing attack and advises them on how they avoid it in the future. The goal of this program is to train users in online security at the "most teachable

moment" when they have just clicked on a link in a Phishing communication. The project authorities encourages '*all brand owners to approve this process, all takedown providers to request the use of this redirect scheme, and all ISPs, registrars, registries, etc. to redirect to this page instead of serving an error page*'.

The Phishing Education Landing Page Program works as follows: as part of the process for shutting down a Phishing website, the project authorities ask ISPs, registrars, and persons who have control of the Phishing page to take the following steps:

- Check whether the brand being attacked has approved having the Phishing website URLs re-used to redirect their users (who have fallen) to training page (i.e. a webpage to educate users about Phishing).
- If the redirection has been approved, instead of showing an error page when a user arrives at the URL, redirect them to the APWG/CMU Phishing Education Landing Page (See Figure 17).

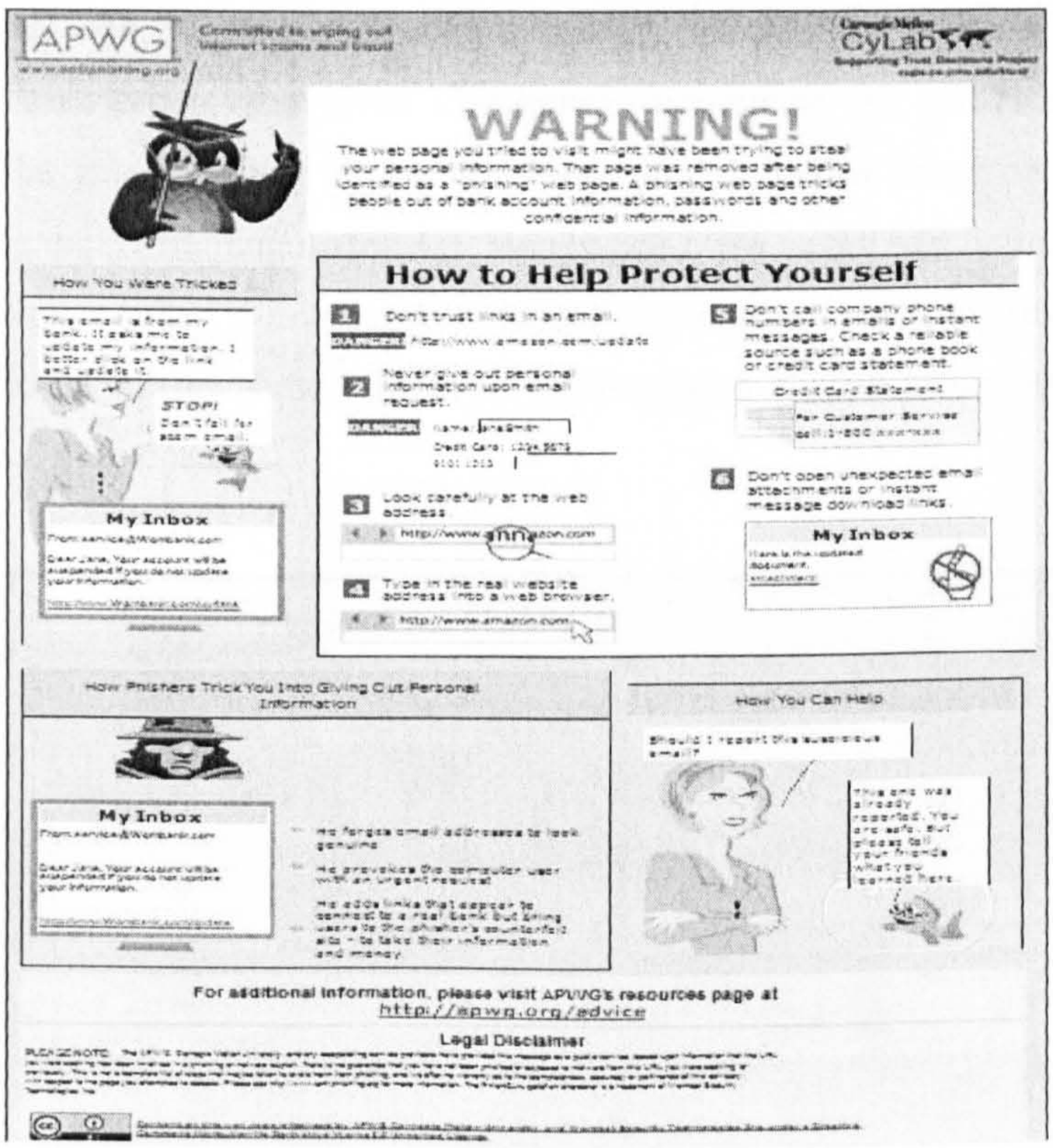


Figure 17: The APWG/CMU Phishing Education Landing Page [PEI08]

An online game (See Figure 18) was proposed in order to teach users good habits to help them avoid Phishing attacks [Sheng et al.07]. Sheng et al. described the story of the game as '*the main character of the game is Phil, a young fish living in the Interweb Bay. Phil wants to eat worms so he can grow up to be a big fish, but has to be careful of phishers that try to trick him with fake worms (representing Phishing attacks)*'.

The game was designed and evaluated through a user study. They included participants who are considered as 'non-experts' in terms of computer technical ability. They used the same criteria used in other Phishing experiments [Downs et al.06, Kumaraguru et al.07a, Kumaraguru et al.07b]. There were three groups. There were 14 participants in each group for a total of 42 participants. Each participant was given the scenario that they have received an email that asks them to click on one of its links. Then, they imagine that they clicked on the link to see if it is a legitimate website or a Phishing one. After that, participants were presented with ten websites and were asked to decide whether a website was legitimate or Phishing. Participants also were asked to tell how confident they were in their decisions (on a scale of 1 to 5, where 1 means not confident at all, and 5 means very confident). Then, participants in each group were given 15 minutes to complete one anti-Phishing training task (playing the game, reading an anti-Phishing tutorial created based on the game, or reading existing online training materials). Finally, the same as the part before the training, participants were presented with ten more websites to decide on. The study revealed that the participants who played the game were better in identifying Phishing websites compared to other conditions' participants.

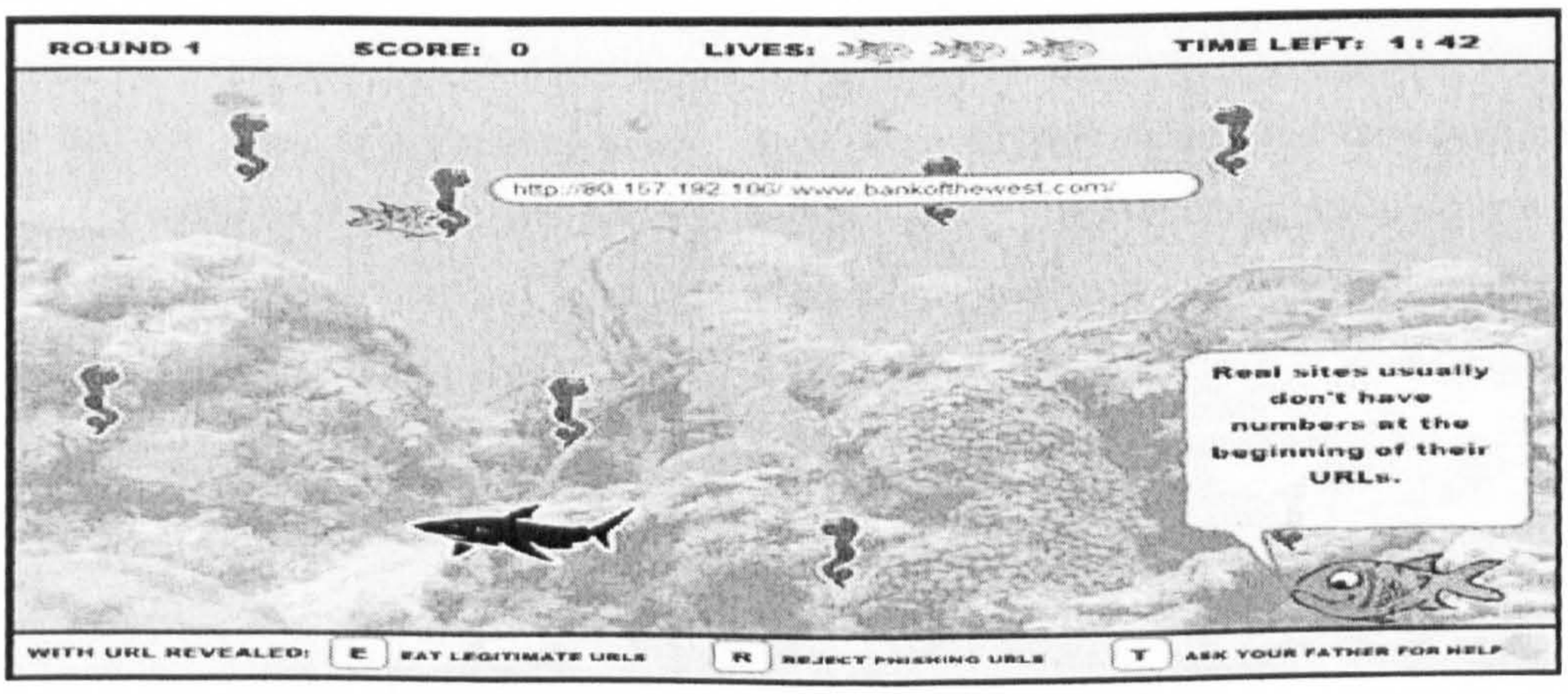


Figure 18: Anti-Phishing training game screen

Some training media using a comic-book format for online fraud have been developed [Jakobsson07]. As illustrated on Figure 19, the comic-book format approach shows some common risks and the users' thoughts about them as well as some advices.

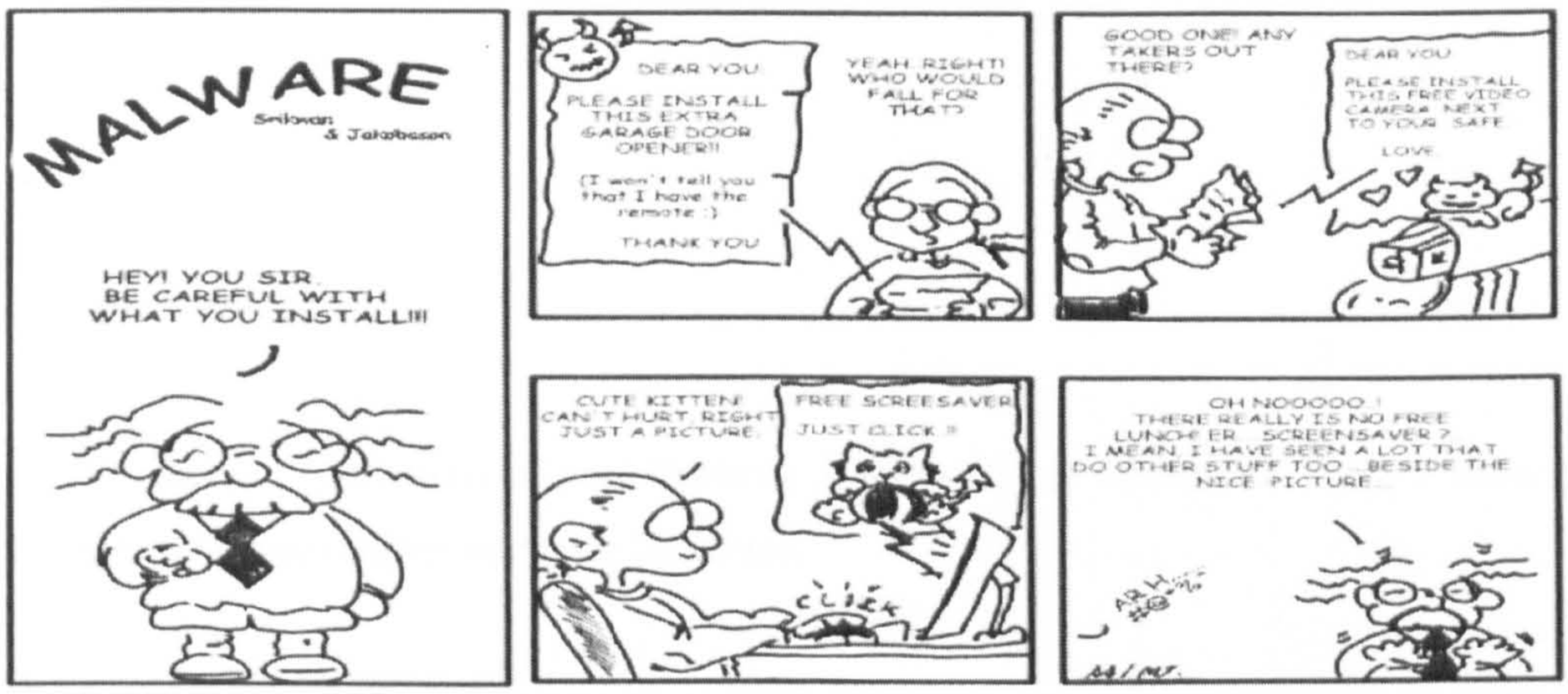


Figure 19: Comic-book format for anti-fraud end-user education [Jakobsson07]

3.7.2.3. Anti-Phishing Knowledge Retention

Previous study tested users immediately after training. Kumaraguru et al. [Kumaraguru et al.07a] designed and showed that embedded training improved users' ability to identify Phishing emails [Kumaraguru et al.07c]. Kumaraguru et al. [Kumaraguru et al.07c] tested users to find out how well they retained knowledge that was received through embedded training and how well they transferred this knowledge to detect other types of Phishing emails. They recruited people who did not know what Phishing was. Participants were not told that the study is a Phishing study. There were 42 participants and they had been randomly assigned to one of three groups: an "embedded" group in which participants were presented the training material when they clicked on links in the Phishing emails, "non-embedded" group in which participants were shown training materials in an email message and "control" group did not receive training materials but received an email from a friend. The study was carried out in two sessions separated by at least 7 days (mean = 7.2).

They found that (a) participants learned more when the training materials were presented after they clicked links on Phishing email (embedded training) than when the training

materials were sent by email (non-embedded training); (b) participants retained more knowledge when trained with embedded training than when trained with non-embedded training; (c) participants transferred more knowledge about how to avoid Phishing emails when trained with embedded training rather than when trained with non-embedded training.

3.8. Discussion

In this section, limitations of anti-Phishing approaches are discussed. Table 5 shows a summary of the approaches and their limitation.

Kumaraguru et al. [Kumaraguru et al.07b] evaluated the effectiveness of 24 existing online training materials that teach people how to protect themselves from Phishing attacks. They found that 'these training materials are surprisingly effective when users actually read them'. Their participants spent approximately 15 minutes reading anti-Phishing training materials and then showed good improvements in their ability to identify Phishing websites when compared to a control group. However, this research did not consider the effectiveness of the users' tips themselves. It did not examine the effectiveness of each individual tip. Therefore, there is need to examine the effectiveness of the most common users' tips for detecting and preventing Phishing websites individually. The effectiveness of each individual tip will be assessed and then the tip will be ranked accordingly. The aim is to identify the most effective anti-Phishing tips that users can focus on to detect and prevent Phishing attacks by themselves.

There have been technical (e.g. toolbars) and training (e.g. tips) approaches to mitigate Phishing. Regarding the training approaches, users usually need to open new web browsers and access online training material to read. Then they go back to their online activity browser to proceed. This scenario happens in the case that users know that there are attacks called Phishing and there are training materials that help in preventing them. Therefore, if the users know nothing about Phishing and anti-Phishing training materials, they are unlikely to access the training materials provided. People do not read anti-Phishing online training materials although they are surprisingly effective when users read them

[Kumaraguru et al.07b]. Moreover, an online game was proposed in order to teach users good habits to help them avoid Phishing attacks [Sheng et al.07]. The game presents anti-Phishing information in an enjoyable way. However, the disadvantage of this approach is the same as the online training materials. Users must have an idea about Phishing in advance in order to access and play the game. Also, there are anti-Phishing training courses such as IQ tests and class assessments [RobilaRagucci06]. The courses explain to users what Phishing attacks are and how to prevent them. The disadvantage of the courses' approach is that typically people are unlikely to attend them.

Many commercial institutions, such as Microsoft [Microsoftb], provide a service that periodically sends anti-Phishing emails that warn people from Phishing emails and websites. The emails provide tips for people to help them detecting Phishing emails websites. However, only subscribed customers can receive these emails.

Kumaraguru et al. [Kumaraguru et al.07a] considered training people about Phishing email during their normal use of email. Their aim was to teach people what Phishing clues to look for located in emails to make better decisions in distinguishing Phishing emails. They found that email training approach works better than the current practice of publishing or sending anti-Phishing tips. However, Kumaraguru et al.'s approach does not consider teaching people with Phishing website-related tips. Phishing websites can be reached via various methods in addition to emails such as online advertisements and typing their web addresses in a web browser. Therefore, helping users on how to make correct decisions in distinguishing Phishing and legitimate websites during their normal use is required.

Several approaches were evaluated using user experiments that involved participants who were recruited based on their technical abilities [Downs et al.06, Kumaraguru et al.07a, Kumaraguru et al.07b, Sheng et al.07]. Participants were classified into 'experts' and 'non-experts' users based on pre-study screening questions. Technical ability was judged on whether the participants had changed preferences or settings in their web browser, created a web page, and helped someone fix a computer problem. The participant who said 'no' to at least two of the screening questions was selected to take part in their experiments. This technical ability assessment was used to recruit low technical people (they called them non-experts) in the previous studies. Participants who were technically considered non-experts could know about Phishing and how to detect attacks before participating in the evaluation

experiments. Having participants with Phishing knowledge in advance may provide biased results in anti-Phishing approaches' evaluation experiments. This is because people who know about Phishing before participating in the evaluation experiments may use their prior knowledge rather than the anti-Phishing approaches that are being tested in the evaluation. Downs et al. [Downs et al.07] studied whether there are correlations between some web environment experiences and the susceptibility to Phishing. They found that people who correctly answered the knowledge question about the definition of Phishing (i.e. Phishing aware people) were significantly less likely to fall to detect Phishing emails. Low technical users (i.e. non-experts) may be Phishing aware and high technical users (i.e. experts) may be Phishing unaware.

An investigation on the effects of technical ability and Phishing knowledge on Phishing websites' detection is required. This clarifies whether the previous screening questions for recruiting low technical users in evaluating anti-Phishing approaches are beneficial. The investigation assesses using Phishing knowledge in the screening questions to recruit participants. If the results of the investigation show that (i) there is no effect for technical ability on Phishing detection and (ii) there is an effect for Phishing knowledge on Phishing websites detection, then there is need to make sure that the participants do not know about Phishing regardless of their technical ability level in evaluating the effectiveness of a new anti-Phishing approach.

Research	Approach	Participants Recruitment Criteria	Limitation(s)
Kumaraguru et al. [Kumaraguru et al.07b]	Evaluating the effectiveness existing online anti-Phishing materials.	Participants technical ability (non-experts were included)	1. Examining the effectiveness of each individual tip was not carried out. 2. Including non-experts without testing their Phishing knowledge in experiments may produce biased results.
Financial and commercial institutions	Anti-Phishing tips for end-users.	N/A	People in general do not read anti-Phishing online training materials.
Sheng et al. [Sheng et al.07]	Anti-Phishing online game.	Participants technical ability (non-experts were included)	1. People in general do not read anti-Phishing online training materials. 2. Including non-experts without testing their Phishing knowledge in experiments may produce biased results.
Robila and Ragucci [RobilaRagucci06]	Anti-Phishing IQ tests and class assessments.	Non-computer science students.	Typically people are unlikely to attend them
Kumaraguru et al. [Kumaraguru et al.07a]	Anti-Phishing embedded training for detecting Phishing emails.	Participants technical ability (non-experts were included)	1. The approach does not consider training people for detecting Phishing websites. 2. Including non-experts without testing their Phishing knowledge in experiments may produce biased results.
Microsoft [Microsoftb]	Anti-Phishing email.	N/A	Only subscribed customers can receive the emails.

Table 5: Summary of anti-Phishing approaches and their limitations

3.9. Summary

This chapter described and considered Phishing attack. Its definition, clues and some examples were shown. Figures about its negative impact on the e-commerce and online banking sectors were described. The chapter reviewed the existing research in suitability to Phishing risks as well as existing approaches in detecting and preventing Phishing emails and websites. The chapter finished with a discussion on limitations of anti-Phishing approaches.

4. Training

4.1. Introduction

This chapter presents an overview of training definition and methodologies. It also presents an overview about embedded training and discussion on its definition, advantage and examples. Retention of knowledge obtained from training is discussed together with the facts that may affect the retention rate.

4.2. Training Definition

The term 'training' is defined as '*a planned process to modify attitude, knowledge or skill behavior through learning experience to achieve effective performance in an activity or range of activities. Its purpose, in the work situation, is to develop the abilities of the individual and to satisfy the current and future manpower needs of the organization*' [KenneyReid86]. Harrison also defined training as a systematic process in which a person is helped to understand defined tasks or areas of skill and knowledge to pre-determined standards [Harrison88, p. 5].

4.3. Training Methods

Training, as a process, can be run through a number of methods. They are as follows [Coffield et al.04, ReadKleiner96, Wilson00, ShuHsiu02]:

1. Lecture.

The trainees gather in a classroom and are given a lecture. The lecture is a traditional method of training and is the most used of all methods despite its limitations [Wilson00]. The lecture alone is a poor training method unless it has good trainees' involvement and valuable feedback to them [ReadKleiner96]. A good possible way to have effective training through lectures is to stop the lecture periodically and ask the trainees to draw conclusions from the information presented. The conclusions should be related to the objectives of the training.

There are assumptions that the lecturer relies on. One of them is that participants are motivated to learn. Another assumption is that that the lecturer can have the attention of the majority of the trainees [Wilson00].

2. Training Manual.

This method involves reading reference material. The material should relate to the topic being studied. Training manuals may include self-assessment questions, progress tests or summaries [Wilson00].

3. Case Studies.

A case study can bring strong realism into the training process. Usually, a case study includes the description of a real problem and leaves the solution of the problem to be developed by the trainees. The problem description may involve the facts needed to create a solution [ReadKleiner96].

4. Cooperative (group) training.

Basically, group training is a method of collaborative learning. Generally, collaborative learning can help trainees to make progress by the activities in which they engage. If the trainees have opportunities to interact with their instructors and other trainees about the

instruction or content, then they have opportunities to build their own knowledge. Trainees also can share their own knowledge with others [ShuHsiu02].

5. Brainstorming.

Small groups try to create new ideas and attempt to answer a problem. They usually use a blackboard or whiteboard. All ideas or solutions for problems should be noted whether they are useful or not useful. Groups' members train to think differently. They also increase confidence in generating ideas. The brainstorming method helps to generate creative ideas under informal conditions [Wilson00].

6. Problem-solving training.

Trainees need to go through steps to perform this method. They need to define a problem. Then, they need to generate data about the problem. After that, trainees need to generate ideas or other courses of action to solve the problem. The three steps can be all done using brainstorming. Then, they need to choose a solution by voting or ranking (with or without criteria). Finally, the trainees are required to implement the solution or decision voted or ranked in the last step [Wilson00].

7. Demonstration.

Demonstration is effective training method because participants use all their senses. It brings alive whatever points the trainers are trying to make. Trainees can experience the idea or technique that they are trying to gain. There are guidelines, for trainers who consider applying the demonstration method, to achieve the most of it. They are careful preparation, explaining the purpose of the training, step-by-step demonstration and providing the opportunity for trainees to practice [HartCrisp91, p. 51].

8. Learning by experience

Learning by experience (it is known as experiential learning) theory defines learning as '*the process whereby knowledge is created through the transformation of experience. Knowledge results from the combination of grasping and transforming experience*' [Kolb84, p. 41]. The unique feature of experiential learning is that the experience of the learner is central place in all considerations of learning [Anderson et al.00]. This experience may involve earlier events in the life of the learner, current life events, or those coming from the learner's participation in activities implemented by teachers

[ibid]. The development process of experiential learning is that learners analyze their experience by reflecting, evaluating and reconstructing it in order to draw meaning from it based on prior experience [ibid].

9. Games.

The use of games is popular. They usually involve competition between trainees as individuals or groups. Wilson [Wilson00] states that '*games are an experiential learning activity governed by rules, entailing a competitive situation with winners and losers*'. Furthermore, the use of simulation games, i.e. a reality-based game, is more widespread due to that they can make fun. People are highly motivated and more likely to participate in training when they have a good time. Games also are useful because they can deliver more than one idea at a time [ReadKleiner96, Wilson00].

10. Simulation-based training.

Simulation is defined as '*a false assumption or display, a surface resemblance or imitation, of something*' [OED]. Simulation-based training makes the skills given by trainers more real to the trainees. Kozlowski et al.01 [Kozlowski et al.01] points out that '*practice is central in simulation-based training, since having trainees practice the skills that are the target of training services serves the purpose of making the skills more "real" to the trainees, rather than leaving them in the abstract, lecture-based domain*'. Simulation-based training provides a good opportunity for trainees to be involved in practical experience (by doing). Practice is an important factor that positively affects the training knowledge retention as discussed later in Section 4.6.2.

11. Computer-based training.

Computer-based training is classified into two groups. They are computer-assisted instruction and computer-managed instruction [ReadKleiner96]. Regarding computer-assisted instruction, training takes place during an interaction between the trainee and the computer which acts as a tutor. The computer asks questions and the trainee responds to them by typing on the keyboard. Then, the information is presented via the monitor. The disadvantage of computer-assisted instruction is that it is time-consuming because each trainee needs one computer.

With regards to computer-managed instruction, the training takes place off-line. The computer allocates each trainee different personalized instruction modules that are completed away from the computer. After completion, the computer evaluates the trainees and dives recommendations in the areas of weakness, and gives additional tasks if needed. The advantage of managed instruction is that trainees spend less time online so a single terminal may be used by many trainees. This can significantly reduce the cost of the training programme [ReadKleiner96].

12. One-on-one instruction.

One-on-one instruction is classified into two methods. They are on-the-job training and off-the-job training. On-the-Job training is any training that occurs while the trainee is actually working. The trainee is doing work in the real work environment under normal working conditions. On-the-Job training ensures that skills achieved from training can be transferred to the job. The other method is off-the-job. It refers to any training that is performed away from the trainee's work area [ReadKleiner96].

13. Role Plays.

Role playing training implies that the trainees act and plays certain roles in the context of a situation that is applicable to the training objectives. Role playing also provides a good opportunity for trainees to be involved in practical experience (doing). It is very useful in gaining insight into the feelings and viewpoints of others. Role play is limited to training situations in which mistakes are treated with tolerance by both the trainer and trainees. This is to make sure that reinforcement is mostly positive [ReadKleiner96]. Role play method is an active version of the case study method and is designed to represent the real world [Wilson00].

14. Training through practice.

This method means any training that is performed and acquired through practice. Knowledge and skills are strengthened through practice (by doing) [Anderson93].

These training methods are common and being used in many areas of skill. No single training method is better to all others. When possible, it is best to pick a method that satisfies two important activities. They are as follows:

- Encouraging active participation by the trainee and

- Providing adequate feedback [ReadKleiner96]. The descriptive feedback should deliver to the trainee what behaviors they did or did not do, what facts or concepts they did or did not learn and what results they did or did not achieve [Kozlowski et al.01].

This increases the likelihood that what is given in training will be later retained and applied. Trainees will also retain more and be more willing to learn if training is followed by positive reinforcement such as praising the trainee and the trainee's internal sense of satisfaction that comes from learning something new [ReadKleiner96].

4.4. Embedded Training

In this section, the embedded training concept is presented. The section discusses the definition of embedded training, the advantage of applying embedded training and an existing example of applying embedded training.

4.4.1. Definition

Embedded training is a training that has the ability to train a task or a skill using the associated operational system including software and machines that people normally use [Kirkley et al.03]. When using embedded training, training is not a separate activity but it is an ongoing activity that is an integral part of the workplace and its system [Kozlowski et al.01].

4.4.2. Advantage of Embedded Training

When training materials incorporate the context of the real world, work, or testing situation, training will be most effective [Anderson et al.96]. Kozlowski et al. point out that many skills and basic knowledge can be acquired in conventional training environments (i.e.

classrooms). However, they can be fully developed and refined in the actual performance environment or very close approximation to it (by practice or doing) [Kozlowski et al.01]. This means that training systems must either (a) push training toward long-term exposure of integrated teams to a multiplicity of task situations in high fidelity and full mission simulation or (b) move more training to the performance context to improve acquisition to key skills and tasks which also can be integrated with suitable instructional support systems (i.e. embed training in the workplace) [ibid]. This means that the training occurs in the trainee actual work.

One of the positive factors that is involved in embedded training is practice. Kozlowski et al. [Kozlowski et al.01] states that '*within the training context, one of the most obvious tools available to trainers is practice'*. Practice is an effective factor that is essential to some training methods such as training through practice as mentioned before.

4.4.3. Examples of Applications that Used Embedded Training

Embedded training has been widely used. Embedded training has been used in the training of military personnel on new Future Combating Systems (FCS) [Kirkley et al.03]. They developed an instructional methodology called problem-based embedded training (PBET). PBET enables designers to create simulated mixed and virtual reality tasks that are able to meet certain training objectives. To validate the methodology, they conducted a heuristic evaluation with five experts in the US military training and instructional design. They found that PBET matches training contexts as closely as possible with real world situations and scenarios and it supports training is just-in-time.

Another example of the use of embedded training is the application of the Advanced Embedded Training System (AETS). AETS uses intelligent tutoring systems (ITS) technology to improving tactical training quality and reducing the need to human in training. AETS is used in one of the USA's Navy's projects. Embedded training is a good choice for applying ITS because it allows the ITS to train in the actual work environment and eliminates the need to create workstations [Zachary et al.99].

4.5. Effects on Training during Training Process

The training process for each trainee, from the state of being untrained to the state of being trained, involves factors that can affect training. Figure 20 illustrates the factors that can affect the training. They simply are as follows [Getley78]:

- The training method,
- The training environment,
- Individual ability to learn and
- Individual motivation to learn.

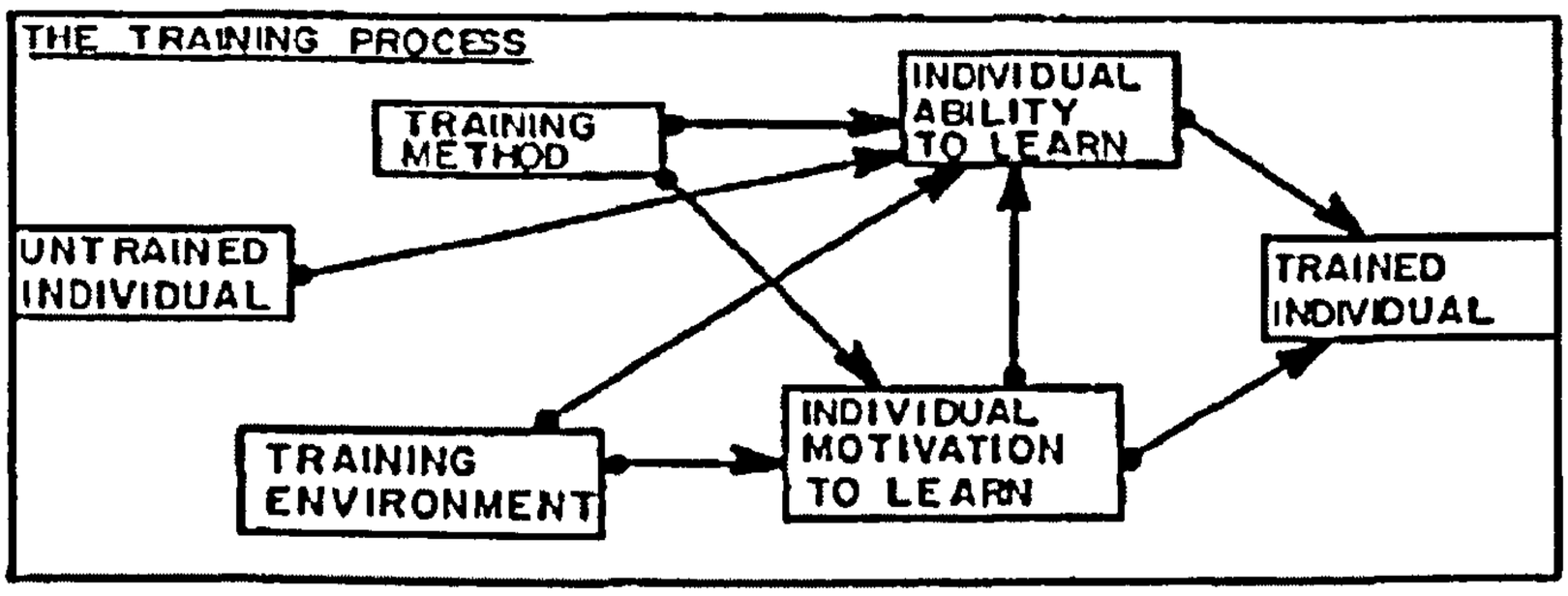


Figure 20: Factors that can affect training in the training process [Getley78]

4.6. Training Knowledge Retention

In this section, the people's training knowledge retention is discussed. Its definition and the interval factors that can affect, positively and negatively, training knowledge retention for individuals are presented.

4.6.1. Definition

Skill retention can be described as the transfer of skills from training to test [StothardNicholson01]. Retention is also defined as the ability of people to recall or retrieve the concepts and knowledge taught when they are evaluated under the same or similar situations after a period of time from the first time of knowledge acquisition [Merrienboer et al.97]. Any trained skill can be applied in situations that differ from the training environment, so trainee can be evaluated in using the skill beyond the original training environment [StothardNicholson01].

4.6.2. Retention Interval Factors

There are factors that can affect people's knowledge retention. They are called '*retention interval factors*'. They all are important factors in knowledge retention rates. They are as follows [StothardNicholson01]:

1. Time interval

   One of the factors is the time interval between training and practice. Therefore, the longer the time between training and practice, the greater skill loss that people can have.

2. Opportunity to practice

   The chance to practice the skill or task would, clearly, reduce the rate of skill loss over time. Therefore, training systems that use the methodology training by practice (by doing) have the ability to reduce skill loss over time.

3. Individual factors

   There are personal factors that also may affect the knowledge retention. One of theses factors is the motivation that individuals have to train in the first place. Another possible factor is personal ability to retrieve information. There differences in people's ability to use their skills beyond the training environment.

4.7. Summary

This chapter has shown issues in training. Training definition and methods were presented. Then, the chapter discussed embedded training and its definition, advantage and examples. Additionally, people's retention of the knowledge obtained from training and the factors was discussed. It was also shown that the knowledge retention rate can be affected by three factors; time gap between training and knowledge retention, practicing the knowledge obtained from training and personal differences between people such as ability to remember and motivation to be trained.

5. Experimental Design and Statistical Analysis

5.1. Introduction

This chapter presents an overview of two issues; the experimental designs and statistical analysis. It discusses the experimental design definition and terminologies. It also shows translating the research question to a hypothesis, the steps to performing an experiment, and the steps to test a hypothesis. The chapter shows an overview of common statistical analysis methods.

5.2. Experimental Design

5.2.1. *Definition of Experimental Design*

In order to define the term 'experimental design', the meaning of 'experiment' needs to be understood. An **experiment** is '*a test or series of tests in which purposeful changes are made to the input variables of a process or system so that we may observe and identify the reasons for changes that may be observed in the output response*' [Santner et al.03, p. 1]. Whereas, the **experimental design** is defined as '*a complete plan for applying differing experimental conditions to your experimental subjects so that you can determine how the conditions affect the behavior or result of some activity*' [Pfleeger95].

5.2.2. *The Experiment Terminology*

There are formal terms that describe the experiment components. Pfleeger and Mason et al. [Pfleeger95, Mason et al.03] define the important ones. These are as follows:

- **Treatment** is the new method or tool the experimenter wishes to evaluate (compared with an existing or different method or tool).
- **Trial** is an individual test in an experiment. Only one treatment is used in any run of an individual test.
- The **experiment** is formally described as the set of trials.
- The **experimental objects** or **experimental units** are defined as the objects to which the treatment is applied.
- **Population** involves all possible items that have one or more common characteristics under specific experimental conditions [Mason et al.03, p. 10].
- A **Sample** is a set of data taken from a population [Mason et al.03, p. 13].
- **Experimental subjects** are those people who are applying the treatment.
- A **control object** is described as an object not using the treatment when the experimenter is comparing using the treatment to not using it. The control provides information that enables to make comparisons.
- The **response variables** (also known as **dependent variables)** are those variables that are the results or outcome of an experiment [Mason et al.03, p. 12].
- Whereas, **state variables** (also known **independent variables)** are those variables that may influence the application of a treatment and then influence the result of the experiment indirectly. For example, state variables describe characteristics of the developers or the processes used to produce a piece of software code.
- A **factor** is known as an independent variable in the experimental design. The dependent variable may change as one or more of the independent variables changes [Mason et al.03, p. 12].
- An **experimental error** is defined as the failure of two identically treated experimental objects to yield identical results. The error can be as a result of problems such as errors of experimentation, errors of observation, errors of measurement or the variation in experimental resources.

5.2.3. *Steps to Performing Experiments*

Having meaningful and useful results from an experiment requires a careful planning [Pfleeger95]. An overview of the planning needed and the steps to conducting an experiment will be considered.

Pfleeger [Pfleeger95] provides an outline of the recommended procedures that lead to a good design of an experiment. They are as follows:

1. Conception

The first step is to define the goals of the experiment. The goals should be considered as research questions that need to be answered. Then, the next procedure is to plan an experiment that will provide the answers.

2. Design

This step includes selection of the response variable, choice of factors and their levels. Additionally, experimenters try to design the experiment so that the effects of irrelevant variables are distributed equally across all the experimental conditions. Realistically, this strategy is better than allowing the irrelevant variables to affect the results of a particular condition. Therefore, there are principles that help to reduce experimental error by giving guidance on forming experimental units. They are replication, randomization and blocking [Santner et al.03]. **Replication** is described as examining the response variables multiple times at the same set of inputs. It allows the experimenter to directly estimate the magnitude and distribution of experimental error. However, **blocking** means running the experiment in relatively homogeneous sets called blocks. The blocking allows observing the relation between the response variables and the inputs within blocks. Because of the homogeneity within a block, experimental error is less within a block than between blocks and then the effects of the inputs is more easily observed. **Randomization** is the process of the random allocation of subjects to groups or of treatments to experimental units. This helps the experimenter to explore how the response variables vary as the inputs vary.

Pfleeger [Pfleeger95] classifies the blocking principle under a wider one called 'local control'. Local control indicates how much control the experimenter has over the placement and the organization of subjects in experimental units. Local control has two characteristics

of the design: **blocking** and **balancing** the units. **Balancing** is defined as making sure that an equal number of subjects is assigned to each treatment wherever possible. Balancing is not necessary. However, it simplifies the statistical analysis.

3. Preparation

This step includes readying the subjects for the experiment. The experiment's instructions must be clear and written properly. Also, it is recommended and useful procedure that a run of the experiment on a small set of people (pilot) is performed. This is to ensure that the design is complete and the instructions are clear.

4. Execution

After preparing for the experiment, it can be carried out. The steps provided in the plan should be followed and the treatment to the experimental subjects should be consistently applied so that comparison of results is sensible.

5. Analysis

This phase involves analysis of the sets of data based on statistical principles. The statistical analysis gives an answer to the original research question addressed in the beginning.

6. Dissemination and decision-making

At the end of the analysis step, conclusions about how the different inputs affected the outcome will be reached. All the aspects involved in the experiment should be documented. This means that the goals, the hypothesis, the experimental subjects and objects, the treatments, the response and state variables, and the results should be carefully documented. Also, documenting both methods and conclusions in a way that will allow the research field people to duplicate your experiment and then confirm your conclusions in a similar setting.

Montgomery [Montgomery05] also stated some steps for obtaining a good performance of experiments. They seem to be similar to Pfleeger's but with different divisions. They are as follows:

1. Recognition of and statement of the problem (Conception).
2. Selection of the response variable (Design).

3. Choice of factors, levels and ranges (Design).
4. Choice of experimental design (Design).
5. Performing the experiment (Execution).
6. Statistical analysis of the data (Analysis).
7. Conclusions and recommendations (Dissemination and decision-making).

The first three steps are pre-experimental ones. The second and the third steps are often done simultaneously or in reverse order.

5.2.4. Translating the Research Goal to a Hypothesis

When a research question is clearly stated, it must be translated into a formal hypothesis. There are two kinds of hypotheses. They are the **null hypothesis** and the **experimental (or known as alternative) hypothesis** [Mason et al.03, p. 52].

The **null hypothesis** is the one that assumes that there is no difference between two treatments with regards to the dependent variable the experimenter is measuring. In contrast, the **experimental hypothesis** believes that there is a significant difference between the two treatments. The null hypothesis is assumed to be true unless the data indicates otherwise. Thus, 'testing the hypothesis' means examining whether the data is convincing enough to reject the null hypothesis and accept the experimental as true [Pfleeger95]. Therefore, in order to answer the research question, the hypothesis needs to be tested.

5.2.5. Hypothesis Testing

Mason et al. summarizes the steps to test a hypothesis into four basic steps [Mason et al.03, p. 77]. They are as follows:

1. State the null and alternative hypotheses.
2. Collect a sample and work out the appropriate test statistic.

3. Compare the significance probability of the test statistic to the significance level selected for the test.
4. Draw the appropriate conclusion and interpret the results.

5.3. Analysis

Having collected data from the experiments, the analysis of the data will be presented in this section. An overview of different statistical analysis methods will be presented. This section discusses also a way to choose the appropriate analysis methods for different experimental designs.

5.3.1. Choosing Statistical Analysis Methods

In this section, choosing the suitable statistical analysis method for the evaluation experiments is presented. There will be some points need to be taken in consideration in order to decide what the appropriate method that suits the experiment data.

Pfleeger [Pfleeger95] has given three major points to consider when choosing the analysis methods. They are the nature of the collect data collected (distribution of data), the type of experimental design used (design considerations) and the aim of carrying out the experiments. Each one of them is considered in turn.

- **Distribution of Data**

It is essential to understand that the data are a sample from a larger population. After that, the relatively small sample might be generalized to larger population. Many statistical methods assume that the data is normally distributed, and the sample is randomly taken from larger distribution [Pfleeger95].

▪ Design Considerations

Pfleeger [Pfleeger95] states that '*the experimental design must be considered in choosing the analysis techniques. At the same time, the complexity of analysis can influence the design chosen*'. For example, multiple groups usually need to use the analysis of variance (ANOVA) method, whereas a simple t-test can be used with two groups.

▪ Purpose of the Experiment

The goal of the experiment plays an important role in choosing the suitable statistical analysis method. There are four major objectives to conduct a formal experiment [Pfleeger95]. They are as follows:

A. Confirming a theory

The experiments often exist to evaluate a theory. For example, an experiment hypothesis believes that the use of a certain technique (the treatment) has an effect on the experimental subjects, making it better than another treatment (usually the existing technique).

If the data is taken from a normal distribution and there are two groups to be compared to each other, the t-test can be used to analyze the effects of the two treatments. If there are more than two groups to compare, the analysis of variance (ANOVA) test, using the *F statistic,* is appropriate. In contrast, if the data is taken from a non-normal distribution, the Mann-Whitney and Wilcoxon Signed Ranks tests are used for comparing two groups and Kruskal-Wallis is suitable to be used for comparing more than two groups [Field05, p. 521].

B. Exploring a relationship

Some experiments are conducted to determine the relationship among data that describes one variable or across many variables. For example, knowing the normal ranges of productivity or quality on many projects, so there is a baseline to compare for the future.

In the case of having an experiment to explore a relationship, there are three techniques can be used: box plots, scatter diagrams, and correlation analysis. A *box plot* can describe a summary of the range of a set of data about one variable. It shows where most of the data is gathered. Regarding the *scatter diagram*, it describes the relationship between two variables. The analyzer can visually determine the likelihood of an underlying relationship between the variables. Finally, the *correlation analysis* uses statistical methods to validate whether there is a real relationship between two attributes.

C. Evaluating the accuracy of a model

In many software engineering projects, a model of behavior is used to predict what should occur. Although the purpose of the experiment is different from confirming a theory, the analysis methods are the same. Consequently, the methods mentioned in confirming a theory purpose can be used as well. This is because the prediction model generates a predicted date set which then can be compared with real data.

D. Validating a measure

Verifying the measure that captures the attribute it claims to reflect can be a purpose of an experiment. Exploring the relationship purpose can be used in the experiments are often designed to validate a measure. This happens when exploring the relationship between the measure and data that is recognized to be correlated with the attribute. Due to this reason, the analysis methods in 'exploring a relationship' purpose are suitable ones for 'validating a measure' purpose.

5.3.2. An Overview of Common Statistical Analysis Methods

Having shown a way of choosing the suitable method, this section presents the common statistical analysis methods. There are different tests based upon the ways of data collection in each test and the number of samples. They are [Urdan05, pp. 89-90 & p. 309]:

5.3.2.1. Tests for Two Independent Samples

When there are two experimental conditions and two different groups of subjects, each group is assigned to one condition (See Figure 21).

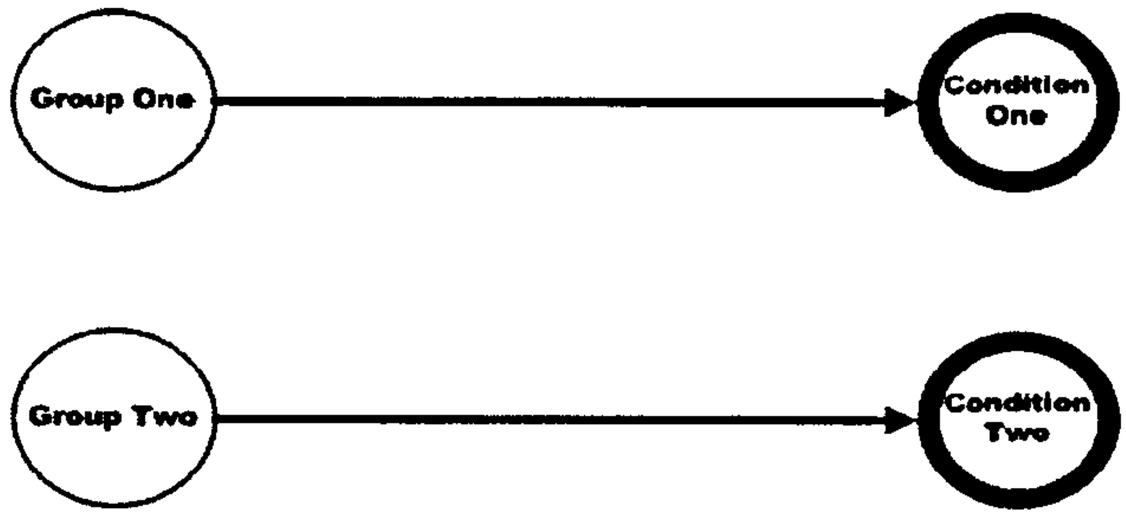


Figure 21: Two independent samples

The tests for two independent samples are different based on data type; parametric and non-parametric.

- **Parametric Tests**

All parametric tests are based on normal distribution. They are reliable under assumptions. These assumptions are [Field05, p. 64]:

- Data is extracted from normally distributed population,
- The dependent variable is measured on an interval scale at least.

If the test is used to test different groups of people, two assumptions are added [Field05, p. 287]. They are as follows:

- In each experimental group, the variances are roughly equal,
- Data from different participants are independent. This means the behavior of one participant does not influence the behavior of another participant.

When a comparison between two means collected from two different groups of subjects taken from a normal population is required, a statistical method called 'Independent t-test' is used [Urdan05, p. 299].

- **Non-Parametric Tests**

There is another type of tests called non-parametric tests [Sprent00, p. 3]. They are also referred to as distribution-free tests because they do not make assumption about population distribution [KinnearGray04, p. 9]. Therefore, the non-parametric tests are used when the data distribution is assumed non-normal.

When there are two means from two different groups of subjects taken from non-normal population, a statistical method called 'Mann-Whitney test' is used [KinnearGray04, p. 9]. Mann-Whitney test is the nonparametric equivalent for 'independent t-test' [ibid].

5.3.2.2. *Tests for Two Dependent (Related) Samples*

When there are two experimental conditions and the same subjects take part in both conditions (See Figure 22), this is called dependant samples, matched-pairs or paired samples test.

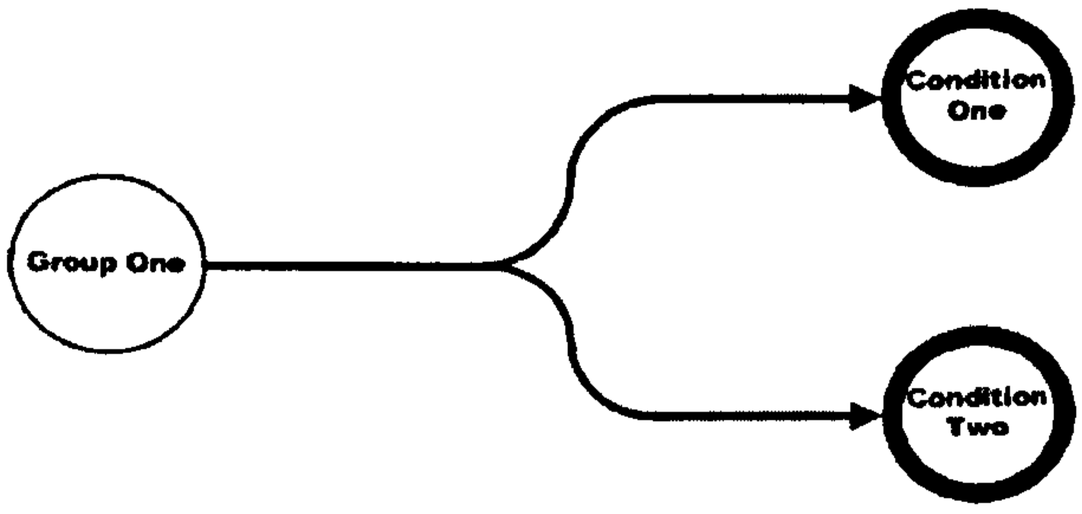


Figure 22: Two dependent samples

As described with the two independent samples, the tests for two dependent samples are different based on data type; parametric and non-parametric.

- **Parametric Tests**

When there are two samples taken from the same subjects, a test called 'Paired samples t-test' is used [Field05, p. 286].

- **Non-Parametric Tests**

When there are two means from the same subjects taken from non-normal population, a statistical method called 'Wilcoxon Signed Ranks test' is used [Field05, p. 534]. The test is the nonparametric equivalent for 'Paired samples t-test' [ibid].

5.3.2.3. Tests for Several Independent Samples

In case of having more than two means or more than two groups of participants (e.g. three groups), different statistical tests are used based on data type; parametric and non-parametric.

- **Parametric Tests**

A statistical analysis test called 'Analysis of Variance' (ANOVA) is used when there are more than two means taken from different groups of subjects taken from a normal population [Urdan05, p. 101]. There are different ANOVA designs. One design is the One-way independent ANOVA which compares several means taken from different participants when there is one independent variable [ibid]. Another design is Factorial ANOVA which is used when there are two or more independent variables (the variables are known as factors) [Field05, pp. 389-390]. Therefore, there could be Two-way ANOVA which indicates that there are two independent variables or Three-way ANOVA when there are three independent variables and so forth [ibid].

- **Non-Parametric Tests**

The nonparametric equivalent test for 'One-way independent ANOVA' is called 'Kruskal-Wallis test' [KinnearGray04, p. 219]. However, the Kruskal-Wallis test tells only of a difference exists between the groups [Field05, pp. 549-550]. In order to see the difference between each group and another group, follow up tests (post hoc tests) are carried out using Mann-Whitney test between every two groups. Using many Mann-Whitney tests might provide inaccurate results. This can be resolved by using *Bonferroni correction* [ibid].

Bonferroni correction means that instead of using .05 as a critical value of significant difference for each test, the value (.05) is divided by the number of tests carried out as post hoc tests [ibid]. For example, if there are three groups need to be compared, there should be three Mann-Whitney tests to compare the groups with each other. Therefore, instead of using (.05) as the critical value of significance, (.05/3=.0167) is used. It is recommended not to use this follow up tests method in case there are many groups because the critical value will be too small [ibid].

In all nonparametric tests mentioned (Mann-Whitney, Wilcoxon Signed Ranks and Kruskal-Wallis tests), there were different significance methods that should be chosen depending on the sample size. If the sample is large, methods called *Asymptotic* or *Monte Carlo* could be used. However, if the sample is small, *Exac* test should be chosen in order to have accurate results [Field05, pp. 528,538,547]

5.4. Summary

This chapter presents an overview of two issues; the experimental designs and statistical analysis. It discusses the experimental design definition and experiment's terminologies. It also shows translating the research question to a hypothesis and then the steps to performing an experiment. Additionally, in this chapter, the way to choose one of various possible statistical methods to use is presented. Finally, an overview of parametric and non-parametric statistical methods used is given.

6. An Evaluation of Users' Tips Effectiveness for Phishing Websites Detection

6.1. Introduction

Recently, Phishing attacks have become a serious problem for end-users, online banking and e-commerce websites. Many anti-Phishing approaches have been proposed to detect and prevent Phishing. The most basic approach is publishing guidelines for the Internet users to follow when they go online. Theses guidelines are referred as users' tips in this thesis. The anti-Phishing tips are published by many governmental and private organizations. All the information used in the training approaches is based on the users' tips. There are many different tips. This chapter examines the effectiveness of most common users' tips for detecting Phishing websites. In this chapter, a novel effectiveness criteria is proposed and used to examine each single tip and rank it based on its effectiveness score. The chapter tries to find fewer anti-Phishing tips that users can focus on to detect Phishing attacks by themselves.

Chapter 3 already reviewed the literature with related to this chapter (Sections 3.3 and 3.7). The remainder of the chapter is organized as follows. The chapter describes the research methodology and then the results. Then, the final section concludes the chapter with a discussion of the findings.

6.2. Research Methodology

Users' anti-Phishing tips for Phishing websites will be examined based on evaluation criteria that consist of four criterions. These criteria are called 'Effectiveness Criteria'.

This methodology is described by three sections. They are as follows:

6.2.1. Collection of Anti-Phishing Tips Sample for Phishing Websites

The anti-Phishing tips are collected in two steps as follows:

6.2.1.1. Survey of Online Fraud Tips

A survey of online fraud tips was carried out. The survey was for both businesses' and users' tips published by government organizations, banks, financial organizations and e-commerce websites. The online fraud tips were for all types of online fraud. The online fraud tips resulting from this survey was 491 different tips. Figure 23 shows the sources of the tips.

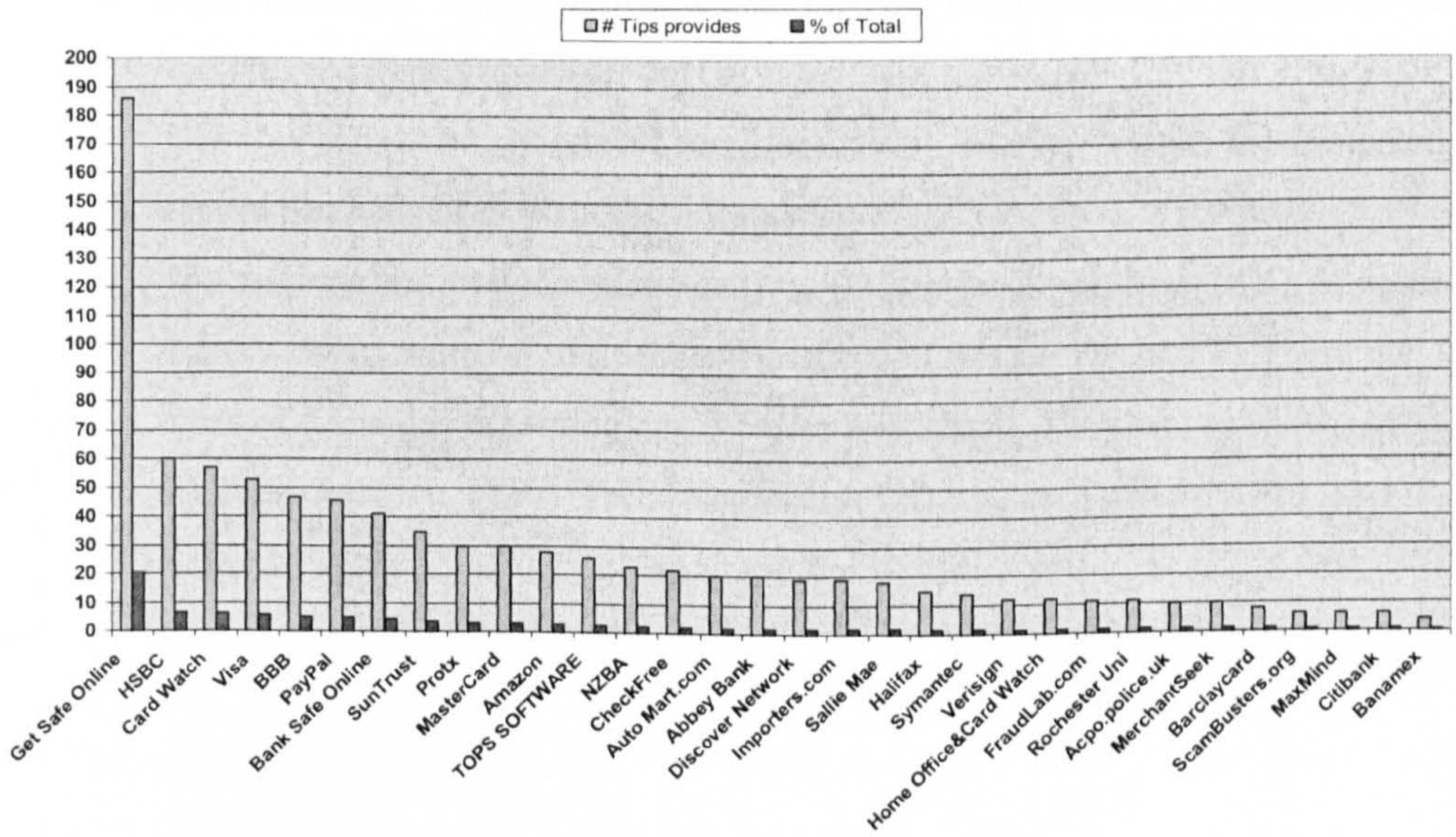


Figure 23: The sources of the online fraud tips

6.2.1.2. Extracting the Anti-Phishing Tips

The anti-Phishing tips are extracted from the online fraud tips that are related to users and for Phishing websites. This step consists of three phases:

a. Extracting from the online fraud tips' survey the tips that are related to users. The number of tips resulted from this phase was 290.
b. Extracting from the users' online fraud tips that are resulted from phase (a) the tips that are related to Phishing attacks. The Phishing emails' and websites' tips extracted from this phase were 57.
c. Extracting from anti-Phishing tips resulted from phase (b) that are applicable to Phishing websites. The anti-Phishing tips for websites was 21. Therefore, the effectiveness evaluation is on these 21 tips.

6.2.2. Effectiveness Criteria

The effectiveness criteria are as follows:

1. The tip detects the most common clue. This requires analysis of Phishing scenarios to find out the most common Phishing clues appear in the scenarios.
2. Solo reliability. This criterion means that the evaluated tip is enough to detect and prevent Phishing attack.
3. The clue cannot be spoofed [Cranor et al.06b]. In other words, the evaluated tip cannot be changed or faked by a fraudster.
4. The tip does not produce false positives (FP) or false negatives (FN). This means that by using the tips, the decision made will not be FN or FP. There are four types of decisions regarding any website legitimacy. They are, as shown in Table 6, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).

Decision / *Website Legitimacy*	**True**	**False**
Positive	***TP***	***FP***
Negative	***TN***	***FN***

Table 6: The possible decisions could be made regarding websites' legitimacy

In order to understand the four types of decisions, they are defined as follows:

- *True Positive (TP)*: The TP case happens when a legitimate website is considered as legitimate.
- *True Negative (TN)*: The TN case happens when a Phishing website is considered as Phishing.
- *False Positive (FP)*: The FP case happens when a legitimate website is considered as Phishing.
- *False Negative (FN)*: The FN case happens when a Phishing website is considered as legitimate.

These criteria are then given weights as shown in Table 7. The effectiveness weight for the criteria is divided into four equal quarters. This means that each criterion has 0.25 of the weight.

#	Criterion	Score (out of 1)
1	The tip prevents the most common clue	0.25
2	Solo reliability	0.25
3	The clue, addressed by the tip, cannot be spoofed	0.25
4	The tip does not possibly produce FP or FN	0.25

Table 7: The effectiveness criteria and their scores

After evaluating each single tip against each single criterion and finding out whether or not it satisfies the criterion, the tip effectiveness can be calculated using the following 'Effectiveness Metric' EM:

$$EM(T) = \sum_{i=1}^{4} W_i \bullet C_i$$

where *EM(T)* is the effectiveness metric of the tip *T*, W_i is the weight of criterion i and C_i is 1 if criterion i is relevant or 0 if criterion i is not relevant or not applicable.

Therefore, the tip with the most effectiveness score will be first in the effectiveness ranking and the second effective tip should be the second and so on. In the case where two or more tips have the same effectiveness score, the tip with the most percentage of clue appearance in Phishing scenarios analysis should come first and so on. This is referred as '*Ranking Role*'.

6.2.3. Applying the Effectiveness Criteria

Applying the effectiveness criteria to each individual tip requires having the common Phishing clues appear in Phishing scenarios. Therefore, an analysis of Phishing scenarios is carried out.

6.2.3.1. Phishing Scenario Analysis

An analysis of 42 real Phishing scenarios presented in the APWG's archive [APWG07c] was carried out. The scenarios analyzed were the latest scenarios that were added to the archive by APWG experts. The scenarios were described and explained in details in the archive. Figure 24 illustrates an example of a Phishing scenario.

Figure 24: An example of Phishing scenario described in APWG archive

The purpose of analyzing the scenarios is to find the most common Phishing clues that appear in the scenarios. In other words, what are the Phishing 'indicators' that appear most in Phishing scenarios?

In the analysis, each clue that appears in a Phishing scenario is counted. One clue could appear in many Phishing scenarios and one Phishing scenarios could have more than one clue. This is illustrated in Figure 25.

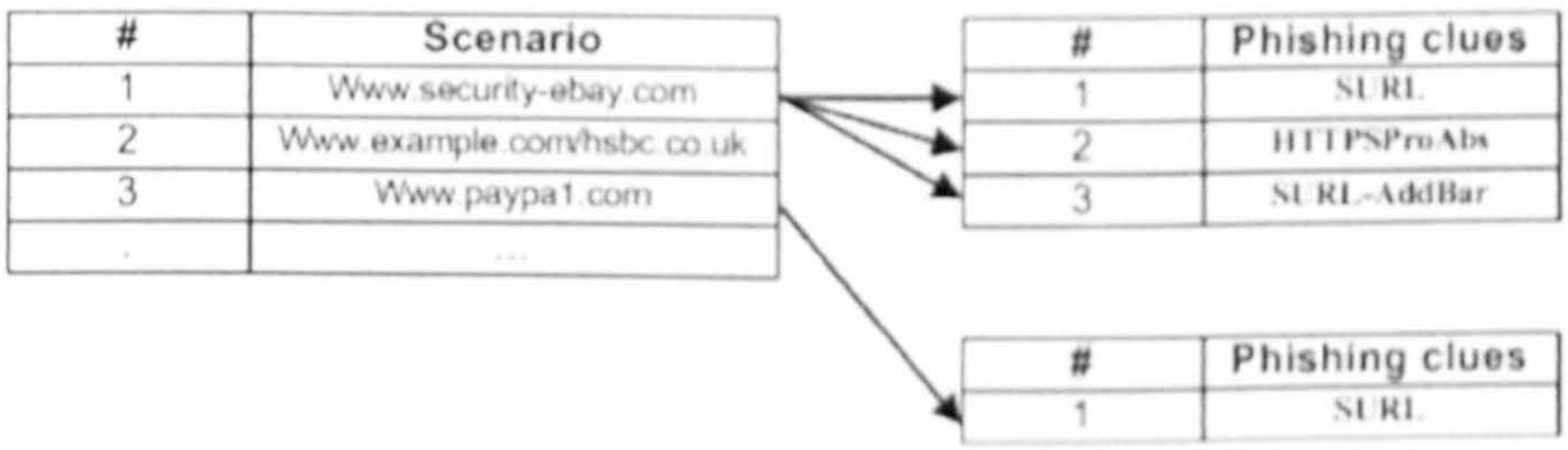


Figure 25: Example of extracting Phishing clues from scenarios

Table 8 presents the clues that appear most in the Phishing scenarios. Table 9 shows the meanings of the abbreviations used for the clues.

Clue	No. of Appearances	% of Appearance
SURL	42	100.00
LockStBarAbs	38	90.48
HTTPSProAbs	31	73.81
SURL-AddBar	30	71.43
AddBarNotVisible	3	7.14
HTMLError	1	2.38
StatBarDisabled	1	2.38

Table 8: Clues that appear in the Phishing scenarios

Abbreviation	Description
SURL	Suspicious URL for the web page
LockStBarAbs	Absence of a 'lock' icon in the status bar
HTTPSProAbs	Absence of the https protocol in the address bar
SURL-AddBar	Suspicious URL in the address bar
AddBarNotVisible	Address bar is not visible.
HTMLError	HTML errors in rendering the page
StatBarDisabled	Status bar is disabled.

Table 9: Clue abbreviations

The analysis of the Phishing clues identified that there were 7 major clues. Table 9 presents these clues. Only 2 of these clues (*SURL* and *HTTPSProAbs*) were the same as these identified by Chou et al and discussed in Section 3.3 in Chapter 3. Thus, few new clues were identified and replaced those of Chou et al.

As Table 8 shows, the clue *SURL* 'suspicious URL for the web page' appeared in all the 42 Phishing scenarios. Furthermore, the clue *SURL-AddBar* 'suspicious URL in the address bar' appeared in approximately 90% of the scenarios. In contrast, clues such as *HTMLError* 'HTML errors in rendering the web page' and *StatBarDisabled* 'status bar is disabled' have the least appearances. Each clue appeared once in the scenarios.

6.3. Results

#	Tip	Criteria				TE	TR
		1	*2*	*3*	*4*		
		0.25	**0.25**	**0.25**	**0.25**		
1	Type in your browser the address of the website you intend to go or use a bookmark that you previously created.	N	Y	NA	Y	**0.5**	2
2	Make sure you are on a secure connection when entering sensitive information. Secure Web pages will have the text https: instead of http:	Y (73.8)	N	N	N	**0.25**	5
3	Do not be fooled by a padlock that appears on the web page itself. It's easy for conmen to copy the image of a padlock. Look for one that is in the window frame of the browser.	N	N	N	N	**0**	>6
4	Look beyond the logo and do not give out your information before you check the privacy and security seals. Scammers often include actual logos and images of legitimate companies.	N	Y	Y	N	**0.5**	~2
5	A fake website's address is different from what you are used to, perhaps there are extra characters or words in it or it uses a completely different name or no name at all, just numbers. Check the address in your browser's address bar after you arrive at a website.	Y (71.4)	N	N	N	**0.25**	6
6	Even though you are asked to enter private information there is NO padlock in the browser window or 'https://' at the beginning of the web address to signify that it is using a secure link and that the website is what it says it is.	Y (90.4)	N	N	N	**0.25**	4
7	A fake website may have this characteristic: The website's address is different from what you are used to, perhaps there are extra characters or words in it or it uses a completely different name or no name at all, just numbers. Check the True URL. The true URL of the website can be seen in the page 'Properties'.	Y (100)	Y	Y	N	**0.75**	1

Table 10: Results of tips effectiveness

As a result of the applying the effectiveness criteria, there are different tips effectiveness TE scores and different tips ranking TR accordingly. Table 10 presents the results of the first seven ranked tips (all 21 tips are presented in Appendix A).

In order to clarify how Table 10 was constructed, an example of one item from the table is shown. The example explains the construction of item 2. The tip representing item 2 in Table 10 is '*make sure you are on a secure connection when entering sensitive information. Secure Web pages will have the text https: instead of http:*'. This tip is examined against every criterion from the 'Effectiveness Criteria' shown in Table 7. As Table 10 presents, each criterion has a weight of 0.25. As explained earlier in Section 6.2.2, if a tip satisfies a criterion then the weight is multiplied with 1 whereas the weight is multiplied with 0 if the tip does not satisfy the criterion or is not applicable. Thus, when tip 2 examined against each criterion, the tip effectiveness TE score was constructed using the effectiveness metric (See Section 6.2.2) as follows:

$$EM(2) = 0.25 \times 1 + 0.25 \times 0 + 0.25 \times 0 + 0.25 \times 0$$
$$EM(2) = 0.25$$

After this calculation, the tip with the most effectiveness score became first in the effectiveness ranking and the second effective tip became the second and so on. Thus, tip 2 was ranked as fifth according to the tip effectiveness score TE.

Regarding the results, there is no tip that satisfies all the criterions defined. The most effective tip is tip number 7. It has met three out of four criterions. Its effectiveness score is 0.75. Tips 1 and 4 come second in the ranking because they have the same score (0.50).

The tips 2, 5 and 6 have the same score (0.25). However, they have different ranking. Their ranking is fifth, sixth and fourth respectively. This is because the '*Ranking Role*' is used. The three tips have different clue appearance's percentages. As shown between brackets in criterion 1 in Table 10, the clue of tip 6 appeared in 90.4% of Phishing scenarios whereas, tip 5 appeared in 73.8% and tip 6 appeared in 71.4%.

It is worth mentioning that the ranking of tips 1 and 4 has not been calculated in the same way as the ranking of the tips 2, 5 and 6. This is because there are no clue appearance's percentages for tips 1 and 4. Therefore, the Ranking Role can not be fully applied.

The tip 3 has the last ranking because its effectiveness score is zero. This is because it does not meet any of the criterions. Its ranking is (>6). It is not given rank seven because all the rest of tips have the same ranking.

6.4. Discussion

There is no completely effective tip (with an effective score of 1). The most effective tip met three out of four criterions. Its effectiveness score is (0.75). It has not met the criterion four. This is because the tip helps finding the true URL of a page but it does not help in verifying whether or not the URL is related to a legitimate website. Thus, it possibly produces FP or FN by using it alone. Using a search engine, such as Google, in verifying the URL after using the tip would overcome its weakness.

Therefore, the most effective anti-Phishing tip is used with a search engine recommendation as follows: "*a fake website's address is different from what you are used to, perhaps there are extra characters or words in it or it uses a completely different name or no name at all, just numbers. Check the True URL (Web Address). The true URL of the site can be seen in the page 'Properties' or 'Page Info': While you are on the website and using the mouse Go Right Click then Go 'Properties' or 'Page Info'. If you don't know the real web address for the legitimate organization, you can find it by using a search engine such as Google*".

6.5. Summary

In this chapter, an evaluation of the effectiveness of most common users' tips for detecting and preventing Phishing websites was carried out. A novel effectiveness criteria was proposed and used to examine each single tip and rank it based on its effectiveness score. The 'Effectiveness Criteria' involves four criterions.

The chapter found the most effective anti-Phishing tips that users can focus on to detect Phishing attacks. The most effective tip met three quarters of the criterions. It has not met the criterion four because the tip helps finding the true URL of a page but it does not help in verifying whether or not the URL is related to a legitimate website. The tip would overcome its weakness by using a search engine, such as Google, after its use to verify the URL. Also, the effective tips can be focused by anti-Phishing training approaches.

7. An Anti-Phishing Approach That Uses Training Intervention for Phishing Websites Detection

7.1. Introduction

This chapter proposes a novel Anti-Phishing Approach that uses Training Intervention for Phishing Websites Detection 'APTIPWD'. The APTIPWD approach considers helping people detecting Phishing websites during their normal use of the Internet. It brings information to end-users and helps them immediately after they have made a mistake in order to recognize Phishing websites for themselves.

The chapter is organized as follows. The New Approach is presented in the second section. Then, the scenarios of the proposed approach are discussed in the third section. After that, this chapter presents a prototype proof of concept implementation of the New Approach. The aim is to validate whether the New Approach is implementable, viable and can be deployed properly. The final section concludes the chapter with a discussion on the New Approach.

7.2. The Proposed Approach

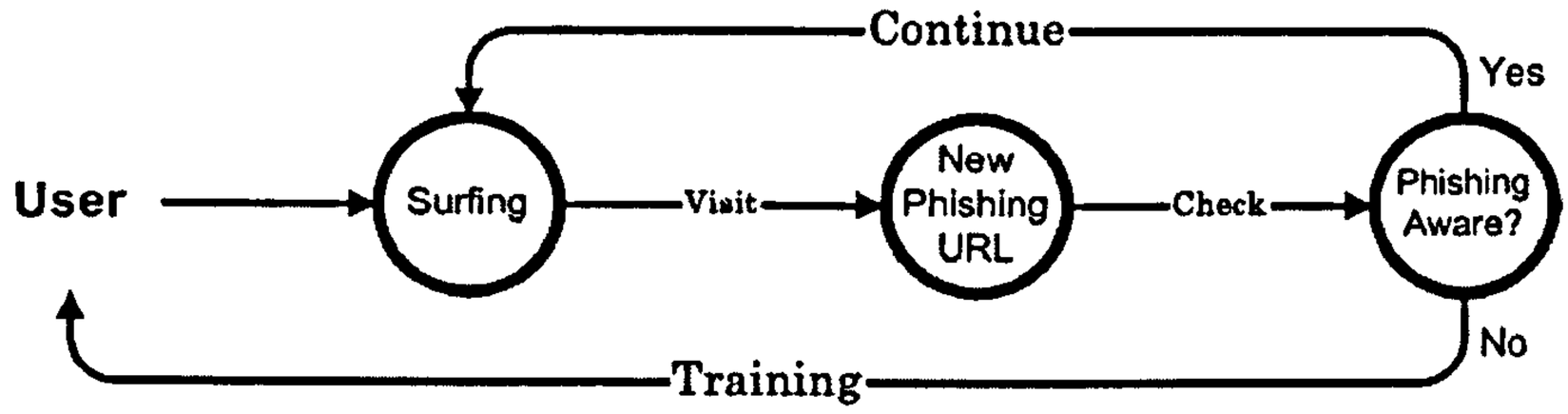


Figure 26: The broad idea of the anti-Phishing proposed approach

The process of the New Approach is shown in Figure 26. The broad idea is to check whether a user is Phishing aware when they surf the Internet and visit a Phishing website. If the user tries to submit their sensitive information to the Phishing website, they are shown intervening message to help them understand what Phishing websites are and how to detect them. The New Approach also keeps anti-Phishing training ongoing process. This means that whenever users try to submit information to Phishing website, they will be trained. In the case where the user is Phishing aware, the approach does nothing and lets the user keep surfing the Internet.

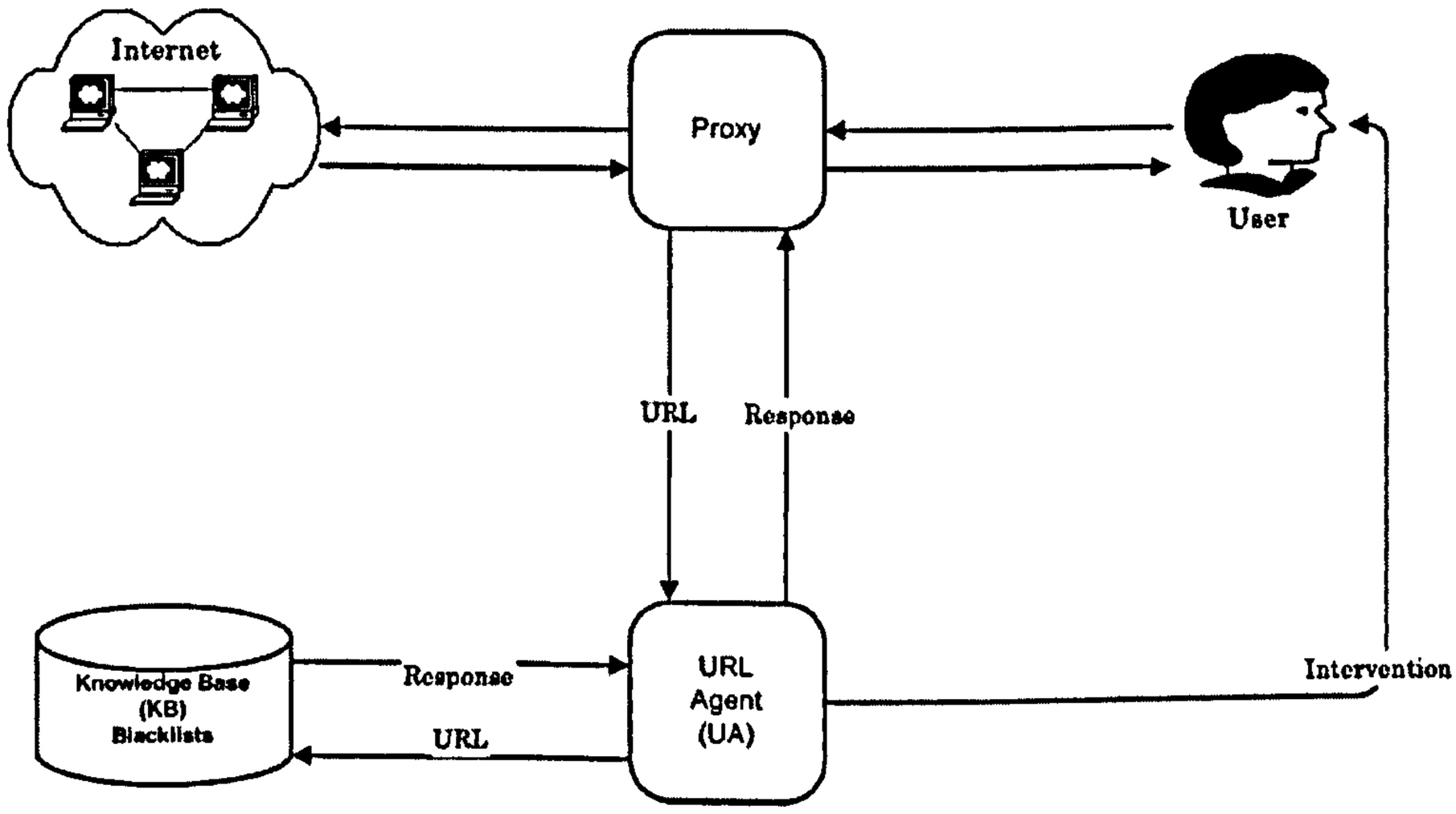


Figure 27: The architecture of the New Approach (APTIPWD)

The New Approach is based on training intervention based around the use of blacklists to detect Phishing websites. Figure 27 shows the main components of the approach. The components are Proxy, URL Agent (UA) and Knowledge Base (KB). The intervention takes place between the Internet and Users. Any URL request made by a user goes through the Proxy. The Proxy communicates with a URL Agent (UA). When the user browses the URL page and clicks to submit information, the UA verifies whether the URL is blacklisted or not by checking the blacklists. If the URL is not blacklisted, the Proxy allows submission process to proceed. If the URL is blacklisted, the Proxy prevents the information being submitted. Then, the UA shows an intervening message to the user in order to help them understanding what Phishing is and how to detect them in the future.

There are many anti-Phishing tips that can be used in the intervening message. The most effective anti-Phishing tip evaluated in Chapter 6 is used. The tip used in the intervening message is as follows: "*a fake website's address is different from what you are used to, perhaps there are extra characters or words in it or it uses a completely different name or no name at all, just numbers. Check the True URL (Web Address). The true URL of the website can be seen in the page 'Properties' or 'Page Info': While you are on the website and using the mouse Go Right Click then Go 'Properties' or 'Page Info'. If you don't know the real web address for the legitimate organization, you can find it by using a search engine such as Google*".

Using the New Approach will present the intervening messages to users who access Phishing websites and try to submit their information. Also, by using this approach, users do not need to attend training courses and do not need to access online training materials. This is because the approach brings information to end-users and helps them immediately after they have made a mistake in order to detect Phishing websites by themselves. The New Approach helps users on how to make correct decisions in distinguishing Phishing and legitimate websites during their normal use of the Internet.

This approach will only work if intervention is shown to be an effective method for training people in detection of Phishing websites. In order to effectively evaluate the New Approach, a series of experiments need to be carried out.

7.3. Simulating the Proposed Approach

Evaluating the New Approach on the real Internet is difficult because the blacklists component is dynamic and therefore is hard to control. A better solution is to evaluate under experimental conditions. If the evaluation reveals that the approach is successful and achieve its goals, the approach will be implemented and evaluated on the Internet with dynamic blacklists.

In order to evaluate the approach accurately under experimental conditions, all possible scenarios of the approach need to be simulated and the blacklists (dynamic components) need to be made fixed. The scenarios are shown in the flow chart diagram illustrated in Figure 28. The possible scenarios are as follows:

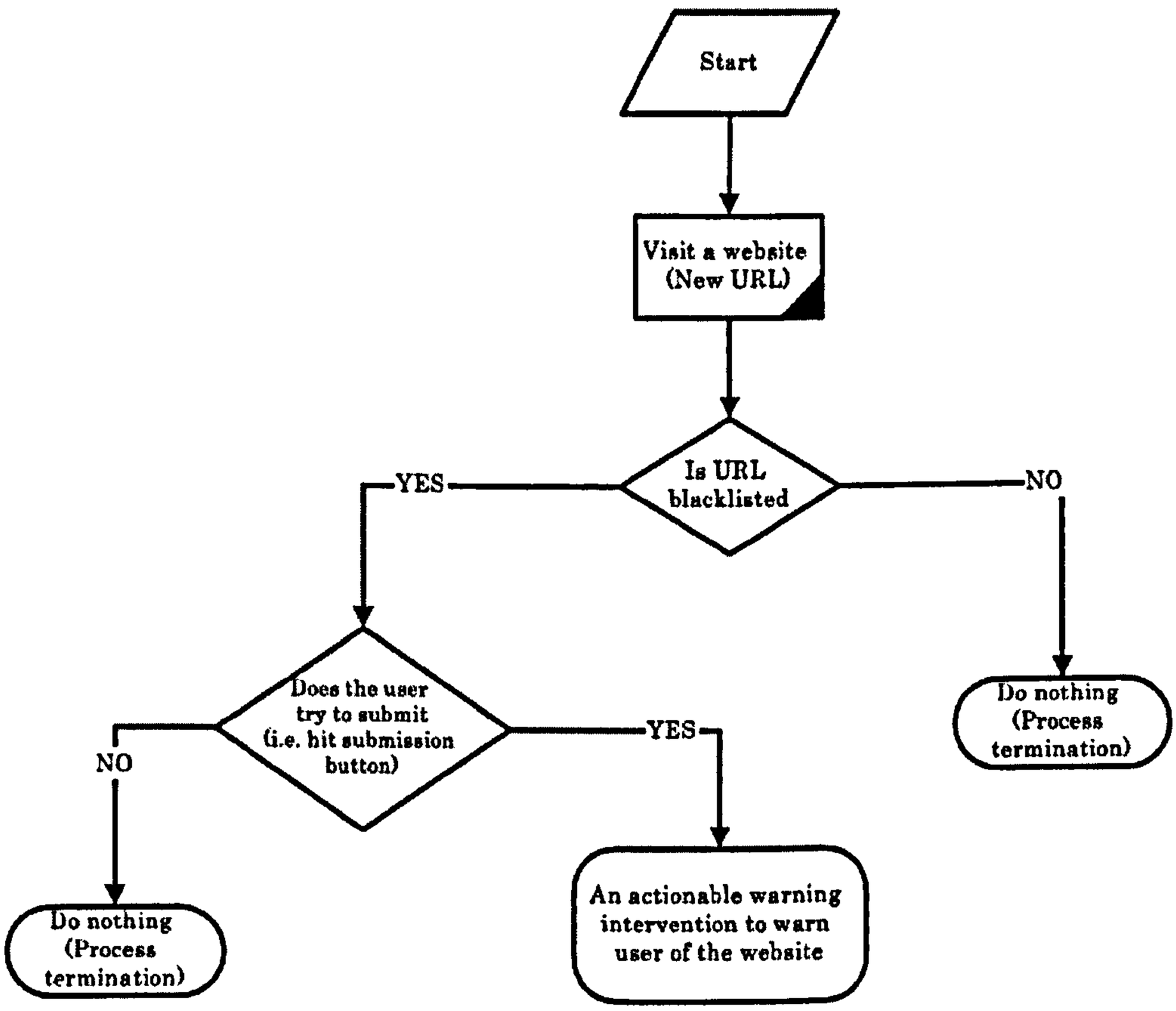


Figure 28: Flow chart diagram for the New Approach's scenarios

I. **Scenario one:**

The user visits a new website (i.e. new URL) and it is checked whether the URL is blacklisted. If the URL is blacklisted and the user does not hit the submission button to submit their information, no action is taken.

II. **Scenario two:**

The user visits a new website (i.e. new URL) and it is checked whether the URL is blacklisted. If the URL is blacklisted and the user hits the submission button to submit their information then an intervening anti-Phishing message is shown.

III. **Scenario three:**

The user visits a new website (i.e. new URL) and it is checked whether the URL is blacklisted. If the URL is not blacklisted then no action is taken.

These scenarios will be implemented and then used in the evaluation experiments described in Chapter 8 and analyzed in Chapter 9.

7.4. An Approach to the Implementation of the APTIPWD

7.4.1. Proxy based Computer Network

7.4.1.1. General Structure

A client-server model is a common design for distributed computing. The client and the server are two components that interact between each other [JiaWanlei04, p. 16].

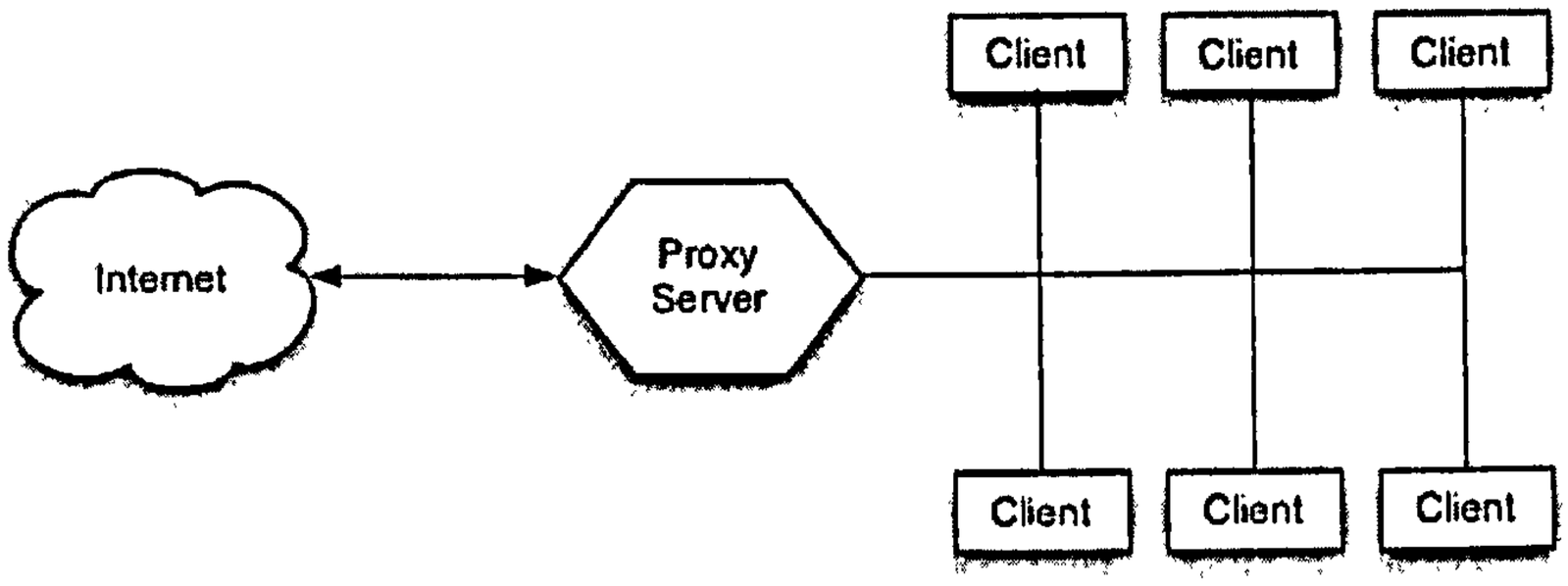


Figure 29: Server-Proxy-Client Interaction[20]

A client-proxy-server model extends the client-server model [Singh05, pp. 36-9]. It introduces an additional component which is a proxy. The proxy is located between the client and the server [ibid]. Figure 29 presents an overview of the interaction between the client, proxy and server. The server component is represented by the "Internet" because in a proxy based computer network, any URL request to the web made by a client is directed to the URL domain server. Proxies have been widely used in many applications to perform various tasks such as

- clients' connections control,
- URLs' request control,
- caching and
- filtering data [XiaoChen08, p. 331].

7.4.1.2. How it Works

The interaction between client and server is as follows [JiaWanlei04, p. 16]:

- Client requests a service from Server.
- Server processes the requests and replies to Client.

However, in the client-proxy-server, the interaction becomes as follows:

- Client sends request for Server to Proxy.
- Proxy passes request to Server.

---

[20] Source: ServerWatch.com, available at: http://www.serverwatch.com/tutorials/article.php/10825_3092521_1, last access on 15 November 2008

- Server processes the request and sends reply for the Client to Proxy.
- Proxy passes reply to Client.

7.4.2. Applying the New Approach to a Proxy based Computer Network

In this section, the New Approach is applied to a proxy based network. The blacklists (dynamic components) in the approach architecture shown in Figure 27 is made fixed list. The design and implementation are described.

7.4.2.1. System Design with Fixed List of Phishing Websites

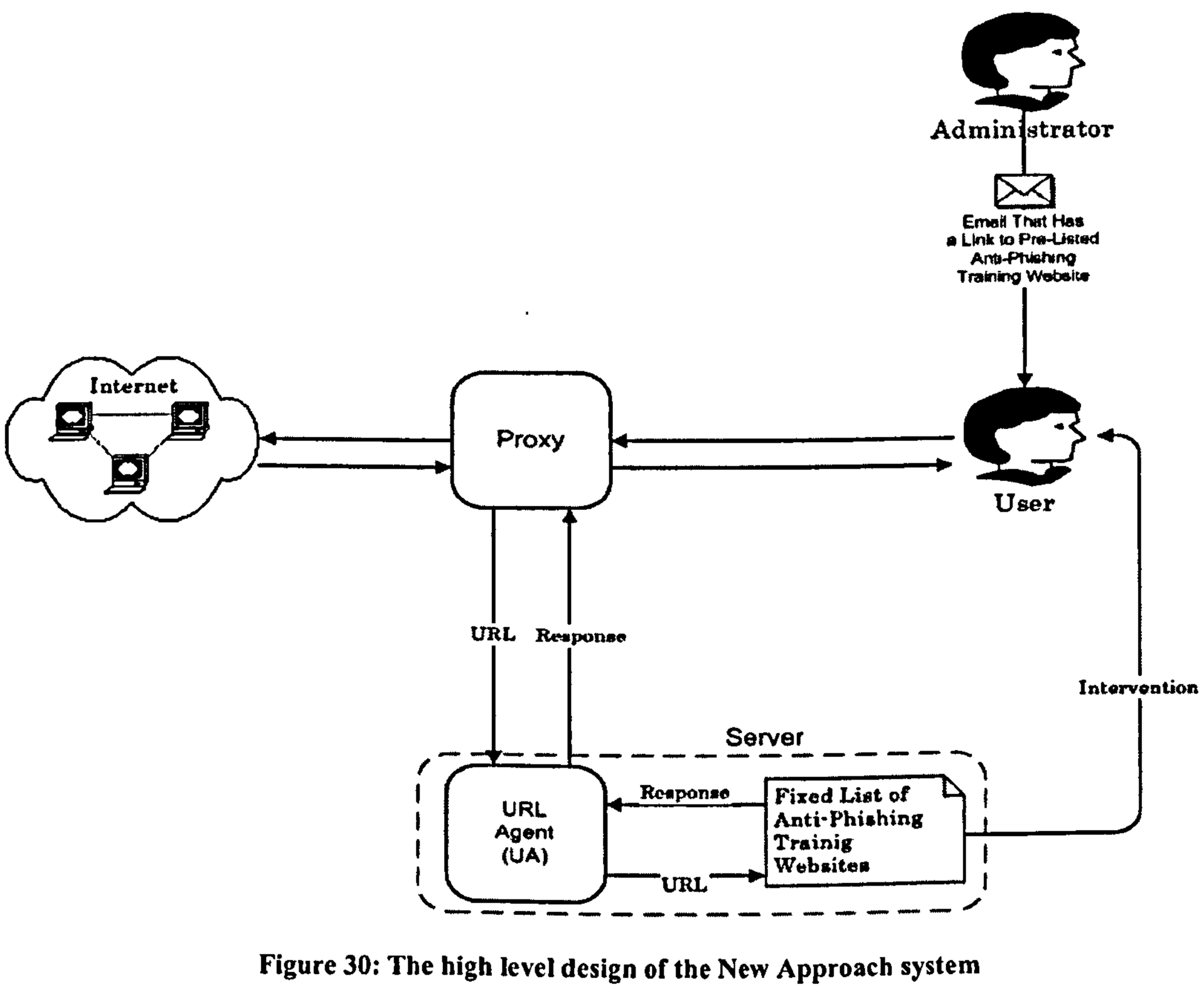


Figure 30: The high level design of the New Approach system

As shown in Figure 30, the design of the New Approach system consists of four components. They are:

- Server,
- Proxy (Gateway),
- Administrator, and
- Client (User).

The Administrator is a person who is in charge of sending Phishing emails to any User in a network. The Proxy is in place between the Internet and Users. The Proxy acts as a gateway for all requests made in the network by its Users. Any URL request made by a User goes through the Proxy. The Proxy then communicates with the Server. The Server contains three sub-components. They are a Fixed List of Anti-Phishing Training Websites (FLAPTW), a URL Agent (UA) and the Intervention message. The FLAPTW contains a fixed number of fake websites that are designed to look the same as the original ones and to be used for anti-Phishing training only, whereas the UA is responsible for checking whether the requested URL passed by the Proxy is in the FLAPTW. The Intervention message is stored in the Server. It is shown to the User in order to help them understand what Phishing is and how to detect it in the future.

The Administrator sends the anti-Phishing training email to (a) specific User(s). The email contains a link (URL) for one of the FLAPTW. If the User goes to the URL, the UA verifies whether or not the URL is listed in the FLAPTW by checking the FLAPTW. If the URL is listed, the proxy redirects the User to a simulated Phishing page (i.e. not Phishing) to browse it. The page submission button is linked with an intervention message so that if the User clicks the button to submit information the intervention message is presented to them. If the URL is not listed, the Proxy allows the User to browse the Internet as normal. This process is similar to the scenarios described in Section 7.3.

7.4.2.2. Assumption

There is an assumption that the Administrator is given the privilege in the network email system to send anti-Phishing training email that bypasses the anti-Phishing filters that might be applied in the network email system. This means that the anti-Phishing training email should have the following characteristics:

- The domain of the sender's email should be the same as the domain of a legitimate website.
- The email content should look as it is legitimate email.

7.4.3. Implementation

In this section, the implementation of the components of the APTIPWD is presented. Each component's implementation is described separately.

7.4.3.1. Server

The Server component was implemented using Apache HTTP Server. Apache HTTP Server is an open-source web Server for popular operating systems such as UNIX and Windows [ApacHttp]. A 1.40GHz Toshiba laptop, which runs Microsoft Windows XP home edition, was used to run the Apache HTTP Server.

The Server's sub-components, the URL Agent (UA), the Fixed List of Anti-Phishing Training Websites (FLAPTW) and the Intervention message, were linked to each other. The UA received any URL from the Proxy and directed it to either the local server (i.e. the prototype's Server) or the requested website on the Internet. This was accomplished by the virtual hosts[21] directives in Apache HTTP Server. The virtual hosts' container is a configuration file that contains all the web addresses that were served locally by the Server when requested (See Figure 31). However, this container had to be pointed by the main Apache HTTP Server's configuration file (See Figure 32).

[21] Virtual Host is defined as the practice of running more than one website, such as www.example1.com and www.example2.com, on a single machine [ApacHTTPVirtual].

```
#
# VirtualHost example:
# Almost any Apache directive may go into a VirtualHost container.
# The first VirtualHost section is used for all requests that do not
# match a ServerName or ServerAlias in any <VirtualHost> block.
#
<VirtualHost *:80>
    DocumentRoot "C:/Program Files/Apache Software Foundation/Apache2.2/htdocs"
    ServerName localhost
    ErrorLog "logs/localhost-error.log"
    CustomLog "logs/localhost-access.log" common
</VirtualHost>

<VirtualHost *:80>
    DocumentRoot C:/mysites/amazonme
    ServerName www.amazon.co.uk.me.com
    ErrorLog "logs/www.amazon.co.uk.me.com-error.log"
    CustomLog "logs/www.amazon.co.uk.me.com-access.log" common
</VirtualHost>

<virtualHost *:80>
    DocumentRoot C:/mysites/citybank
    ServerName www.citybank.co.uk
    ErrorLog "logs/www.citybank.co.uk-error.log"
    CustomLog "logs/www.citybank.co.uk-access.log" common
</VirtualHost>

<virtualHost *:80>
    DocumentRoot C:/mysites/halifaxme
    ServerName www.halifax-online.co.uk.me.com
    ErrorLog "logs/www.halifax-online.co.uk.me.com-error.log"
    CustomLog "logs/www.halifax-online.co.uk.me.com-access.log" common
</VirtualHost>

<virtualHost *:80>
    DocumentRoot C:/mysites/argosmyshop
    ServerName www.argos.co.uk.myshop.com
    ErrorLog "logs/www.argos.co.uk.myshop.com-error.log"
    CustomLog "logs/www.argos.co.uk.myshop.com-access.log" common
</VirtualHost>

<virtualHost *:80>
    DocumentRoot C:/mysites/cometonline
    ServerName www.comet-online.co.uk
    ErrorLog "logs/www.comet-online.co.uk-error.log"
    CustomLog "logs/www.comet-online.co.uk-access.log" common
</VirtualHost>
```

Figure 31: Examples of virtual hosts' directives in their container

```
# virtual hosts
Include conf/extra/httpd-vhosts.conf
```

Figure 32: Pointing virtual hosts' container in Apache configuration file

In addition, the DNS[22] host files in the Windows operating system were modified so that web browsers displayed the URL of the actual Phishing websites. As Figure 33 illustrates, the web addresses listed were pointed to the local machine IP address (127.0.0.1) so that any request to one of the addresses that arrived at the Apache HTTP Server was directed to and served by the local server. Thus, the users were not actually at risk since they used local web pages.

[22] DNS stands for Domain Name System. The DNS main task is mapping symbolic host names to their IP addresses [Friedlander et al.07].

```
# Copyright (c) 1993-1999 Microsoft Corp.
#
# This is a sample HOSTS file used by Microsoft TCP/IP for Windows.
#
# This file contains the mappings of IP addresses to host names. Each
# entry should be kept on an individual line. The IP address should
# be placed in the first column followed by the corresponding host name.
# The IP address and the host name should be separated by at least one
# space.
#
# Additionally, comments (such as these) may be inserted on individual
# lines or following the machine name denoted by a '#' symbol.
#
# For example:
#
#      102.54.94.97     rhino.acme.com          # source server
#       38.25.63.10     x.acme.com              # x client host

127.0.0.1       localhost

127.0.0.1       www.ebay-security.com
127.0.0.1       www.paypa1.com
127.0.0.1       www.online.ll0ydstsb.co.uk
127.0.0.1       www.amazon.co.uk.me.com
127.0.0.1       www.barclaysbanking.co.uk
127.0.0.1       www.halifax-online.co.uk.me.com
127.0.0.1       www.citybank.co.uk
127.0.0.1       www.capital0ne0nline.co.uk
127.0.0.1       www.co-operattivebank.co.uk
127.0.0.1       www.comet-online.co.uk
127.0.0.1       www.argos.co.uk.myshop.com
```

Figure 33: Screenshot of the modified DNS host file used for the prototype

As seen in Figure 31, the every single virtual host pointed a single location for a website pages directory stored in the Server. Thus, there was a directory for each anti-Phishing training website. As shown in Table 11, eleven websites were used. They were a fixed list of anti-Phishing training websites (FLAPTW). There were different URL syntax tricks (i.e. Phishing clues). They formed the URLs for the Phishing websites. They were as follows:

- URLs with a different domain from a well-known domain,
- URLs with misspelled known websites and
- URLs with large host names that contained a part of a well-known web addresses.

#	Anti-Phishing Training Websites	URL	Tricks
1	eBay	www.ebay-security.com	Different domain
2	Paypal	www.paypa1.com	Misspelled
3	Lloyds TSB Bank	www.online.ll0ydstsb.co.uk	Misspelled
4	Amazon	www.amazon.co.uk.me.com	Large host name
5	Barclays Bank	www.barclaysbanking.co.uk	Different domain
6	Halifax Bank	www.halifax-online.co.uk.me.com	Large host name
7	Citibank	www.citybank.co.uk	Misspelled
8	Capital One	www.capital0ne0nline.co.uk	Misspelled
9	Cooperative Bank	www.co-operattivebank.co.uk	Misspelled
10	Comet	www.comet-online.co.uk	Different domain
11	Argos	www.argos.co.uk.myshop.com	Large host name

Table 11: The fixed list of anti-Phishing training websites used in the prototype

Each one of the websites was linked to the intervention message by modifying the submission button so that it transferred the traffic to the intervention message. The intervention message was a simple HTML page adjusted by JAVA scripts to appear as a pop up window and to locate in the middle of the screen. Figure 34 presents the intervention message used in the prototype.

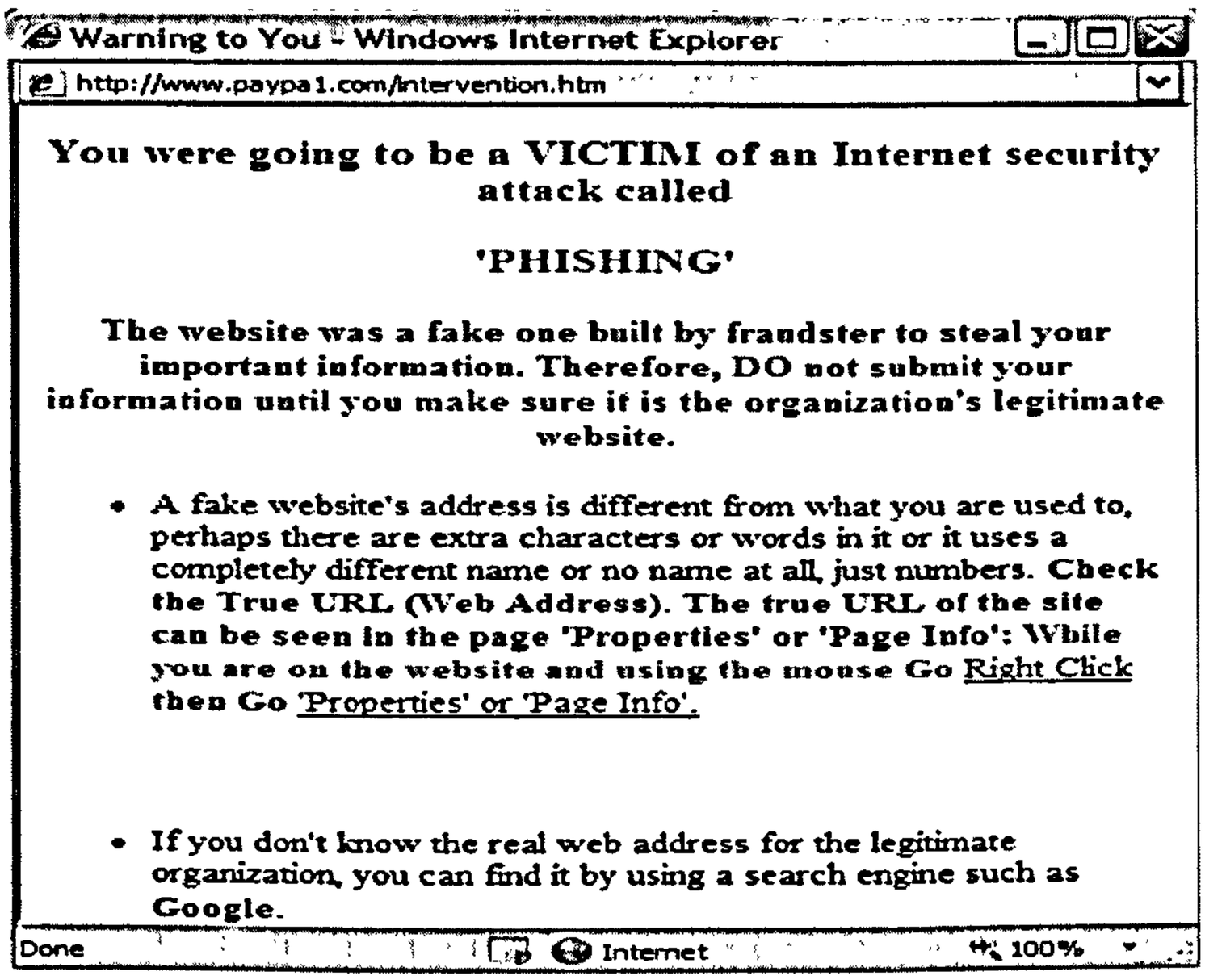


Figure 34: The intervention message used in the prototype

7.4.3.2. Proxy (Gateway)

The Proxy component was implemented using Apache HTTP Server because it has proxying capabilities that are useful and very easy to implement. The Proxy was implemented by activating the proxy module in the Server. As shown in Figure 35, the Apache HTTP Server configuration file was modified so that the proxy was able to do caching and to handle http and secure http requests. Therefore, the Proxy deals with all requests made to a specific port, which is 80.

```
LoadModule proxy_module modules/mod_proxy.so
LoadModule proxy_ftp_module modules/mod_proxy_ftp.so
LoadModule proxy_http_module modules/mod_proxy_http.so
LoadModule cache_module modules/mod_cache.so
LoadModule disk_cache_module modules/mod_disk_cache.so
LoadModule proxy_connect_module modules/mod_proxy_connect.so


<IfModule mod_proxy.c>
ProxyRequests On
AllowCONNECT 80 443
<Proxy *>
Order deny,allow
Deny from all
Allow from 192.168.1.65
</Proxy>
</IfModule>
```

Figure 35: The proxy module in the Server's configuration file

7.4.3.3. Administrator

There was no Graphical User Interface (GUI) implemented for the Administrator part. Microsoft Outlook was used instead. Microsoft Outlook has Email Accounts settings where people can provide sender name and email address. Therefore, the Administrator provided false sender name and email address that appeared as it was issued by a legitimate organization such as eBay (See Figure 36).

Due to that the fake emails were read using *Maktoob* email portal [Maktoob], the fake emails were sent by using *Maktoob*'s MX Record[23] as the outgoing mail or server. The outgoing mail settings were adjusted in Microsoft Outlook (See Figure 36).

After setting the Email Account information, the Administrator could send an email with content that looked authentic and similar to that used by a legitimate organization. As shown in Figure 37, the emails sent by the Administrator had links to anti-Phishing training websites stored and run by the Apache HTTP Server discussed previously.

---

[23] It stands for *mail exchange record*. It is an entry in a domain name database that identifies the mail server that is responsible for handling emails for that domain name. More information can be found at http://www.goecart.com/domain-name-terms-glossary.asp, last access on 19 September 2008.

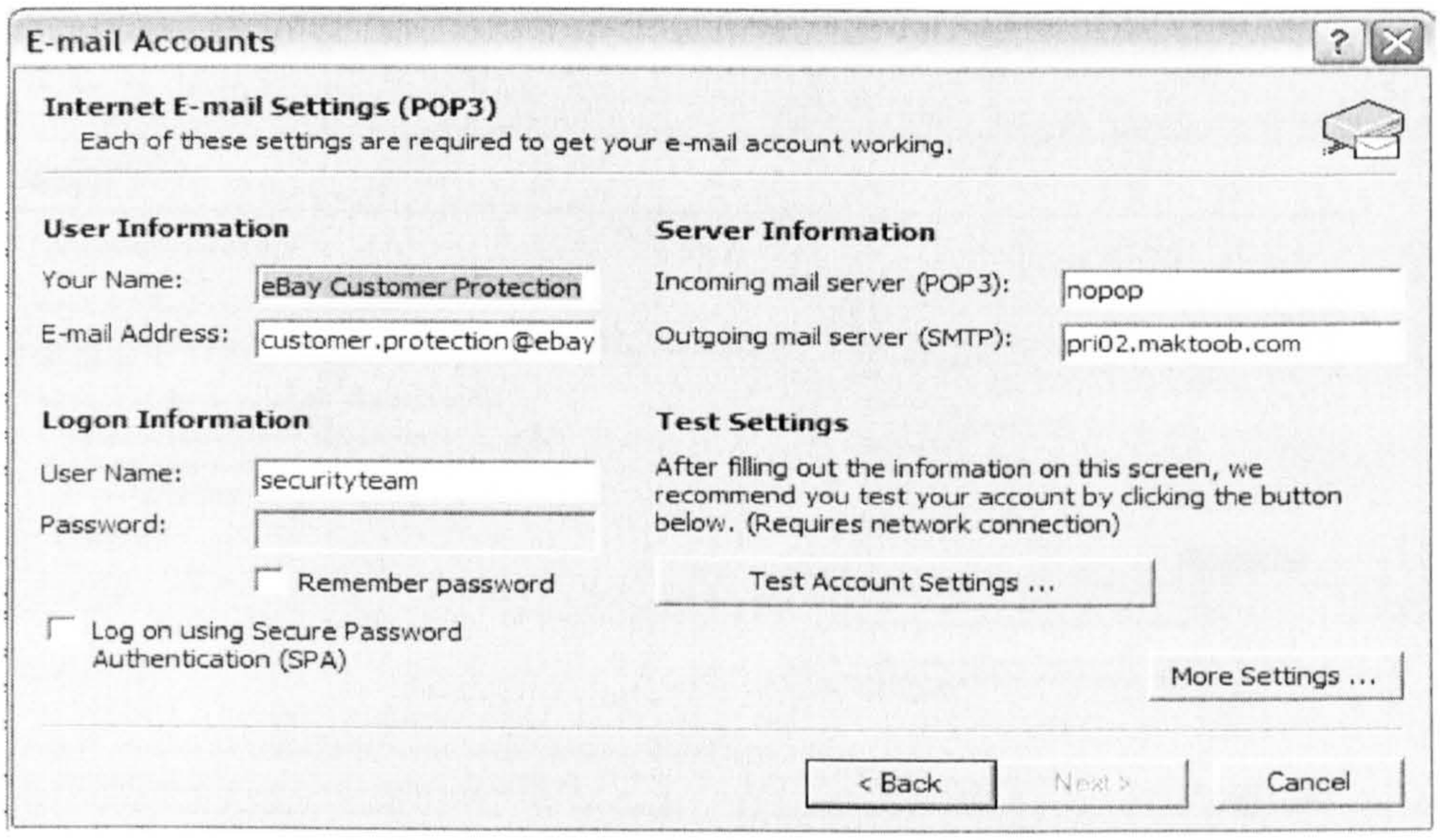


Figure 36: MS Outlook account's settings

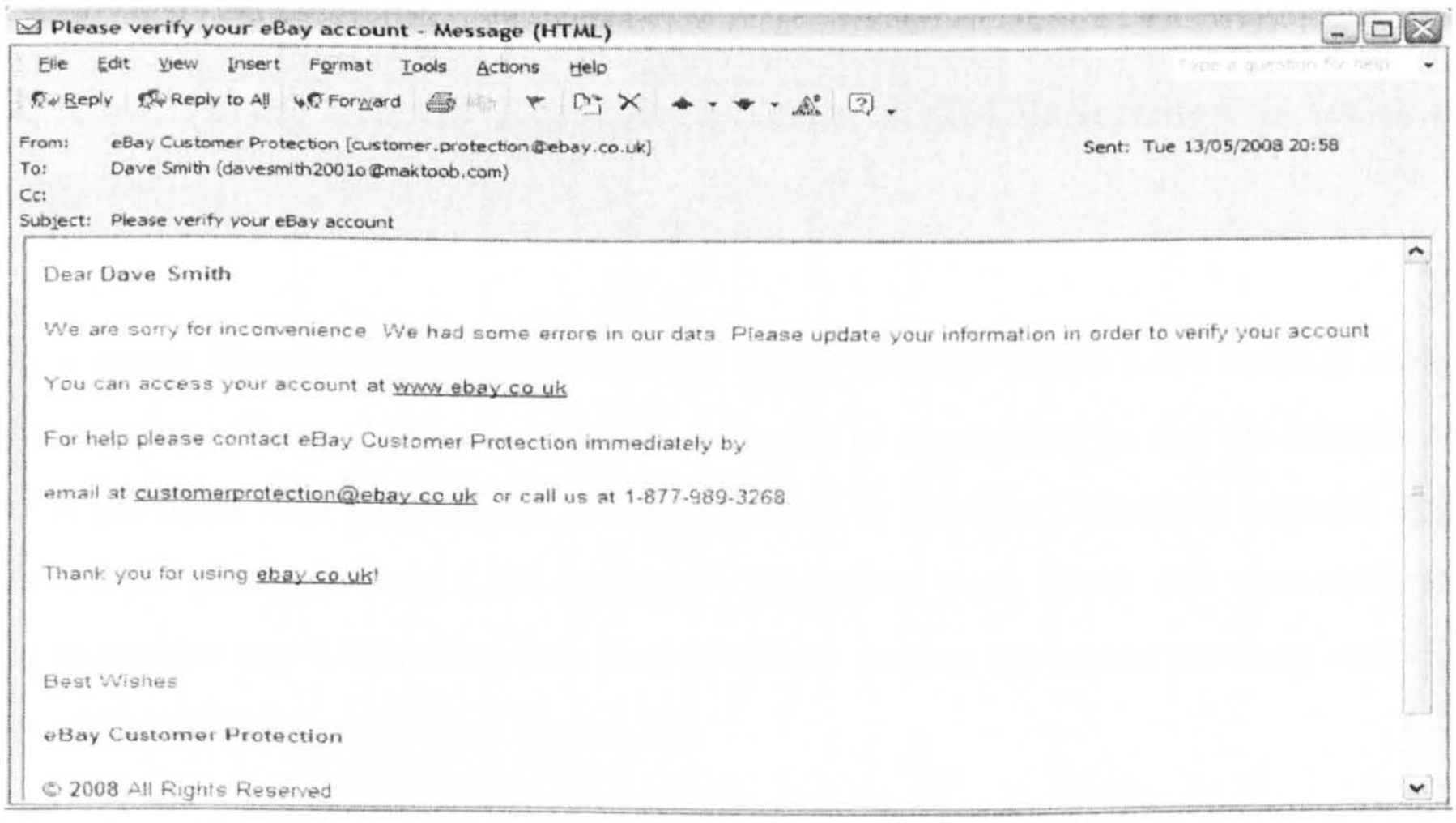


Figure 37: Example of Phishing email created and sent using MS Outlook

7.4.3.4. Client (User)

There was no implementation required for the client side of the prototype. The user used the Internet Explorer (IE) 7 browser for accessing emails and websites through *Maktoob* mail portal [Maktoob]. Figure 38 shows a screenshot of the eBay anti-Phishing training website used in the APTIPWD System.

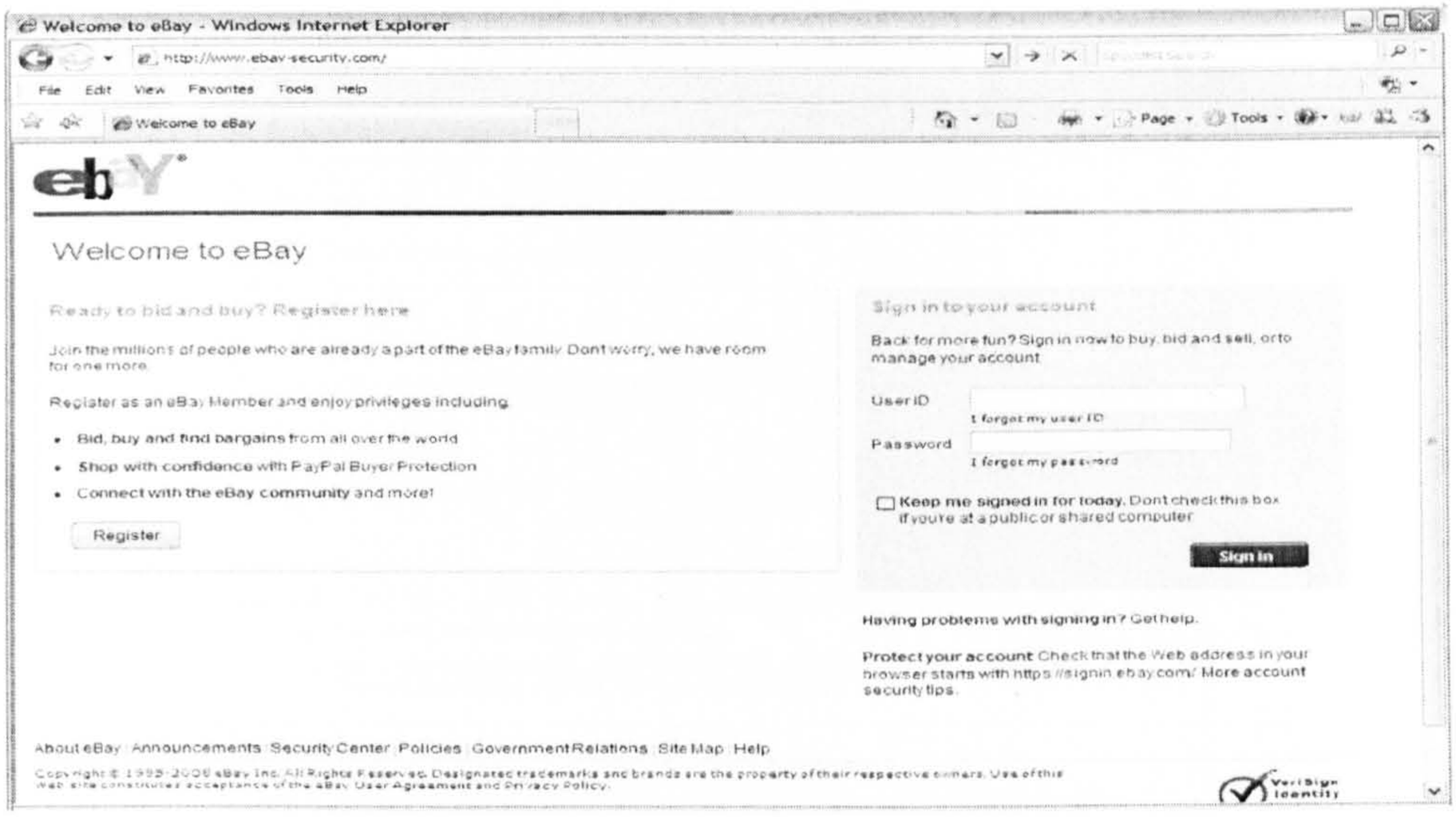


Figure 38: Screenshot of eBay-like anti-Phishing website

7.4.4. Configuring Clients' Local Area Network (LAN) Settings to Speak to the Proxy

In a proxy based computer network, the proxy settings in the LAN settings of every single machine (client) that is connected to it should be configured so that the address of the proxy is provided with its port. For example, clients in Durham University network applied the university proxy in their LAN settings[24]. Therefore, each client was connected to the Proxy to request any URL. This was carried out by putting the server machine as its LAN proxy on the default port 80 (See Figure 39).

[24] Computer network settings in Durham University. Available at: http://www.dur.ac.uk/its/services/network/lan/quicksettings, last access on 2 December 2008.

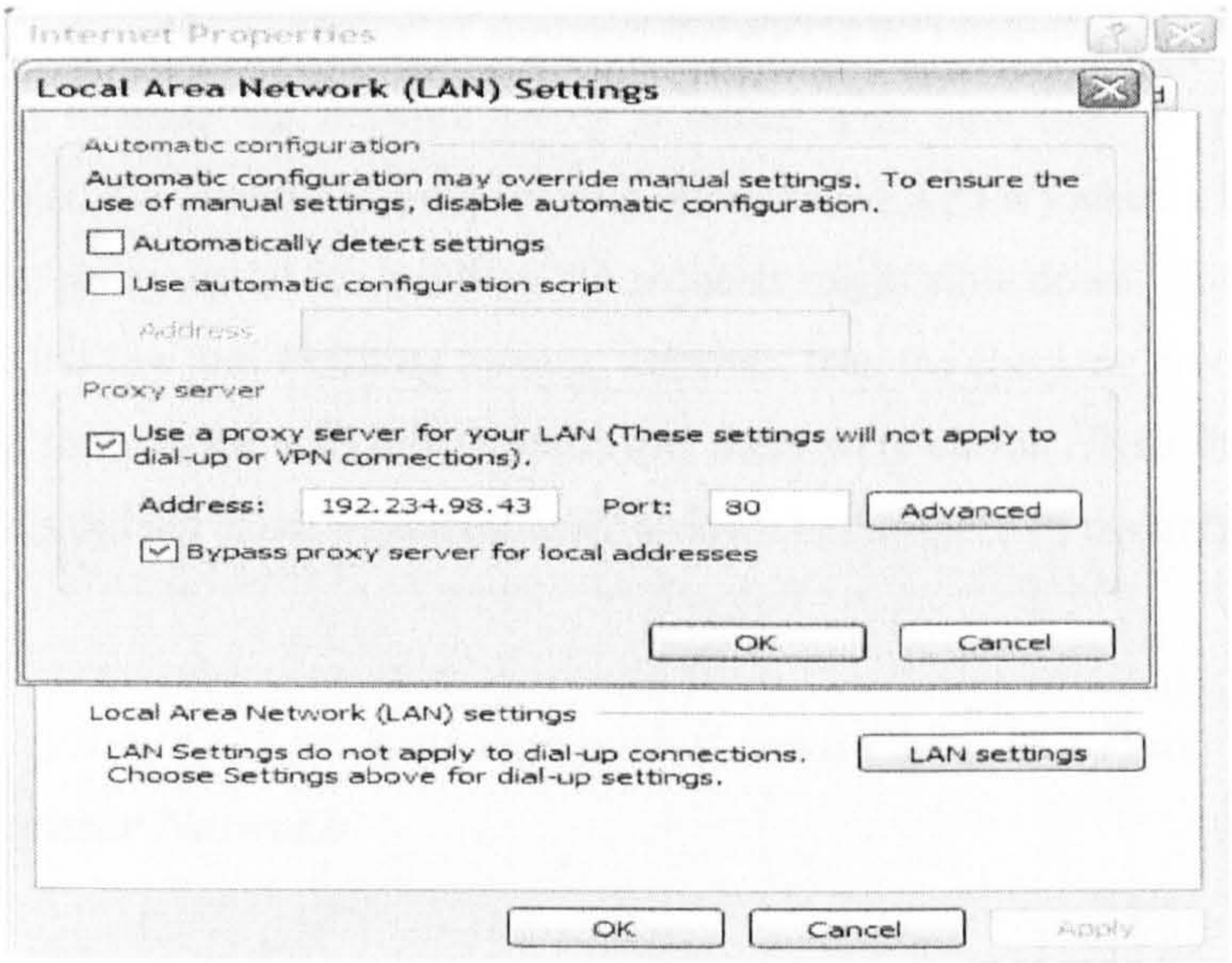


Figure 39: The Internet Explorer's LAN settings

7.4.5. Discussion

7.4.5.1. Advantages and Limitation

Applying the APTIPWD to a proxy based computer network has advantages and a possible limitation. The advantages can be summarized as follows:

1. It is easy to implement.
2. There is no need to write any programming code.
3. It is a browser independent tool. Thus, there is no specific browser that is required for the tool to be run.
4. Since the training is sent by email, the Administrator is able to send anti-Phishing training to specific users.
5. The aim of training users without informing them that it is anti-Phishing training is satisfied. Therefore, the limitation of using the role-play protocol in anti-Phishing training is resolved.
6. The aim of training users while they normally use the Internet is satisfied.

In contrast, a possible limitation of applying the APTIPWD to a proxy based computer network is that because the network proxy is added with new tasks to perform (i.e. checking the fixed list of anti-Phishing training websites (FLAPTW) when a URL request is received), the proxy speed for handling the requests might slow down. However, when the APTIPWD has few anti-Phishing training websites, then the checking process does not consume much time. In the APTIPWD prototype, there were eleven URLs that needed to be checked. This did not cause a noticeable slow down to the speed of the traffic.

7.4.5.2. Deploying the New Approach with its own Proxy in a Proxy based Computer Network

Applying the New Approach to an existing proxy based computer network has been described. This means that the proxy used in applying the New Approach is the network proxy that handles the URLs requests made by the network's clients. The proxy needs to be configured to communicate with the Apache HTTP Server and the clients.

In addition to this, the New Approach can be applied to a proxy based computer network (in this instance, Durham University network) without configuring its proxy. This was accomplished by having a proxy only for running the New Approach. This meant that there were two proxies when the New Approach was running; the Durham University network's proxy and the New Approach's own proxy. The New Approach's proxy was planted between the Durham University network's proxy and the Client. For this to be done, a simple alteration to the Apache HTTP Server configuration file, shown in Figure 35, was performed. As presented in Figure 40, the New Approach's proxy forwarded all URLs requests to the University proxy unless the URLs requested were listed to be served in the New Approach local server.

```
LoadModule proxy_module modules/mod_proxy.so
LoadModule proxy_ftp_module modules/mod_proxy_ftp.so
LoadModule proxy_http_module modules/mod_proxy_http.so
LoadModule cache_module modules/mod_cache.so
LoadModule disk_cache_module modules/mod_disk_cache.so
LoadModule proxy_connect_module modules/mod_proxy_connect.so


<IfModule mod_proxy.c>
ProxyRequests On
AllowCONNECT 80 443
<Proxy *>
Order deny,allow
Deny from all
Allow from 192.168.1.65
</Proxy>
</IfModule>



ProxyRemote  *  http://wwwcache.dur.ac.uk:8080
NoProxy         www.ebay-security.com www.paypal.com www.online.llOydstsb.co.uk www.amazon.co.uk.me.com
```

Figure 40: Pointing the Durham University's proxy in the Server's configuration file

7.5. Summary

In this chapter, a novel Anti-Phishing Approach that uses Training Intervention for Phishing Websites Detection (APTIPWD) was proposed and discussed. The New Approach presents an intervening message to users who access Phishing websites and try to submit their information. The intervention message uses the most effective anti-Phishing tip evaluated in Chapter 6. By using this approach, users do not need to attend training courses and do not need to access online training materials. This is because the approach brings information to end-users and helps them immediately after they have made a mistake in order to detect Phishing websites by themselves.

Due to the fact that the blacklists component is dynamic and therefore is hard to control, evaluating the New Approach on the real Internet is difficult. A better solution is to evaluate under experimental conditions. In order to evaluate the approach under experimental conditions, all possible scenarios of the approach were simulated and the blacklists (dynamic components) were made fixed.

A prototype proof of concept implementation of the New Approach was presented. It also showed that the New Approach is feasible and can be implemented easily without writing a single line of a programming code and without undue disruption of the users system.

8. Experiments

8.1. Introduction

In this chapter, the evaluation experiments are considered. The hypotheses and their themes are discussed together with the way in which the experiments' participants were recruited and their demographic information. Effectiveness ratios that are used in evaluating the hypotheses are defined. The chapter also reviews comparisons between real Phishing attacks and Phishing experiments in order to decide what should be simulated in the experiments. It then concludes with a discussion on the story board of the experiments.

8.2. Hypotheses and Themes

Before discussing the themes and hypotheses, it is useful to define few terms. The New Approach is an Anti-Phishing approach that uses Training Intervention for Phishing Websites Detection (APTIPWD), discussed in Chapter 7. The Old Approach is the current practice of sending anti-Phishing tips by email, as discussed in Chapter 3.

High Technical Ability (HTA) people are those who are considered experts in terms of computer technical ability. In contrast, Low Technical Ability (LTA) people are those who are considered non-experts in terms of computer technical ability. The criteria for classifying experts and non-experts as well as Phishing aware and Phishing unaware people will be discussed in Section 8.3.2.

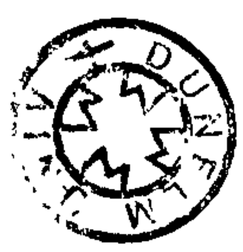
The hypotheses in this section are experimental hypotheses. The null hypotheses are shown in Chapter 9.

8.2.1. Theme 1: Evaluating the New Approach

Hypothesis 1: There is no difference between using the New Approach and the Old Approach in helping people recognize legitimate websites and detecting Phishing websites.

8.2.2. Themes 2 and 3: Technical Ability and Phishing Knowledge

A number of anti-Phishing approaches have been evaluated using participants who were recruited based on their technical abilities [Downs et al.06, Kumaraguru et al.07a, Kumaraguru et al.07b, Sheng et al.07]. Participants were classified into 'expert' and 'non-expert' users based on pre-study screening questions. Technical ability was judged on whether the participants had changed preferences or settings in their web browser, created a web page, and helped someone fix a computer problem. Any participant who said 'no' to at least two of the screening questions was selected to take part in their experiments. This technical ability assessment was used to recruit people with low technical ability (they called them 'non-experts'). However, the participants who were considered non-experts could know about Phishing and how to detect attacks before participating in the evaluation experiments. Having participants with Phishing knowledge in advance may provide biased results in evaluation experiments on anti-Phishing approaches. This is because people who know about Phishing before participating in the evaluation experiments may use their own Phishing knowledge rather than the anti-Phishing approaches of the evaluation in which they are participating. Downs et al. [Downs et al.07] studied whether there are correlations between some web environment experiences and susceptibility to Phishing. They found that people who correctly answered the knowledge question about the definition of Phishing (i.e. Phishing aware people) were significantly less likely to be deceived by Phishing emails. Low technical users may be Phishing aware and high technical users may be Phishing unaware.



It is necessary to make sure that the participants do not know about Phishing regardless of their technical ability level. Therefore, a research hypothesis is expressed as follows:

Hypothesis 2: In evaluating an anti-Phishing approach, it is better to recruit subjects based on their Phishing knowledge rather than their technical ability.

The hypothesis has two main issues. They are technical ability and Phishing knowledge. In order to find out which issue has an effect on people's decisions on legitimate and Phishing websites, each issue must be assessed separately. The hypotheses *2.1* and *2.2* discuss the technical ability and Phishing knowledge respectively.

8.2.2.1. Theme 2: The Effect of High and Low Technical Abilities on Phishing Websites Detection

Hypothesis 2.1: There is no difference between high technical people and low technical people in recognizing legitimate websites and detecting Phishing websites.

8.2.2.2. Theme 3: The Effect of Phishing Awareness and Phishing Unawareness on Phishing Websites Detention.

Hypothesis 2.2: Phishing aware people are better than Phishing unaware people in recognizing legitimate websites and detecting Phishing websites.

8.2.3. Theme 4: Anti-Phishing Knowledge Retention

If the New Approach, which uses training intervention, demonstrates that users are better than the users of the Old Approach of sending anti-Phishing tips by email in detecting Phishing attacks when they are evaluated immediately after they are trained, the question arises about whether the New Approach users can retain the knowledge that they gained

during training after a period of time better than the Old Approach users. As a result of this, a hypothesis is expressed as follows:

Hypothesis 3: People who used the New Approach retain their anti-Phishing knowledge better than people who used the Old Approach of sending anti-Phishing emails.

In evaluating these hypotheses, other hypotheses are extracted from them and are shown in Chapter 9. The extracted hypotheses are then evaluated and analyzed.

8.3. Recruiting Participants and Demographic Information

Before running the evaluation experiments, the participants had to be classified according to their technical ability and their Phishing knowledge. Regarding their technical ability, they were classified into two categories, high and low. In terms of their Phishing awareness, they were classified into two categories, Phishing aware and Phishing unaware.

In order to do these classifications, a pre-study survey was conducted. There were both online and offline surveys to be answered by respondents. Invitation posters to participate in the experiments were distributed in different places on the Durham University campus. Invitations were also distributed to the university's students by emails using the colleges' mailing lists.

There was time gap between participants filling in the pre-study survey and their participation in the experiments. In this gap, the participants' technical ability and Phishing awareness situations might have changed from low to high and from unaware to aware respectively. Due to this, a 'pre-session survey' for the participants just before participation in the experiments was conducted. This was to make sure that the information given by the participants in the pre-study survey was still valid.

8.3.1. Pre-Study Survey

The survey was built from multiple resources. It asked questions about Internet and email usage, participant's technical ability, their web browser knowledge and computer terminology. The computer terminology section included the question about Phishing knowledge. The questions about Phishing knowledge and participant's technical ability were the main concerns in the survey.

Initially, potential participants were asked to provide their email addresses so that they could be contacted if they were selected to take part in the study. Also, participants were not asked about their demographic information in the pre-study survey in order to save their time and because they might be deterred from filling in the survey if it took more than 10 minutes. Therefore, the participants were asked about their demographic information just before taking the study.

The pre-study survey is presented in Appendix B. An overview of the survey sections is discussed as below.

8.3.1.1. Internet and Email Usage

Participants were asked questions about their email usage and skills [Health e-Tech]. In addition, they were asked questions about their online transactions experience [Downs et al.07]. The reason for having this section was to convey the idea that the experiment was just a study about the participant's use of email systems and Internet.

8.3.1.2. Technical Ability

The participants were asked questions on computer technical tasks in order to assess their technical ability. The questions were as follows [Sheng et al.07]:

- Have you changed preferences or settings in your web browser?
- Have you created a web page?
- Have you helped someone fix a computer problem?

The goal of having this section was to classify the participants into low and high with regards to their technical abilities.

8.3.1.3. Web Browser Knowledge

Participants were asked questions about their knowledge of URLs (i.e. interpreting the URLs syntax) and padlock icons. Participants were shown an image of the padlock icon found within the browser chrome and were asked whether they had seen "this padlock image" before (See Figure 41) [Downs et al.07].

Figure 41: The padlock image

The goal of having this section was to reinforce the idea that the experiment was a study about the use of email systems and the Internet.

8.3.1.4. Computer Terminology

Participants were asked to choose the best definition for six computer related terms [Downs et al.07]. They were cookie, spyware, Google, virus, messenger and Phishing. Participants were given the same list of ten possible definitions to choose from for each definition, as well as options to indicate familiarity with the word or not. Each term had one correct answer on the list. The goal of having this section was to classify the participants into Phishing Aware and Phishing Unaware in terms of their Phishing knowledge.

8.3.2. Classification Criteria

The selection of the experiment's subjects was based on their answers on the pre-study survey. Because 'the Internet and email usage' and 'web browser knowledge' sections were

included simply to convey the idea that the experiment was a study just about the participant's use of email systems and Internet, these sections were not included in the selection criteria. The questions about Phishing knowledge and technical ability are the main concerns in the survey. The participants were not told that the experiments were about Phishing. Therefore, the answers to the survey questions were used to classify the survey respondents as:

- Low or high technical people in terms of technical ability and
- Phishing Aware or Phishing Unaware.

Regarding the technical ability questions, respondents who say 'no' to more than one of the three questions were considered 'low technical people'. Otherwise, the respondents were considered 'high technical people'.

In terms of Phishing awareness, the section *Computer Terminology* (8.3.1.4) included a Phishing definition question. Those who defined Phishing correctly were regarded as 'Phishing Aware'. Otherwise, the respondents were considered as 'Phishing Unaware'.

8.3.3. Participants

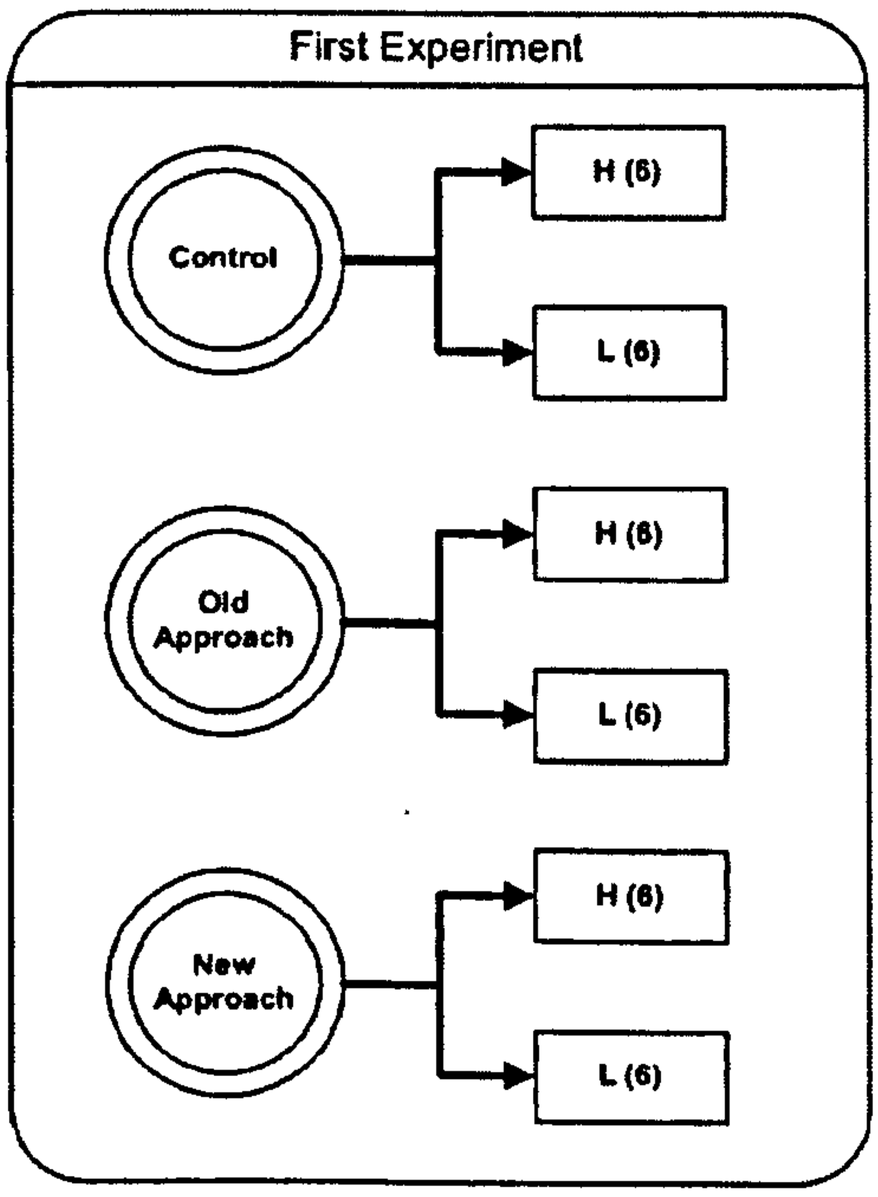
As a result of the pre-study survey invitations, 219 people responded to the survey. Of these, 13 skipped the survey's questions, providing their names and contacts only. Therefore, they were excluded from the experiments.

With regards to participants' classifications, 133 out of 206 were 'Phishing Aware', which represent 64.6% of all respondents. This means that they knew the Phishing definition. In contrast, 73 out of 206 were 'Phishing Unaware'. This means that they did not know the Phishing definition and represented 35.4% of all respondents. In terms of the respondents' technical ability, 125 out of 206 were 'high technical ability', which represent 60.68%. However, 81 out of 206 were 'low technical ability', representing 39.32%.

In terms of the pre-session survey given to the participants before starting the participation in the experiments, two participants had changes in their technical ability and

Phishing awareness. One of them had high technical ability based on the information given in the pre-session survey though classified as low technical ability from the pre-study survey. The other participant was Phishing Aware according to the pre-session survey whereas the participant was Phishing Unaware as suggested by the information given in the pre-study survey.

8.3.4. Demographic Information



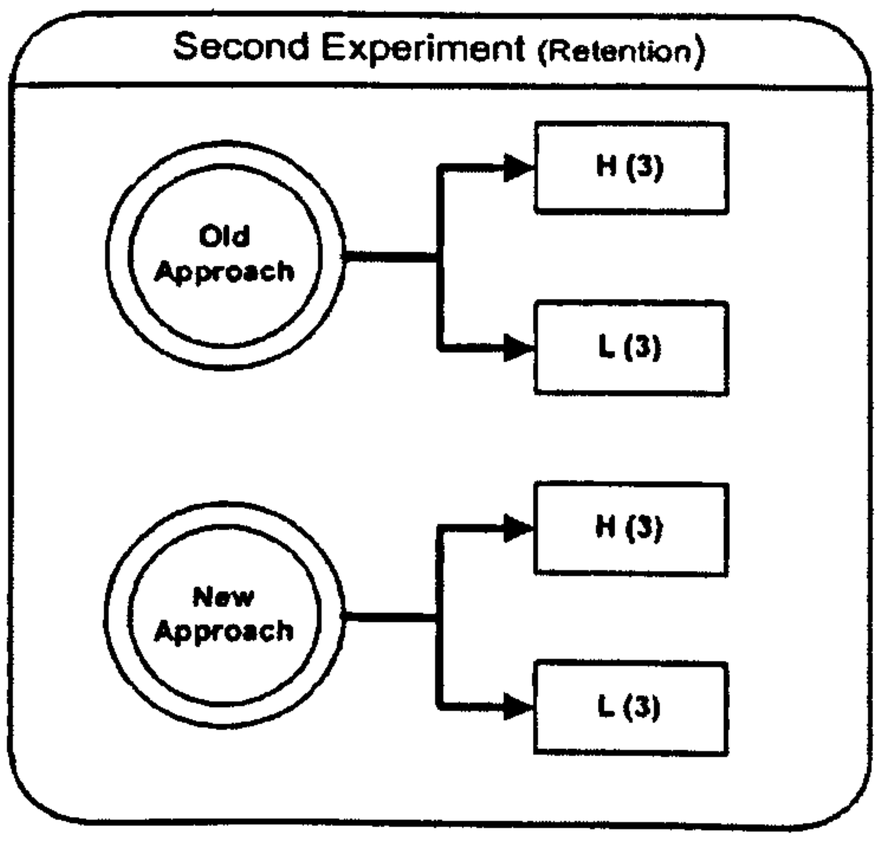
Keys:

H means people with High technical ability.
L means people with Low technical ability.
(n) the number between brackets indicates the number of participants.

The websites' and treatment's order followed by each group.

Order

Order	Trick	
Phishing	Long Host names	Pre-Treatment
Friend	N/A	
Work	N/A	
Legitimate	N/A	
Phishing	Misspelled Website	
Work	N/A	
Legitimate	N A	
Treatment	N/A	
Legitimate	N A	Post Treatment
Phishing	Misspelled Website	
Friend	N/A	
Legitimate	N/A	
Friend	N/A	
Phishing	Different Domain	



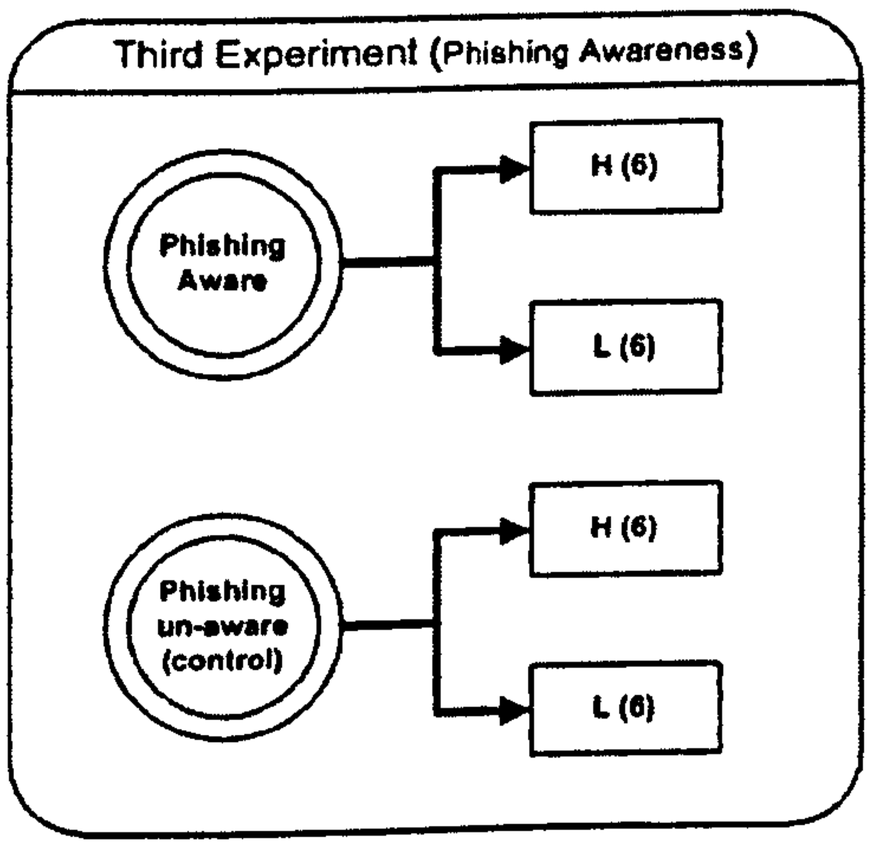


Figure 42: The three experiments & the websites' and treatment's order

For the evaluation experiments, three different experiments were carried out and their overview is shown in Figure 42. A total of 48 different people participated in the experiments and Table 12 shows their demographics. In all three experiments, participants were randomly placed in the experiments' groups.

Age	**18-25**	**42** (87.5%)
	26-32	**5** (10.4%)
	33-39	**1** (2.1%)
	40-46	**0**
	46+	**0**
Gender	**Male**	**16** (33.33%)
	Female	**32** (66.66%)
Language	**Native English Speaker**	**42** (87.5%)
	Non-native English Speaker	**6** (12.5%)
Level of Study	**Undergraduate**	**37** (77.1%)
	Postgraduate	**11** (22.9%)

Table 12: The demographics of the total subjects participated in the three experiments

In the first experiment, to assess the anti-Phishing approach, there were three groups, Control, Old Approach and New Approach. There were 36 participants in the experiment. Each group had 12 participants divided into two subgroups, High and Low technical ability. Thus, each subgroup had 6 participants. Table 13 presents the demographics of the first experiment's participants.

Age	**18-25**	**31** (86.1%)
	26-32	**4** (11.1%)
	33-39	**1** (2.8%)
	40-46	**0**
	46+	**0**
Gender	**Male**	**12** (33.33%)
	Female	**24** (66.66%)
Language	**Native English Speaker**	**31** (86.1%)
	Non-native English Speaker	**5** (13.9%)
Level of Study	**Undergraduate**	**29** (80.6%)
	Postgraduate	**7** (19.4%)

Table 13: The demographics of the subjects participated in the first experiment

The second experiment, to assess anti-Phishing knowledge retention, is called also the Retention experiment. This experiment had two groups, Old Approach and New Approach. There were 12 participants in the experiment. Each group had 6 participants divided into

two subgroups, High and Low technical ability. Thus, each subgroup had 3 participants. All subjects in the second experiment also participated in the first experiment. This means that they participated in the first experiment and then, after a period of time, they were called back and asked to participate in the second experiment. The period between the two experiments varied from subject to subject. However, the average period was 16.7 days. Table 14 presents the demographics of the second experiment's participants.

Age	**18-25**	**11** (91.7%)
	26-32	**1** (8.3%)
	33-39	**0**
	40-46	**0**
	46+	**0**
Gender	**Male**	**5** (41.7%)
	Female	**7** (58.3%)
Language	**Native English Speaker**	**10** (86.1%)
	Non-native English Speaker	**2** (13.9%)
Level of Study	**Undergraduate**	**11** (91.7%)
	Postgraduate	**1** (8.3%)

Table 14: The demographics of the subjects participated in the second experiment (Retention)

The third experiment is named the Phishing Aware experiment. 24 participants were divided into two groups, Phishing Aware participants and Phishing Unaware participants (Control). Each group had 12 participants. Each group was divided into two subgroups, High and Low technical ability. Thus, each subgroup had 6 participants. Table 15 shows the demographics of the third experiment's participants.

Age	**18-25**	**20** (83.3%)
	26-32	**3** (12.5%)
	33-39	**1** (4.2%)
	40-46	**0**
	46+	**0**
Gender	**Male**	**8** (33.33%)
	Female	**16** (66.67%)
Language	**Native English Speaker**	**21** (87.5%)
	Non-native English Speaker	**3** (12.5%)
Level of Study	**Undergraduate**	**16** (66.67%)
	Postgraduate	**8** (33.33%)

Table 15: The demographics of the subjects participated in the third experiment (Phishing Aware)

In total, 48 participants took part in the experiments. 60 trials (i.e. experiment runs) were carried out (48 participants, of whom 12 participated in both the first experiment and the Retention experiment). While this might seem limited, the numbers for an experiment of this type quickly grow to unmanageable sizes. A decision was taken to conduct small well-controlled experiments and hence the experiments setup described.

There were two different blocks in the experiments, Technical Ability (TA) and Phishing Awareness (PA). Each block had two different levels. Technical Ability had High and Low whereas the Phishing Awareness had Aware and Unaware. The first experiment had three groups, Control, Old Approach and New Approach. In order to compare these groups properly, it was necessary to include the blocks Technical Ability (TA) and Phishing Awareness (PA) in each group. As an initial number, 10 participants for each level in each block in each group were set as shown in Table 16. This means 40 participants were required in each group. This implies 120 participants were required for the experiments. This number was considered to be unmanageable in size and requiring too much effort in time and funds to be conducted.

Technical Ability / Phishing Knowledge	High	Low	Total
Aware	10	10	20
Unaware	10	10	20
Total	20	20	40

Table 16: The initial size for each group in the first experiment

Then, the number of participants in each level was reduced several times until it was set as 6 participants. This means that 24 participants were required in each group. Consequently, 72 participants were needed for the experiments. Due to the fact that only Phishing Unaware people were needed to participate in the first experiment, Phishing Aware people were excluded. Therefore, there were 6 participants in each Technical Ability level, which in turn means 12 participants were required in each group. Thus, 36 participants were asked to participate in the first experiment.

In the retention experiment, as many of the first experiment' participants who were available came back to take part. For the third experiment, Phishing Aware people were

required to participate and their results were needed to be compared with the results of Phishing Unaware people (Control group in the first experiment). Therefore, a decision was taken to have a similar number to the Control group (12 participants) considered 'Phishing Aware'. There were 6 participants in each Technical Ability level. Then, they were asked to take part in the experiment. In total, there were 48 participants and their distribution is presented in Table 17.

Phishing Knowledge \ Technical Ability	High	Low	Total
Aware	6	6	12
Unaware	18	18	36
Total	24	24	48

Table 17: The final sample size for all groups participated in the experiments

8.4. Effectiveness Ratios

In this section, the effectiveness ratios that were used in evaluating the hypotheses are described.

8.4.1. Decisions for Website Legitimacy

8.4.1.1. Definitions

There were four types of decisions. They are, as shown in Table 18, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). They are defined as follows:

User Decision / *Website Legitimacy*	**True**	**False**
Positive	***TP***	***FP***
Negative	***TN***	***FN***

Table 18: The possible decisions could be made regarding websites' legitimacy

True Positive (TP): The TP case happens when a legitimate website is considered as legitimate.

True Negative (TN): The TN case happens when a Phishing website is considered as Phishing.

False Positive (FP): The FP case happens when a legitimate website is considered as Phishing.

False Negative (FN): The FN case happens when a Phishing website is considered as legitimate.

8.4.1.2. Decision vs. Result's Type

The four different decisions have different types of results in terms of their effects on the end-user's sensitive security information. The types of the decision's results are illustrated in Table 19.

Decision	Type of Result (Good Decision?)
TP	Correct.
TN	Correct.
FP	Incorrect. It is a matter of *inconvenience* since the user does not hand in their sensitive information to legitimate website because of their fears.
FN	Incorrect. It is considered as *the most undesirable* result since the user hands over their sensitive information to fraudsters.

Table 19: Decisions vs. results' types

8.4.2. *Ratios*

8.4.2.1. Calculation

The ratios described here were used in evaluating the experiments. The end-user's decisions on judging a website could be either correct or wrong. The Correct Decision means either submitting information to a legitimate website or not submitting information to a Phishing website. Otherwise, the decisions are classified as wrong ones. Therefore, the first ratio is the Correct Decision Rate (CDR). The CDR is calculated as shown in Formula 1.

$$CDR = \frac{NumberOfCorrectDecisions}{NumberOfWebsites} \qquad (1)$$

The other two ratios used in the evaluation are False Positive Rate (FPR) and False Negative Rate (FNR). The FPR's and FNR's calculations are shown in Formulas 2 and 3 respectively.

$$FPR = \frac{NumberOfFalsePositives}{NumberOfLegitimateSites} \qquad (2)$$

$$FNR = \frac{NumberOfFalseNegatives}{NumberOfPhishingSites} \qquad (3)$$

8.4.2.2. Values

The values of the three ratios (CDR, FPR and FNR) are between 0 and 1. Because CDR is based on correct answers, a higher CDR result from the experiments is better than a lower value. In contrast, because FPR and FNR are based on wrong answers, the lower FPR and FNR the better is the result.

8.4.2.3. Ratios' Use in Evaluating the Hypotheses

Each hypothesis in the evaluation is stated based on the three ratios (CDR, FPR and FNR). This means that each hypothesis then becomes three different hypotheses because there is one hypothesis for each ratio.

The three ratios were used for different purposes in evaluating the research hypotheses. The Correct Decision Rate (CDR) is the main and the decisive ratio. Therefore, the final result is based on the CDR comparisons. However, the comparisons of False Positive Rate (FPR) and False Negative Rate (FNR) were used to give clear descriptions of the decisions made about the legitimate websites and Phishing websites respectively. The False Positive Rate (FPR) was based on the legitimate websites and the False Negative Rate (FNR) was based on the Phishing websites.

8.5. Comparisons between Real Phishing Attacks and Experiments

In the real world, a variety of issues are involved in the majority of Phishing attacks. These include the following:

1. The user does not know that they have received a Phishing attack.
2. The Phishing attack has some clues that are indicative that it is a Phishing attack. For example, an eBay user may be advised to follow a link spelled *www.paypa1.com* (the '*l*' of 'pal' is replaced by the digit one '1') and provide their information believing it to be the genuine Paypal website.
3. The user may reveal their sensitive information to the phisher via a fake email or website.

4. In context-aware Phishing attacks[25], the user may receive an email that uses their contextual information such as the user's real name.

For the Phishing experiments, in order to simulate real Phishing attacks, there should be some considerations of the following issues:

1. The subjects should not have knowledge about Phishing before taking part in the experiment. This is because knowing about Phishing in advance will affect the subjects' behaviors. Thus, having subjects who are considered 'unaware' about Phishing is better than having aware subjects. Therefore, the more accurate the selection of the unaware subjects, the more accurate will be the results of the experiment.
2. The subjects should not know that they are being tested about Phishing attacks.
3. The Phishing attacks in experiments should be similar to the real ones. This means that the Phishing clues in emails or web browsers should be similar to the real ones.
4. The experiment should not put the participants at any risk. This means that their sensitive information (passwords, PINs, credit card details, and so on) should be safe, secure and anonymous to anyone, even to the experimenters themselves [JakobssonRatkiewicz06].
5. In context-aware Phishing experiments, the attacks should use some contextual information about the participants.
6. The dependent and independent variables, and the metrics to be calculated, need to be clear.

8.6. Methodology

In order to evaluate the research hypotheses, a pilot study was undertaken and then three experiments were conducted.

---

[25] A context-aware Phishing attack happens when the phisher gains knowledge (name, date of birth, part of credit card number, etc.) about the victim and then use it to customize an attack that appears to be from a genuine website [RobilaRagucci06].

8.6.1. Pilot Study

8.6.1.1. Objective

In order to carry out well-designed evaluation experiments, a pilot study was run before the experiments. The main benefit of having a pilot study was to discover the mistakes and errors that could occur in the experiment. Subsequently, the errors were corrected in order to have well-designed experiments.

8.6.1.2. Scenario Overview

A pilot study that involved 8 participants was carried out. The participants used the experimental scenario as well as the software of the proposed anti-Phishing approach. There were no different groups. All participants had the same scenario. Each participant performed in a separate session of nearly 20 minutes. The scenario, email and websites used, were similar to the scenario used in the first experiment and shown in the next section (8.6.2).

8.6.1.3. Errors

The errors that occurred in the pilot study are divided into two types, technical and procedural.

I. Technical Errors

The technical errors were related to the technologies used in the experiments. There were few technical errors, such as the email client (*Maktoob* email portal [Maktoob]) marking the emails as unsafe. So the emails were signed suspicious and the client did not show the organization's legitimate logo in the content of the emails. Additionally, the firewall running in the machine popped up messages.

II. Procedural Errors

The procedural errors were related to the experiment procedures and scenarios. There were few procedural errors, such as the Halifax password written on the Scenario

Information Sheet (See Appendix B) was not completed. One error also was that a participant checked the emails randomly (i.e. not in order).

8.6.1.4. Debugging

All the errors reported in the pilot study were resolved. The technical errors were fixed and the procedural errors were corrected and re-designed.

8.6.2. The First Experiment

The experimental participants undertook email and web role-play protocol. The use of role-play in the experiment, while not being ideal, does give a close approximation to real world behaviour [Downs et al.07]. The evaluation protocol was used successfully in other studies [Downs et al.06, Kumaraguru et al.07a]. Participants were asked to deal with emails because Phishing websites are usually reached through emails that ask users to click on a link. Each participant played the role of an imaginary person named "Dave Smith". Dave Smith is an employee of a company and works in the marketing department. Participants were asked to interact with the emails and websites in the way they would normally do. Participants were told that the experiment would investigate "how people effectively manage and use the Internet and emails". They undertook a pre-study survey about their email usage to enforce the idea that this was an experiment about their use of email systems and the Internet. All participants in this experiment were considered 'Phishing Unaware' because it was necessary to have participants with no knowledge about Phishing in evaluating the New Approach. Having participants with Phishing knowledge in advance may provide biased results in Phishing experiments. People who know about Phishing (i.e. Phishing Aware) might use their own knowledge to detect a Phishing website rather than the knowledge they receive from the New Approach.

The study was recorded (audio and screen) using Camtasia software in order to re-play the experiments for further analysis if required.

Accounts with some well-known organizations such as eBay, PayPal and Amazon were created for the user, Dave Smith. Participants were given an information sheet that included a description of the experiment's scenario as well as usernames and passwords for the employee's accounts at the organizations (See Appendix B).

The experiment was divided into two parts: pre-treatment and post-treatment. In the pre-treatment part, all participants in all groups dealt with the emails and websites without having their treatment. In the post-treatment part, participants had different treatments according to the group in which they had been placed. The groups were as follows:

- **Control group:** In this group, the treatment was an email from work (in this instance, an ordinary email from work, essentially a null treatment).
- **Old Approach group:** In this group, the treatment was anti-Phishing tips sent by email. It was an email with online training material on Phishing.
- **New Approach group:** In this group, the treatment was the New Approach.

Each participant was shown 13 email messages (7 messages are for the pre-treatment part and 6 are for the post-treatment part). Five messages were legitimate email messages, with no embedded links, that Dave Smith received from colleagues at his company and friends. These messages were just to re-enforce the idea that the experiment was about how people effectively use and manage the Internet and email. Dave Smith was expected to perform simple tasks such as replying. The other 8 email messages (implies 8 related websites) were divided into 4 simulated legitimate emails from organizations with which Dave Smith had an account and 4 Phishing emails. In the New Approach group, the intervention was run based on visiting a blacklisted (Phishing) website. This website is Paypal Phishing website (See Table 20).

The order of the emails and websites for both the pre-treatment and post-treatment parts were predefined in all groups and are shown in Table 20. The emails and websites highlighted are the groups' different treatments.

#	eMails and Websites Order			Website	Tricks	URL
	Control	Old Approach	New Approach			
1	Phishing	Phishing	Phishing	Amazon	Large Host names	www.amazon.co.uk.me.com
2	Friend	Friend	Friend	N/A	N/A	N/A
3	Work	Work	Work	N/A	N/A	N/A
4	Legitimate	Legitimate	Legitimate	Halifax	N/A	www.halifax-online.co.uk
5	Phishing	Phishing	Phishing	Citibank	Misspelled Website	www.citybank.co.uk
6	Work	Work	Work	N/A	N/A	N/A
7	Legitimate	Legitimate	Legitimate	eBay	N/A	www.ebay.com
8	Work	Anti-Phishing email	Intervention	N/A	N/A	For the Intervention website, misspelled website was used (www.paypa1.com)
9	Legitimate	Legitimate	Legitimate	Amazon	N/A	www.amazon.co.uk
10	Phishing	Phishing	Phishing	Lloyds	Misspelled Website	www.online.ll0ydstsb.co.uk
11	Friend	Friend	Friend	N/A	N/A	N/A
12	Legitimate	Legitimate	Legitimate	Barclays	N/A	www.barclays.co.uk
13	Friend	Friend	Friend	N/A	N/A	N/A
14	Phishing	Phishing	Phishing	eBay	Different Domain	www.ebay-security.com

Table 20: The emails and websites order for each group in the first experiment

There were different URL syntax tricks (i.e. Phishing clues) used in the experiment. They formed the URLs for the Phishing websites. They were as follows:

- URLs with a different domain from a well-known domain,
- URLs with misspelled known websites and
- URLs with large host names that contained a part of a well-known web addresses.

8.6.3. The Second Experiment (Retention Experiment)

About 16 days after the first experiment, users of the New Approach and Old Approach groups in the first experiment were asked to perform a follow up experiment. This was the Second experiment. The target number of participants was as many as could come back. However, participants with low technical ability (LTA) and high technical ability (HTA) were needed within each group in order to ensure equal chances. Therefore, 12 participants were available and participated in the second experiment. Each group (Old Approach and New Approach) had 6 participants, divided into two subgroups, High and Low technical ability. Each subgroup had 3 participants.

This experiment was similar to the first one. The participants undertook email and web role-play protocol. Each participant played the role of an imaginary person named "Dave Smith". They had the same scenarios. However, there were differences. There were no treatments given to the participants. They had different emails and websites. Each participant was shown 14 email messages. Six messages were legitimate email messages, with no embedded links, that Dave Smith received from colleagues at his company and from friends. The other 8 email messages (implies 8 related websites) were divided into 4 simulated legitimate emails from organizations with which Dave Smith had an account and 4 Phishing emails.

The order of the emails and websites was predefined and identical for both groups and they are shown in Table 21. Re-using the emails and websites used in the first experiment might have allowed the participants to use their memory rather than their Phishing knowledge to react to the websites; consequently, the emails and websites in the second experiment were different from the ones used in the first experiment. However, the emails and websites in the second experiment used the same order followed in the first experiment. The websites also used the same URL tricks (i.e. Phishing clues) used in the first experiments and in the same order (See Tables 20 and 21).

#	eMails and Websites Order	Website	Tricks	URL
1	Phishing	Argos	Large Host names	www.argos.co.uk.myshop.com
2	Friend	N/A	N/A	N/A
3	Work	N/A	N/A	N/A
4	Legitimate	Abbey Bank	N/A	www.abbeynational.co.uk
5	Phishing	Capital One	Misspelled Website	www.capital0ne0nline.co.uk
6	Work	N/A	N/A	N/A
7	Legitimate	Comet	N/A	www.comet.co.uk
8	Work	N/A	N/A	N/A
9	Legitimate	Argos	N/A	www.argos.co.uk
10	Phishing	Co-operative Bank	Misspelled Website	www.co-operattivebank.co.uk
11	Friend	N/A	N/A	N/A
12	Legitimate	Egg Bank	N/A	www.egg.com
13	Friend	N/A	N/A	N/A
14	Phishing	Phishing	Different Domain	www.comet-online.co.uk

Table 21: The emails and websites order for each group in the second and third experiment

8.6.4. The Third Experiment (Phishing Awareness Experiment)

This experiment was exactly the same as the second one. The participants undertook email and web role-play protocol. Each participant again played the role of an imaginary person named "Dave Smith". They had the same scenarios. There were no treatments given to the participants. They had the same emails and websites with the same order as shown in Table 21. They also had the same URL tricks.

There were 12 participants. They were divided into two subgroups, High and Low Technical Ability. Each subgroup had 6 participants. All participants were 'Phishing Aware' people because there is need to have participants with knowledge about Phishing in order to compare their CDRs, FPRs and FNRs with Phishing Unaware people. These people were the Control group participants who took part in the first experiment.

8.7. Summary

This chapter has discussed the evaluation experiments. The chapter started with the hypotheses and their themes. Then, it presented the recruitment of the experiments' participants and their demographic information. This included the design of pre-study survey and its goals. After that, the chapter showed the effectiveness ratios, CDR, FPR and FNR, used in evaluating the hypotheses. Then, in order to have an overview of what should be simulated in the experiments, the comparisons between real Phishing attacks and Phishing experiments were considered. Finally, the methodology of the experiments was then discussed.

9. Evaluation

9.1. Introduction

This chapter discusses the evaluation of the research hypotheses presented earlier in Chapter 8. The hypotheses are classified in four research themes. They are evaluating the New Approach, the effect of high and low technical abilities on Phishing prevention, the effect of Phishing awareness and Phishing unawareness on Phishing detection and anti-Phishing knowledge retention. Each research theme has its own research hypotheses. Achieving the conclusion of each theme is accomplished by statistically analyzing its hypotheses.

9.2. Analysis

The statistical methods used were Kruskal-Wallis, Mann-Whitney and Wilcoxon Signed Ranks tests because the data is not normally distributed. The Confidence Interval was 95%.

9.2.1. Evaluating the New Approach

The theme of evaluating the New Approach has three aspects. Each aspect is discussed separately. They are as follows:

9.2.1.1. Aspect: Assessing Users without Treatments

Hypothesis (1.1):

- **Null Hypothesis (1.1):** Before using any treatment in the three groups (Control, Old Approach and New Approach groups), there are no differences between the correct decisions rates (CDRs) between all groups.
- **Statistical analysis method used:** Kruskal-Wallis test and Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Group	Level of Technical Ability	Mean	Std. Deviation	N
Control Group	Low	.5417	.10206	6
	High	.5000	.00000	6
	Total	**.5208**	.07217	**12**
Old Approach Group	Low	.5000	.00000	6
	High	.5000	.00000	6
	Total	**.5000**	.00000	**12**
New Approach Group	Low	.5417	.10206	6
	High	.5000	.00000	6
	Total	**.5208**	.07217	**12**
Total	Low	.5278	.08085	18
	High	.5000	.00000	18
	Total	.5139	.05808	36

Table 22: Descriptive statistics for CDRs' comparisons related to hypothesis 1.1

The correct decisions' rates (CDRs) for the New Approach, Old Approach and Control groups are based on the decisions taken by the groups' subjects before having their treatments. As Table 22 shows, the CDRs for the New Approach, Old Approach, and Control groups are nearly the same. They are .52, .50 and .52 respectively.

There are no significant differences between the rates of the three groups. The statistical difference between all groups before having treatments is p=.464. In order to see which group differs from another, Mann-Whitney tests were carried out to follow up the different results. A *Bonferroni* correction was applied so that the critical level of significance becomes 0.0167. Therefore, the statistical difference between the New Approach group and the Old Approach group is p=.522 and the difference between the

New Approach group and the Control group is p=1.000. The statistical difference also between the Old Approach group and the Control group is p=.522. Thus, the null hypothesis 1.1 is accepted. As a result of this, it is clear that the CDRs for all the three groups were the same before having any treatments.

Hypothesis (1.2):

- **Null Hypothesis (1.2):** Before using any treatment in the three groups (Control, Old Approach and New Approach groups), there are no differences between the false positive rates (FPRs) between all groups.
- **Statistical analysis method used:** Kruskal-Wallis test and Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Group	Level of Technical Ability	Mean	Std. Deviation	N
Control Group	Low	.0000	.00000	6
	High	.0833	.20412	6
	Total	**.0417**	.14434	**12**
Old Approach Group	Low	.0833	.20412	6
	High	.1667	.25820	6
	Total	**.1250**	.22613	**12**
New Approach Group	Low	.0833	.20412	6
	High	.0000	.00000	6
	Total	**.0417**	.14434	**12**
Total	Low	.0556	.16169	**18**
	High	.0833	.19174	**18**
	Total	.0694	.17537	36

Table 23: Descriptive statistics for FPRs' comparisons related to hypothesis 1.2

The false positive rates (FPRs) for the New Approach, Old Approach and Control groups are based on the decisions taken by the groups' subjects before having their treatments. As Table 23 demonstrates, the FPRs for the New Approach and Control groups are lower than the Old Approach group's rate. The rate for both is approximately 0.04 whereas the FPR for the Old Approach group is about 0.13.

There are no significant differences between the three rates for the three groups. This is because the statistical difference between them is p=.316. In order to see which group differs from another, Mann-Whitney tests were carried out to follow up the different results. Thus, the difference between the New Approach group and the Control group is p=1.000. The difference between the New Approach and Old Approach groups is p=.317 and the difference between Old Approach group and Control group is also p=.317 as well. As a result of this analysis, the null hypothesis 1.2 is accepted.

Hypothesis (1.3):

- **Null Hypothesis (1.3):** Before using any treatment in the three groups (Control, Old Approach and New Approach groups), there are no differences between the false negative rates (FNRs) between all groups.
- **Statistical analysis method used:** Kruskal-Wallis test and Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Group	Level of Technical Ability	Mean	Std. Deviation	N
Control Group	Low	.9167	.20412	6
	High	.9167	.20412	6
	Total	**.9167**	.19462	**12**
Old Approach Group	Low	.9167	.20412	6
	High	.8333	.25820	6
	Total	**.8750**	.22613	**12**
New Approach Group	Low	.8333	.25820	6
	High	1.0000	.00000	6
	Total	**.9167**	.19462	**12**
Total	Low	.8889	.21390	**18**
	High	.9167	.19174	**18**
	Total	.9028	.20069	36

Table 24: Descriptive statistics for FNRs' comparisons related to hypothesis 1.3

The false negative rates (FNRs) for the New Approach, Old Approach and Control groups are based on the decisions taken by the groups' subjects before having their treatments. Table 24 shows the FNRs for all groups. The FNR for the Old Approach

group is about 0.88. The FNRs for the New Approach and Control groups are higher than the Old Approach group's rate. The rate for both is approximately 0.92.

All the three rates are high. There are no significant differences between them (p=1.000). In order to see which group differs from another, Mann-Whitney tests were carried out to follow up the different results. Thus, it was found that the statistical difference between each group and the other groups is not significant; p=1.000. Therefore, the null hypothesis 1.3 is accepted.

Aspect Discussion:

Hypothesis	Result
Hypothesis 1.1	Accepted
Hypothesis 1.2	Accepted
Hypothesis 1.3	Accepted

Table 25: A summary of hypotheses' analysis results in assessing users before treatments

It is clear from Table 25 that users in the three groups were nearly equal with regards to their decisions about legitimate and Phishing websites before using any treatment. There is no difference in their correct decisions' rate (CDR). There are also no differences in their false positive and false negative rates.

9.2.1.2. Aspect: Assessing the New Approach in Comparison with the Old Approach

Hypothesis (1.4):

- **Null Hypothesis (1.4):** There is no difference between the correct decisions rate (CDR) for the New Approach group after using the New Approach and the CDRs for the Old Approach group and the Control group after using their treatments.
- **Statistical analysis method used:** Kruskal-Wallis test and Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Reject.
- **Discussion:**

Group	Level of Technical Ability	Mean	Std. Deviation	N
Control Group	Low	.4583	.10206	6
	High	.5833	.12910	6
	Total	**.5208**	.12873	**12**
Old Approach Group	Low	.5000	.15811	6
	High	.5417	.18819	6
	Total	**.5208**	.16714	**12**
New Approach Group	Low	.8333	.12910	6
	High	.7083	.18819	6
	Total	**.7708**	.16714	**12**
Total	Low	.5972	.21246	**18**
	High	.6111	.17620	**18**
	Total	**.6042**	.19249	36

Table 26: Descriptive statistics for CDRs' comparisons related to hypothesis 1.4

The correct decisions' rates (CDRs) for the New Approach, Old Approach and Control groups are based on the decisions taken by the groups' subjects after having their treatments. As Table 26 shows, the CDR for the New Approach group is higher than the Old Approach group's rate. The rate is approximately 0.77 out of 1 for the New Approach group whereas it is about 0.52 for the Old Approach group and the Control group.

There is a significant difference between the groups (p=.001). In order to see which group differs from another, Mann-Whitney tests were carried out to follow up the different results. The significant difference between the New Approach and Old Approach groups is p=.002. Additionally, there is a significant difference between the rates of the New Approach group and the Control group (p=.001). This means that there was a significant positive effect of using the New Approach in comparison with the Old Approach and having no approaches (Control group). By using the New Approach, the subjects were highly protected from making mistakes in judging legitimate and Phishing websites. Therefore, the null hypothesis is rejected. Regarding the comparison between the Old Approach and the Control groups, there is no significant statistical difference between the two groups which is p=.500.

As a result of this, it is shown that the CDR for the New Approach group after using the New Approach was better than the CDRs for the Old Approach group and the

Control group after using their treatments. The Old Approach also had no significant effect on judging both the legitimate and Phishing websites.

Hypothesis (1.5):

- **Null Hypothesis (1.5):** There is no difference between the false positive rate (FPR) for the New Approach group after using the New Approach and the FPRs for the Old Approach group and the Control group after using their treatments.
- **Statistical analysis method used:** Kruskal-Wallis test and Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Group	Level of Technical Ability	Mean	Std. Deviation	N
Control Group	Low	.1667	.25820	6
	High	.0833	.20412	6
	Total	**.1250**	.22613	**12**
Old Approach Group	Low	.1667	.25820	6
	High	.3333	.25820	6
	Total	**.2500**	.26112	**12**
New Approach Group	Low	.0833	.20412	6
	High	.1667	.25820	6
	Total	**.1250**	.22613	**12**
Total	Low	.1389	.23044	18
	High	.1944	.25082	18
	Total	.1667	.23905	36

Table 27: Descriptive statistics for FPRs' comparisons related to hypothesis 1.5

The false positive rates (FPRs) for the three groups (the New Approach group, Old Approach group and Control group) were counted after the subjects had their different treatments. Because the false positive was a wrong decision, the lower the FPR result the better was the judgment by subjects about the legitimate websites. As Table 27 shows, the FPRs for the New Approach and Control groups are lower than the Old Approach group's rate. The rate is about 0.13 out of 1 in both the New Approach and Control groups, whereas it is 0.25 in the Old Approach group.

There were no significant differences between the three rates for the three groups (p=.490). In order to see which group differs from another, Mann-Whitney tests were carried out to follow up the different results. The statistical difference between the New Approach group and the Control group is p=.680. However, the difference also between the New Approach group and the Old Approach group and the difference between the Old Approach group and the Control group are equal (p=.200). Therefore, there were no significant effects of using the New Approach and the Old Approach on properly judging the legitimate websites by the subjects because they did not significantly differ from the Control group who did not have any approach. As a result of this analysis, the null hypothesis 1.5 is accepted. The New and the Old Approaches had no effect on helping the subjects to judge the legitimate websites.

The reason why there are no differences in false positive rates (FPRs) across the three groups is because nearly all the participants responded to the legitimate websites regardless of whether or not they were cautious of Phishing websites. This is due to the fact that if they knew about Phishing, they most probably responded to the website because they considered it to be 'legitimate' and if they did not know about Phishing, they responded because they believed that it was a legitimate website.

Hypothesis (1.6):

- **Null Hypothesis (1.6):** There is no difference between the false negative rate (FNR) for the New Approach group after using the New Approach and the FNRs for the Old Approach group and the Control group after using their treatments.
- **Statistical analysis method used:** Kruskal-Wallis test and Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Reject.
- **Discussion:**

Group	Level of Technical Ability	Mean	Std. Deviation	N
Control Group	Low	.9167	.20412	6
	High	.7500	.41833	6
	Total	**.8333**	.32567	**12**
Old Approach Group	Low	.8333	.25820	6
	High	.5833	.20412	6
	Total	**.7083**	.25746	**12**
New Approach Group	Low	.3333	.25820	6
	High	.4167	.37639	6
	Total	**.3750**	.31079	**12**
Total	Low	.6944	.34890	**18**
	High	.5833	.35355	**18**
	Total	.6389	.35074	36

Table 28: Descriptive statistics for FNRs' comparisons related to hypothesis 1.6

The false negative rates (FNRs) for the three groups are based on the subjects' decisions on Phishing websites after they had their different treatments. Because the false negative was a wrong decision, the lower the FNR result the better was the subjects' judgment on the Phishing websites. Table 28 presents the FNRs for the three groups. The New Approach group has the lowest rate. Their rate is 0.38 out of 1. In contrast the subjects in Control group have the highest rate, which is approximately 0.83. The Old Approach group's rate is in between the New Approach group and the Control group at about 0.71.

The difference between the groups is significant (p=.002). In order to see which group differs from another, Mann-Whitney tests were carried out to follow up the different results. There is a significant difference between the New Approach and the Old Approach groups (p=.012). In addition, there is a significant effect on the New Approach group in comparison with the Control group (p=.001). Therefore, the null hypothesis 1.6 is rejected. However, although the rate of the false negative decisions for the Old Approach subjects is better than for the ones for the Control group, there is no statistically significant difference between them. The difference is p= .107.

It is clear that there are significant effects of using the New Approach group in comparison with the Old Approach group and no treatments (the Control group) in

helping subjects to judge the Phishing websites properly and thus to enable them to detect the Phishing attacks. Thus, the false negative rate (FNR) for the New Approach group after using the New Approach is better (less) than the FNR for the Old Approach group and the Control group.

Aspect Discussion:

Hypothesis	Result
Hypothesis 1.4	Rejected
Hypothesis 1.5	Accepted
Hypothesis 1.6	Rejected

Table 29: A summary of hypotheses' analysis results in assessing the New Approach in comparison with the Old Approach

Table 29 shows a summary of the results of the three hypotheses discussed so far for assessing the New Approach in comparison with the Old Approach. It has been found that there is a significant positive effect of using the New Approach in comparison with the Old Approach. The users of the New Approach were highly protected from making mistakes in judging legitimate and Phishing websites. In detail, the New and the Old Approaches were equal on helping the subjects to judge the legitimate websites. However, there was a significant effect of using the New Approach in comparison with the Old Approach in helping subjects to judge the Phishing websites properly and this enabled them to detect the Phishing attacks. This means that the New Approach worked better than the Old Approach. This is demonstrated by the comparisons of the correct decisions' rate (CDR) of the two approaches discussed earlier.

Regarding the comparison between the Old Approach and the Control groups, the Old Approach had no significant effect on their ability to judge both the legitimate and Phishing websites.

9.2.1.3. Aspect: Assessing Users Before and After Using the Treatments

Hypothesis (1.7):

- **Null Hypothesis (1.7):** In the Control group, there is no difference between the correct decisions rates (CDRs) after having the treatment (in this instance ordinary email from work essentially not treatment) and the CDRs before having the treatment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Results:** Accept.
- **Discussion:**

Response Variable	Mean	N	Std. Deviation
CDR after treatment	.5208	12	.12873
CDR before treatment	.5208	12	.07217

Table 30: Descriptive statistics for CDRs' comparisons related to hypothesis 1.7

The correct decisions' rates (CDRs) for the Control group compared in the hypothesis 1.7 are based on the decisions taken by the group's subjects before and after having their treatment which was an ordinary email from work. Table 30 shows the two rates. The CDR for the Control group before the treatment is exactly the same as the CDR for the same group after taking the treatment. The CDRs before and after the treatment are 0.52. There is also no statistical difference between the two CDRs. This is because the statistical difference between them is p=1.000. As a result of this analysis, the null hypothesis 1.7 is accepted. Due to the fact that the Control group did not have an actual treatment, the subjects reacted to both the Phishing and legitimate websites before and after the treatment at nearly the same average rate in the experiments.

Hypothesis (1.8):

- **Null Hypothesis (1.8):** In the Control group, there is no difference between the false positive rate (FPR) after having the 'treatment' (in this instance ordinary email from work essentially not treatment) and the FPR before having the treatment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.

- **Result:** Accept.
- **Discussion:**

Response Variable	Mean	N	Std. Deviation
FPR after treatment	.1250	12	.22613
FPR before treatment	.0417	12	.14434

Table 31: Descriptive statistics for FPRs' comparisons related to hypothesis 1.8

The false positive rates (FPRs) for the Control group compared in hypothesis 1.8 are based on the decisions taken by the group's subjects before and after having their treatment, which was an ordinary email from work. As Table 31 shows, the false positive rate FPR for the Control group before the treatment is lower than the FPR for the same group after taking the treatment. The FPR before the treatment is 0.04, whereas it is approximately 0.13 after the treatment. However, there is no statistical difference between the two FPRs. This is because the statistical difference between them is p=.500. As a result of this analysis, the null hypothesis 1.8 is accepted. Due to the fact that the Control group did not have an actual treatment, the subjects reacted to the legitimate websites before and after the treatment at rates close to each other in the experiments.

Hypothesis (1.9):

- **Null Hypothesis (1.9):** In the Control group, there is no difference between the false negative rate (FNR) after having the treatment (in this instance ordinary email from work essentially not treatment) and the FNR before having the treatment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Results:** Accept.
- **Discussion:**

Response Variable	Mean	N	Std. Deviation
FNR after treatment	.8333	12	.32567
FNR before treatment	.9167	12	.19462

Table 32: Descriptive statistics for FNRs' comparisons related to hypothesis 1.9

The false negative rates (FNRs) for the Control group compared in hypothesis 1.9 are based on the decisions taken by the group's subjects before and after having their treatment, which in this instance was an ordinary email from work, essentially not treatment. Table 32 shows the two rates. The FNR for the Control group before the treatment is higher than the FNR for the same group after taking the treatment. The FNR before the treatment is 0.92 but it is 0.83 after the treatment. However, there is no statistical difference between the two FNRs. This is because the statistical difference between them is p=.500. As a result of this analysis, the null hypothesis 1.9 is accepted. Due to the fact that the Control group did not have an actual treatment, the subjects reacted to the Phishing websites before and after the treatment at nearly similar average rate in the experiments.

Hypothesis (1.10):

- **Null Hypothesis (1.10):** In the Old Approach group, there is no difference between the correct decisions rate (CDR) after having the treatment (i.e. an anti-Phishing training email) and the CDR before having the treatment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Response Variable	Mean	N	Std. Deviation
The rate after treatment	.5417	12	.20871
The rate before treatment	.4792	12	.07217

Table 33: Descriptive statistics for CDRs' comparisons related to hypothesis 1.10

The correct decisions rates (CDRs) for the Old Approach group shown in this hypothesis are based on the decisions taken by the group's subjects before and after having their treatment which was an anti-Phishing training email. As presented in Table 33, the CDR for the Old Approach group before the treatment is 0.48 whereas it is 0.54 after the treatment. There is no statistical difference between the two CDRs; p=.250. As a result of this analysis, the null hypothesis 1.10 is accepted.

Hypothesis (1.11):

- **Null Hypothesis (1.11):** In the Old Approach group, there is no difference between the false positive rate (FPR) after having the treatment (i.e. an anti-Phishing training email) and the FPR before having the treatment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Response Variable	Mean	N	Std. Deviation
FPR after treatment	.2500	12	.26112
FPR before treatment	.1667	12	.24618

Table 34: Descriptive statistics for FPRs' comparisons related to hypothesis 1.11

The false positive rates (FPRs) for the Old Approach group shown in this hypothesis are based on the decisions taken by the group's subjects before and after having their treatment, which was an anti-Phishing training email. As shown in Table 34, the FPR for the Old Approach group before the treatment is lower (0.17) than the FPR after the treatment (0.25). This is deterioration but there is no statistical difference between the two FPRs; p=.344. As a result of this analysis, the null hypothesis 1.11 is accepted.

The deterioration in the FPRs is because, before having the treatment, nearly all the subjects responded to the legitimate websites because they believed that they were legitimate websites. After having the treatment, they became worried about the legitimacy of websites. Then, if they had a legitimate website to respond to, they preferred not to respond in order to be on the safe side. Thus, the FPR after the treatment is higher (worse) than the FPR before the treatment.

Hypothesis (1.12):

- **Null Hypothesis (1.12):** In the Old Approach group, there is no difference between the false negative rate (FNR) after having the treatment (i.e. an anti-Phishing training email) and the FNR before having the treatment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.

- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Response Variable	Mean	N	Std. Deviation
FNR after treatment	.6667	12	.32567
FNR before treatment	.8750	12	.22613

Table 35: Descriptive statistics for FNRs' comparisons related to hypothesis 1.12

The false negative rates (FNRs) for the Old Approach group shown in this hypothesis are based on the decisions taken by the group's subjects before and after having their treatment, which was an anti-Phishing training email. As shown in Table 35, the FNR for the Old Approach group before the treatment is higher (0.88) than its FNR after the treatment (0.67).

The statistical difference between the FNR after the treatment and the FNR before the treatment is not significant; p=.063. As a result of this analysis, the null hypothesis 1.12 is accepted. Therefore, in the Old Approach group, the FNR after having the treatment (i.e. an anti-Phishing training email) is better (less) than the FNR before having the treatment but there is no significant effect.

Hypothesis (1.13):

- **Null Hypothesis (1.13):** In the New Approach group, there is no difference between the correct decisions rate (CDR) after having the treatment (i.e. anti-Phishing intervention) and the CDR before having the treatment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Reject.
- **Discussion:**

Response Variable	Mean	N	Std. Deviation
The rate after treatment	.7708	12	.16714
The rate before treatment	.5208	12	.07217

Table 36: Descriptive statistics for CDRs' comparisons related to hypothesis 1.13

The correct decisions' rates (CDRs) for the New Approach group shown in this hypothesis are based on the decisions taken by the group's subjects before and after having their treatment, which was an anti-Phishing intervention. As presented in Table 36, the CDR for the New Approach group before the treatment is 0.52 whereas it is 0.77 after the treatment. There is a significant statistical difference between the two CDRs; p=.002 (for the CDR after the treatment). As a result of this analysis, the null hypothesis 1.13 is rejected.

Thus, the New Approach had a significant positive effect on the subjects' decisions. Therefore, in the New Approach group, the CDR after having the approach (i.e. anti-Phishing intervention) is significantly better (more) than the CDR before having the approach.

Hypothesis (1.14):

- **Null Hypothesis (1.14):** In the New Approach group, there is no difference between the false positive rate (FPR) after having the treatment (i.e. anti-Phishing intervention) and the FPR before having the treatment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Response Variable	Mean	N	Std. Deviation
FPR after treatment	.1250	12	.22613
FPR before treatment	.0417	12	.14434

Table 37: Descriptive statistics for FPRs' comparisons related to hypothesis 1.14

The false positive rates (FPRs) for the New Approach group shown in this hypothesis are based on the decisions taken by the group's subjects before and after having their treatment, which was an anti-Phishing intervention. As presented in Table 37, the FPR for the New Approach group before the treatment is 0.04 whereas it is about 0.13 after the treatment. This is deterioration because the FPR before the treatment is better than the FPR after the treatment. However, there is no statistical difference

between the two FPRs; p=.313. As a result of this analysis, the null hypothesis 1.14 is accepted.

The deterioration in the FPRs is because of the same reason that was given and discussed in hypothesis 1.11.

Hypothesis (1.15):

- **Null Hypothesis (1.15):** In the New Approach group, there is no difference between the false negative rate (FNR) after having the treatment (i.e. anti-Phishing intervention) and the FNR before having the treatment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Reject.
- **Discussion:**

Response Variable	Mean	N	Std. Deviation
FNR after treatment	.3333	12	.32567
FNR before treatment	.9167	12	.19462

Table 38: Descriptive statistics for FNRs' comparisons related to hypothesis 1.15

The false negative rates (FNRs) for the New Approach group shown in this hypothesis are based on the decisions taken by the group's subjects before and after having their treatment, which was an anti-Phishing intervention. As shown in Table 38, the FNR for the New Approach group before the treatment is higher (0.92) than the FNR after the treatment (0.33).

There is a statistical difference between the FNR after the treatment and the FNR before the treatment; p=.001. As a result of this analysis, the null hypothesis 1.15 is rejected. Thus, the New Approach had a significant positive effect on the subjects' decisions on judging the Phishing websites. Therefore, in the New Approach group, the FNR after having the approach (i.e. anti-Phishing intervention) is significantly better (less) than the FNR before having the approach.

Aspect Discussion:

Hypothesis	Result
Hypothesis 1.7	Accepted
Hypothesis 1.8	Accepted
Hypothesis 1.9	Accepted
Hypothesis 1.10	Accepted
Hypothesis 1.11	Accepted
Hypothesis 1.12	Accepted
Hypothesis 1.13	Rejected
Hypothesis 1.14	Accepted
Hypothesis 1.15	Rejected

Table 39: A summary of hypotheses' analysis results in assessing users before and after using the treatments

Table 39 shows an overview of the hypotheses results in assessing users before and after having the treatments. In the Control group, there are no differences between the correct decisions rates (CDRs), false positive rates (FPRs) and false negative rates (FNRs) before and after having the treatment (in this instance, an ordinary email from work).

In the Old Approach group, there is no difference between the FPRs and the FNRs. The CDRs also were not significantly different. The CDR is more important because it is indicative of users' decisions on the total of both legitimate and Phishing websites. Therefore, the Old Approach had no significant effect on users' decisions on the legitimacy of websites.

Regarding the New Approach group, there is no significant difference between the FPRs before and after having the treatment (i.e. an anti-Phishing intervention). However, there is a significant difference between the FNRs (for the rate after the treatment). More importantly, the CDRs are significantly different. The CDR after having the treatment is better than the rate before the treatment. Therefore, the New Approach had a significant effect on users' decisions on the legitimacy of websites.

9.2.1.4. Theme Summary

To sum up, users of the three groups were nearly equal with regards to their decisions about legitimate and Phishing websites before using any treatment. After using the treatments, there is a significant positive effect of using the New Approach in comparison with the Old Approach. The New Approach is better than the Old Approach in helping users properly judging the legitimacy of websites.

9.2.2. Effect of High and Low Technical Abilities on Phishing Detection

The theme of effect of high and low technical ability on Phishing detection has three different aspects. Each aspect is discussed individually. They are as follows:

9.2.2.1. Aspect: Assessing the Effect of the Technical Ability Level among Phishing Unaware Users

Hypothesis (2.1):

- **Null Hypothesis (2.1):** In the Control group, there is no difference between the correct decisions' rate (CDR) for high technical ability (HTA) people and the CDR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.5000	.07906
High	6	.5417	.06455

Table 40: Descriptive statistics for CDRs' comparisons related to hypothesis 2.1

The correct decisions rates (CDRs) for the Control group shown in this hypothesis are based on the decisions taken by the group's subjects with both high technical ability (HTA) and low technical ability (LTA). As presented in Table 40, the CDRs for the LTA subjects is 0.50 whereas it is 0.54 for the HTA subjects. The two rates are nearly the same. There is no significant difference between the two CDRs; p=.636. As a result of this analysis, the null hypothesis 2.1 is accepted.

Hypothesis (2.2):

- **Null Hypothesis (2.2):** In the Control group, there is no difference between the false positive rate (FPR) for high technical ability (HTA) people and the FPR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.0833	.12910
High	6	.0833	.20412

Table 41: Descriptive statistics for FPRs' comparisons related to hypothesis 2.2

The false positive rates (FPRs) for the Control group shown in this hypothesis are based on the decisions taken by the group's subjects with both high technical ability (HTA) and low technical ability (LTA). As presented in Table 41, the FPR for the LTA subjects and the FPR for the HTA subjects are exactly the same. They are both 0.08. There is no difference between the two FPRs; p=1.000. As a result of this analysis, the null hypothesis 2.2 is accepted.

Hypothesis (2.3):

- **Null Hypothesis (2.3):** In the Control group, there is no difference between the false negative rate (FNR) for high technical ability (HTA) people and the FNR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.9167	.20412
High	6	.8333	.30277

Table 42: Descriptive statistics for FNRs' comparisons related to hypothesis 2.3

The false negative rates (FNRs) for the Control group shown in this hypothesis are based on the decisions taken by the group's subjects with both high technical ability (HTA) and low technical ability (LTA). As presented in Table 42, the FNR for the LTA subjects is 0.92 whereas it is 0.83 for the HTA subjects. The two rates are high. There is no significant difference between the two FNRs; p=.727. As a result of this analysis, the null hypothesis 2.3 is accepted.

Hypothesis (2.4):

- **Null Hypothesis (2.4):** In the Old Approach group and before having the treatment (i.e. anti-Phishing training email), there is no difference between the correct decisions rate (CDR) for high technical ability (HTA) people and the CDR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.5000	.00000
High	6	.4583	.10206

Table 43: Descriptive statistics for FNRs' comparisons related to hypothesis 2.4

The correct decisions rates (CDRs) for the Old Approach group shown in this hypothesis are based on the decisions taken by the group's subjects with both high technical ability (HTA) and low technical ability (LTA) before having the treatment (i.e. anti-Phishing training email). As presented in Table 43, the CDR for the LTA subjects is about 0.50 whereas it is 0.46 for the HTA subjects. There is no statistical difference between the two CDRs; p=1.000. As a result of this analysis, the null hypothesis 2.4 is accepted.

Hypothesis (2.5):

- **Null Hypothesis (2.5):** In the Old Approach group and before having the treatment (i.e. anti-Phishing training email), there is no difference between the false positive rate (FPR) for high technical ability (HTA) people and the FPR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.0833	.20412
High	6	.2500	.27386

Table 44: Descriptive statistics for FPRs' comparisons related to hypothesis 2.5

The false positive rates (FPRs) for the Old Approach group shown in this hypothesis are based on the decisions taken by the group's subjects with both high technical ability (HTA) and low technical ability (LTA) before having the treatment (i.e. anti-Phishing training email). As Table 44 shows, the FPR for the LTA subjects is about 0.08 whereas

it is 0.25 for the HTA subjects. There is no statistical difference between the two FPRs; p=.545. As a result of this analysis, the null hypothesis 2.5 is accepted.

Hypothesis (2.6):

- **Null Hypothesis (2.6):** In the Old Approach group and before having the treatment (i.e. anti-Phishing training email), there is no difference between the false negative rate (FNR) for high technical ability (HTA) people and the FNR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.9167	.20412
High	6	.8333	.25820

Table 45: Descriptive statistics for FNRs' comparisons related to hypothesis 2.6

The false negative rates (FNRs) for the Old Approach group shown in this hypothesis are based on the decisions taken by the group's subjects with both high technical ability (HTA) and low technical ability (LTA) before having the treatment (i.e. anti-Phishing training email). As presented in Table 45, the FNR for the LTA subjects is about 0.92, whereas it is 0.83 for the HTA subjects. There is no statistical difference between the two FNRs; p=1.000. As a result of this analysis, the null hypothesis 2.6 is accepted.

Hypothesis (2.7):

- **Null Hypothesis (2.7):** In the New Approach group and before having the treatment (i.e. anti-Phishing intervention), there is no difference between the correct decisions' rate (CDR) for high technical ability (HTA) people and the CDR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.

- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.5417	.10206
High	6	.5000	.00000

Table 46: Descriptive statistics for CDRs' comparisons related to hypothesis 2.7

The correct decisions rates (CDRs) for the New Approach group shown in this hypothesis are based on the decisions taken by the group's subjects with both high technical ability (HTA) and low technical ability (LTA) before having the treatment (i.e. anti-Phishing intervention). As shown in Table 46, the CDR for the LTA subjects is about 0.54, whereas it is 0.50 for the HTA subjects. There is no statistical difference between the two CDRs; p=1.000. As a result of this analysis, the null hypothesis 2.7 is accepted.

Hypothesis (2.8):

- **Null Hypothesis (2.8):** In the New Approach group and before having the treatment (i.e. anti-Phishing intervention), there is no difference between the false positive rate (FPR) for high technical ability (HTA) people and the FPR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.0833	.20412
High	6	.0000	.00000

Table 47: Descriptive statistics for FPRs' comparisons related to hypothesis 2.8

The false positive rates (FPRs) for the New Approach group shown in this hypothesis are based on the decisions taken by the group's subjects with both high technical ability (HTA) and low technical ability (LTA) before having the treatment (i.e. anti-Phishing intervention). As Table 47 shows, the FPR for the LTA subjects is about 0.08, whereas it is 0 for the HTA subjects. They are both low rates. There is no significant statistical difference between the two groups; p=1.000. As a result of this analysis, the null hypothesis 2.8 is accepted.

Hypothesis (2.9):

- **Null Hypothesis (2.9):** In the New Approach group and before having the treatment (i.e. anti-Phishing intervention), there is no difference between the false negative rate (FNR) for high technical ability (HTA) people and the FNR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.8333	.25820
High	6	1.0000	.00000

Table 48: Descriptive statistics for FNRs' comparisons related to hypothesis 2.9

The false negative rates (FNRs) for the New Approach group shown in this hypothesis are based on the decisions taken by the group's subjects with both high technical ability (HTA) and low technical ability (LTA) before having the treatment (i.e. anti-Phishing intervention). As presented in Table 48, the FNR for the LTA subjects is about 0.83, whereas it is 1.00 for the HTA subjects. They are both high rates. There is no significant statistical difference between the two FNRs; p=.455. As a result of this analysis, the null hypothesis 2.9 is accepted.

Hypothesis (2.10):

- **Null Hypothesis (2.10):** With regard to technical ability levels and regardless of which group (Control, Old Approach or New Approach) they belonged to and before having the treatments, there is no difference between the correct decisions rate (CDR) for high technical ability (HTA) people and the CDR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	18	.5278	.08085
High	18	.4861	.05893

Table 49: Descriptive statistics for CDRs' comparisons related to hypothesis 2.10

This comparison does not consider the group to which the subjects belonged. The focus is on their level of technical ability. Therefore, the correct decisions rates (CDRs) shown in this hypothesis are based on the decisions taken by subjects with both high technical ability (HTA) and low technical ability (LTA) before having their treatments. Table 49 presents the two rates; one rate for LTA subjects and the other for the HTA subjects. The CDR for the LTA subjects is about 0.53, whereas it is 0.49 for the HTA subjects. They are nearly the same. There is no significant statistical difference between the two groups; p=.257. As a result of this analysis, the null hypothesis 2.10 is accepted.

Hypothesis (2.11):

- **Null Hypothesis (2.11):** With regard to technical ability levels and regardless of the group (Control, Old Approach or New Approach) to which they belonged and before having the treatments, there is no difference between the false positive rate (FPR) for high technical ability (HTA) people and the FPR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.

- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	18	.0556	.16169
High	18	.1111	.21390

Table 50: Descriptive statistics for FPRs' comparisons related to hypothesis 2.11

As in the previous discussion on hypothesis 2.10, this comparison does not consider the group to which the subjects belonged. The focus is on their level of technical ability. Therefore, the false positive rates (FPRs) shown in this hypothesis are based on the decisions taken by subjects with both high technical ability (HTA) and low technical ability (LTA) before having their treatments. Table 50 presents the two rates; one rate for LTA subjects and the other for the HTA subjects. The FPR for the LTA subjects is about 0.06, whereas it is 0.11 for the HTA subjects. They are both low rates. There is no significant statistical difference between the two FPRs; p=.658. As a result of this analysis, the null hypothesis 2.11 is accepted.

Hypothesis (2.12):

- **Null Hypothesis (2.12):** With regard to technical ability levels and regardless of the group (Control, Old Approach or New Approach) to which the subjects belonged and before having the treatments, there is no difference between the false negative rate (FNR) for high technical ability (HTA) people and the FNR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	18	.8889	.21390
High	18	.9167	.19174

Table 51: Descriptive statistics for FNRs' comparisons related to hypothesis 2.12

As for hypotheses 2.10 and 2.11, the groups in this comparison are not important. The focus is on the level of technical ability for all Phishing unaware subjects. Therefore, the false negative rates (FNRs) shown in this hypothesis are based on the decisions taken by subjects with both high technical ability (HTA) and low technical ability (LTA) before having their treatments. Table 51 presents the two rates; one rate for LTA subjects and the other is for the HTA subjects. The FNR for the LTA subjects is about 0.89, whereas it is 0.92 for the HTA subjects. They are both high rates. There is no significant statistical difference between the two FNRs; p=1.000. As a result of this analysis, the null hypothesis 2.12 is accepted.

Aspect Discussion:

Hypothesis	Result
Hypothesis 2.1	Accepted
Hypothesis 2.2	Accepted
Hypothesis 2.3	Accepted
Hypothesis 2.4	Accepted
Hypothesis 2.5	Accepted
Hypothesis 2.6	Accepted
Hypothesis 2.7	Accepted
Hypothesis 2.8	Accepted
Hypothesis 2.9	Accepted
Hypothesis 2.10	Accepted
Hypothesis 2.11	Accepted
Hypothesis 2.12	Accepted

Table 52: A summary of hypotheses' analysis results in assessing the effect of the technical ability level among Phishing Unaware users

As presented in Table 52, all hypotheses in assessing the effect of high and low technical abilities in Phishing Unaware users are accepted. The hypotheses' analysis evaluated users without having treatments in the three groups (Control, New Approach and Old Approach). The result is that there is no significant difference between the decisions rates of high and low technical ability users. Therefore, the level of the technical ability has no effect on Phishing detection or on recognizing legitimate websites among Phishing unaware people.

9.2.2.2. Aspect: Assessing the Effect of the Technical Ability Level among Phishing Aware Users

Hypothesis (2.13):

- **Null Hypothesis (2.13):** In the Phishing Aware people group, there is no difference between the correct decisions rate (CDR) for high technical ability (HTA) people and the CDR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.7292	.12290
High	6	.7083	.10206

Table 53: Descriptive statistics for CDRs' comparisons related to hypothesis 2.13

The correct decisions rates (CDRs) shown in this hypothesis are based on the decisions taken by subjects who are considered 'Phishing Aware' with both high technical ability (HTA) and low technical ability (LTA). Thus, the focus is on the Phishing Aware people group. As presented in Table 53, the CDR for the LTA subjects is about 0.73, whereas it is 0.71 for the HTA subjects. The two rates are nearly the same, which reflects the similarity between the two groups. Additionally, there is no significant statistical difference between the two groups; p=1.000. As a result of this analysis, the null hypothesis 2.13 is accepted.

Hypothesis (2.14):

- **Null Hypothesis (2.14):** In the Phishing Aware people group, there is no difference between the false positive rate (FPR) for high technical ability (HTA) people and the FPR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.

- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.1667	.20412
High	6	.2500	.15811

Table 54: Descriptive statistics for FPRs' comparisons related to hypothesis 2.14

The false positive rates (FPRs) shown in this hypothesis are based on the decisions taken by subjects who are considered 'Phishing Aware' with both high technical ability (HTA) and low technical ability (LTA). Thus, the focus is on the Phishing Aware people group. As shown in Table 54, the FPR for the LTA subjects is about 0.17, whereas it is 0.25 for the HTA subjects. There is also no significant statistical difference between the two groups; p=.494. As a result of this analysis, the null hypothesis 2.14 is accepted.

Hypothesis (2.15):

- **Null Hypothesis (2.15):** In the Phishing Aware people group, there is no difference between the false negative rate (FNR) for high technical ability (HTA) people and the FNR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	6	.3750	.34460
High	6	.3333	.25820

Table 55: Descriptive statistics for FNRs' comparisons related to hypothesis 2.15

The false negative rates (FNRs) shown in this hypothesis are based on the decisions taken by subjects who are considered 'Phishing Aware' with both high technical ability (HTA) and low technical ability (LTA). Thus, the focus is on the Phishing Aware people group. Table 55 presents the two rates. The FNR for the LTA subjects is about 0.38,

whereas it is 0.33 for the HTA subjects. They are approximately similar. There is no significant statistical difference between the two groups; p=1.000. As a result of this analysis, the null hypothesis 2.15 is accepted.

Aspect Discussion:

Hypothesis	Result
Hypothesis 2.13	Accepted
Hypothesis 2.14	Accepted
Hypothesis 2.15	Accepted

Table 56: A summary of hypotheses' analysis results in assessing the effect of the technical ability level among Phishing Aware users

Regarding Phishing Aware users, Table 56 shows a summary of the hypotheses analysis results. There is no difference between the high and low technical ability users in Phishing detection and prevention. There is also no difference between the two groups (high and low technical ability) in properly judging legitimate websites. Therefore, technical ability has no effect on the decisions of Phishing Aware users in Phishing detection and in recognizing legitimate websites.

9.2.2.3. Aspect: Assessing the Effect of the Technical Ability Level Regardless of the Phishing Knowledge (Phishing Aware and Unaware)

Hypothesis (2.16):

- **Null Hypothesis (2.16):** For all subjects with regard to their technical ability levels and regardless of their Phishing knowledge (Phishing Aware or Phishing Unaware) and before having the treatments, there is no difference between the correct decisions rate CDR for high technical ability (HTA) people and the CDR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	24	.5729	.11608
High	24	.5417	.12039

Table 57: Descriptive statistics for CDRs' comparisons related to hypothesis 2.16

Phishing knowledge in this comparison is not considered. The focus is on the level of technical ability for all subjects who participated in all the experiments. Therefore, the correct decisions rates (CDRs) shown in this hypothesis are based on the decisions taken by subjects with both high technical ability (HTA) and low technical ability (LTA) before having their treatments. Table 57 presents the two rates; one rate for LTA subjects and the other is for the HTA subjects. The CDR for the LTA subjects is about 0.57, whereas it is 0.54 for the HTA subjects. They are nearly the same. There is no significant statistical difference between the two CDRs; p=.541. As a result of this analysis, the null hypothesis 2.16 is accepted.

Hypothesis (2.17):

- **Null Hypothesis (2.17):** For all subjects with regard to their technical ability levels and regardless of their Phishing knowledge (Phishing Aware or Phishing Unaware) and before having the treatments, there is no difference between the false positive rate (FPR) for high technical ability (HTA) people and the FPR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	24	.0833	.19035
High	24	.1458	.23215

Table 58: Descriptive statistics for FPRs' comparisons related to hypothesis 2.17

As for the previous hypothesis 2.16, Phishing knowledge in this comparison is not considered. The focus is on the level of technical ability for all subjects who participated in all the experiments. Therefore, the false positive rates (FPRs) shown in this hypothesis are based on the decisions taken by subjects with both high technical ability (HTA) and low technical ability (LTA) before having their treatments. Table 58 presents the two rates; one rate for LTA subjects and the other is for the HTA subjects. The FPR for the LTA subjects is about 0.08, whereas it is 0.15 for the HTA subjects. There is no significant statistical difference between the two rates; p=.494. As a result of this analysis, the null hypothesis 2.17 is accepted.

Hypothesis (2.18):

- **Null Hypothesis (2.18):** For all subjects with regard to their technical ability levels and regardless of their Phishing knowledge (Phishing Aware or Phishing Unaware) and before having the treatments, there is no difference between the false negative rate (FNR) for high technical ability (HTA) people and the FNR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	24	.7708	.32900
High	24	.7708	.32900

Table 59: Descriptive statistics for FNRs' comparisons related to hypothesis 2.18

As for the previous hypotheses 2.16 and 2.17, Phishing knowledge in this comparison is not considered. The focus is on the level of technical ability for all subjects who participated in all the experiments. Therefore, the false negative rates (FNRs) shown in this hypothesis are based on the decisions taken by subjects with both high technical ability (HTA) and low technical ability (LTA) before having their treatments. Table 59 presents the two rates; one rate for LTA subjects and the other is for the HTA subjects. The FNRs for both the LTA subjects and the HTA subjects are

exactly the same. They are 0.77. Therefore, there is no significant statistical difference between the two FNRs; p=1.000. As a result of this analysis, the null hypothesis 2.18 is accepted.

Hypothesis (2.19):

- **Null Hypothesis (2.19):** For all subjects in the Control and Phishing Aware groups with regard to their technical ability levels and regardless of their Phishing knowledge (Phishing Aware or Phishing Unaware), there is no difference between the correct decisions rate (CDR) for high technical ability (HTA) people and the CDR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	12	.6146	.15501
High	12	.6250	.11918

Table 60: Descriptive statistics for CDRs comparisons related to hypothesis 2.19

Phishing knowledge in this comparison is not considered. The focus is on the level of technical ability of subjects in the Control and Phishing Aware groups. The correct decisions rates (CDRs) shown in this hypothesis are based on the decisions taken by subjects with both high technical ability (HTA) and low technical ability (LTA). Table 60 presents the CDR for LTA subjects and the CDR for the HTA subjects. The CDR for the LTA subjects is about 0.61, whereas it is 0.63 for the HTA subjects. They are nearly the same. There is no significant statistical difference between the two groups; p=.830. As a result of this analysis, the null hypothesis 2.19 is accepted.

Hypothesis (2.20):

- **Null Hypothesis (2.20):** For all subjects in the Control and Phishing Aware groups with regard to their technical ability levels and regardless of their Phishing knowledge

(Phishing Aware or Phishing Unaware), there is no difference between the false positive rate (FPR) for high technical ability (HTA) people and the FPR for low technical ability (LTA) people.

- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	12	.1250	.16855
High	12	.1667	.19462

Table 61: Descriptive statistics for FPRs' comparisons related to hypothesis 2.20

Phishing knowledge in this comparison is not considered. The focus is on the level of technical ability of subjects in the Control and Phishing Aware groups. Therefore, the false positive rates (FPRs) shown in this hypothesis are based on the decisions taken by subjects with both high technical ability (HTA) and low technical ability (LTA). Table 61 shows the two rates; one rate for LTA subjects and the other for the HTA subjects. The FPR for the LTA subjects is about 0.13, whereas it is 0.17 for the HTA subjects. There is no significant statistical difference between the two groups; p=.744. As a result of this analysis, the null hypothesis 2.20 is accepted.

Hypothesis (2.21):

- **Null Hypothesis (2.21):** For all subjects in the Control and Phishing Aware groups with regard to their technical ability levels and regardless of their Phishing knowledge (Phishing Aware or Phishing Unaware), there is no difference between the false negative rate (FNR) for high technical ability (HTA) people and the FNR for low technical ability (LTA) people.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Technical Ability	N	Mean	Std. Deviation
Low	12	.6458	.39107
High	12	.5833	.37437

Table 62: Descriptive statistics for FNRs' comparisons related to hypothesis 2.21

As for the previous hypotheses in this aspect, Phishing knowledge is not considered in this comparison. The focus is on the level of technical ability for subjects in the Control and Phishing Aware groups. Therefore, the false negative rates (FNRs) shown in this hypothesis are based on the decisions taken by subjects with both high technical ability (HTA) and low technical ability (LTA). Table 62 presents the two rates; one rate for LTA subjects and the other for the HTA subjects. The FNR for LTA subjects is 0.65 and the FNR for HTA subjects is 0.58. There is no significant statistical difference between the two FNRs; p=.626. As a result of this analysis, the null hypothesis 2.21 is accepted.

Aspect Discussion:

Hypothesis	Result
Hypothesis 2.16	Accepted
Hypothesis 2.17	Accepted
Hypothesis 2.18	Accepted
Hypothesis 2.19	Accepted
Hypothesis 2.20	Accepted
Hypothesis 2.21	Accepted

Table 63: A summary of hypotheses' analysis results in assessing the effect of the technical ability level among both Phishing Aware and Unaware users

Table 63 shows a summary of the hypotheses analysis results in assessing the effect of technical ability level among both Phishing Aware and Phishing Unaware users. There is no difference between the two groups (high and low technical ability) in properly judging both Phishing and legitimate websites. Therefore, technical ability has no effect on the decisions of Phishing Aware and Phishing Unaware users in Phishing detection and in recognizing legitimate websites.

9.2.2.4. Theme Summary

With regards to both Phishing Aware and Phishing Unaware users, it is found that their technical ability has no effect on their decisions in Phishing detection and in recognizing legitimate websites.

9.2.3. Effect of Phishing Awareness and Phishing Unawareness on Phishing Detection

Hypothesis (3.1):

- **Null Hypothesis (3.1):** There is no difference between the correct decisions rate (CDR) for the Phishing Aware people group and the CDR for the Control group (Phishing Unaware).
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Reject.
- **Discussion:**

Level of Phishing Awareness	N	Mean	Std. Deviation
Unaware	12	.5208	.07217
Aware	12	.7188	.10825

Table 64: Descriptive statistics for CDRs' comparisons related to hypothesis 3.1

The comparisons are focused on two groups. They are the Phishing Aware group and the Control group (Phishing Unaware). The correct decisions rates (CDRs) shown in this hypothesis are based on the decisions taken by subjects who are considered as 'Phishing Aware' and 'Phishing Unaware' people. As presented in Table 64, the CDR for the Phishing Unaware subjects is about 0.52, whereas it is 0.72 for the Phishing Aware subjects. The Phishing Aware group has a higher rate than the Phishing Unaware group. Additionally, there is a significant statistical difference between the two groups; p=.000. As a result of this analysis, the null hypothesis 3.1 is rejected. Therefore, the CDR for the Phishing Aware people group is better (more) than the CDR for the Control group

(Phishing Unaware). This means that Phishing awareness has a significant effect on properly judging legitimate and Phishing websites. Therefore, Phishing awareness has a significant effect on Phishing detection.

Hypothesis (3.2):

- **Null Hypothesis (3.2):** There is no difference between the false positive rate (FPR) for the Phishing Aware people group and the FPR for the Control group (Phishing Unaware).
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Reject.
- **Discussion:**

Level of Phishing Awareness	N	Mean	Std. Deviation
Unaware	12	.0833	.16283
Aware	12	.2083	.17944

Table 65: Descriptive statistics for FPRs' comparisons related to hypothesis 3.2

The comparison is between two groups; Phishing Aware group and Control group (Phishing Unaware). The false positive rates (FPRs) shown in this hypothesis are based on the decisions taken by subjects who are considered 'Phishing Aware' and 'Phishing Unaware' people. As shown in Table 65, the FPR for the Phishing Unaware subjects is about 0.08, whereas it is 0.21 for the Phishing Aware subjects. The Phishing Aware group has a higher (worse) rate than the Phishing Unaware group. There is a significant statistical difference between the two groups; p=.043. As a result of this analysis, the null hypothesis 3.2 is rejected.

The reason why the Phishing Aware people have a worse FPR than the Phishing Unaware people is that, in the Control group, nearly all the subjects responded to the legitimate websites because they believed that they were legitimate websites. However, in the Phishing Aware group, subjects were worried about the websites' legitimacy because they already knew about the existence of Phishing websites in the real world. Then, if they had a legitimate website to respond to, they preferred not to respond in

order to be on the safe side. Thus, the FPR in the Phishing Aware group is higher (worse) than the FPR in the Phishing Unaware group.

Hypothesis (3.3):

- **Null Hypothesis (3.3):** There is no difference between the false negative rate (FNR) for the Phishing Aware people group and the FNR for the Control group (Phishing Unaware).
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Reject.
- **Discussion:**

Level of Phishing Awareness	N	Mean	Std. Deviation
Unaware	12	.8750	.25000
Aware	12	.3542	.29113

Table 66: Descriptive statistics for FNRs' comparisons related to hypothesis 3.3

The comparison is between the Phishing Aware group and the Control group (Phishing Unaware). The false negative rates (FNRs) shown in this hypothesis are based on the decisions taken by subjects who are considered 'Phishing Aware' and 'Phishing Unaware' people. The FNR for the Phishing Unaware subjects is about 0.88, whereas it is 0.35 for the Phishing Aware subjects (See Table 66). The Phishing Aware group has a lower (better) rate than the Phishing Unaware group. There is also a significant statistical difference between the two groups; p=.000. As a result of this analysis, the null hypothesis 3.3 is rejected. Therefore, the FNR for Phishing Aware people group is better (less) than the FNR for the Control group (Phishing Unaware). This means that Phishing awareness has a significant effect on Phishing websites detection.

Hypothesis (3.4):

- **Null Hypothesis (3.4):** For all subjects with regard to their Phishing knowledge (Phishing Aware or Phishing Unaware) and regardless of their technical ability level and before having the treatments, there is no difference between the correct decisions

rate (CDR) for the Phishing Aware people group and the CDR for the Phishing Unaware people group.

- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Reject.
- **Discussion:**

Level of Phishing Awareness	N	Mean	Std. Deviation
Unaware	36	.5069	.07285
Aware	12	.7083	.09731

Table 67: Descriptive statistics for CDRs' comparisons related to hypothesis 3.4

For all subjects who participated in all the experiments, the comparisons are focused on two groups. They are the Phishing Aware group and the Phishing Unaware group. The correct decisions rates (CDRs) shown in this hypothesis are based on the decisions taken before having the treatments by subjects who are considered 'Phishing Aware' and 'Phishing Unaware' people. As shown in Table 67, the CDR for the Phishing Unaware subjects is 0.51, whereas it is nearly 0.71 for the Phishing Aware subjects. This means that the Phishing Aware group has a higher (better) rate than the Phishing Unaware group. The statistical difference between the two groups is significant; p=.000. As a result of this analysis, the null hypothesis 3.4 is rejected. This means that Phishing awareness has a significant effect on properly judging the legitimacy of websites.

Hypothesis (3.5):

- **Null Hypothesis (3.5):** For all subjects with regard to their Phishing knowledge (Phishing Aware or Phishing Unaware) and regardless of their technical ability level and before having the treatments, there is no difference between the false positive rate (FPR) for the Phishing Aware people group and the FPR for the Phishing Unaware people group.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Level of Phishing Awareness	N	Mean	Std. Deviation
Unaware	36	.0833	.18898
Aware	12	.2083	.25746

Table 68: Descriptive statistics for FPRs' comparisons related to hypothesis 3.5

For all subjects who participated in all the experiments, the comparison is between two groups; the Phishing Aware group and the Phishing Unaware group. The false positive rates (FPRs) shown in this hypothesis are based on the decisions taken before having the treatments by subjects who are considered 'Phishing Aware' and 'Phishing Unaware' people. As shown in Table 68, the FPR for the Phishing Unaware subjects is approximately 0.08, whereas it is about 0.21 for the Phishing Aware subjects. The Phishing Aware group has a higher (worse) rate than the Phishing Unaware group. However, there is no significant statistical difference between the two groups; p=.086. As a result of this analysis, the null hypothesis 3.5 is accepted.

The reason why the Phishing Aware people have a worse FPR than the Phishing Unaware is the same reason as that discussed in the hypothesis 3.2. In the Phishing Unaware group, nearly all the subjects responded to the legitimate websites because they believed that they were legitimate websites. However, in the Phishing Aware group, subjects were worried about the websites' legitimacy because they already knew about the existence of Phishing websites in the real world. Then if they had a legitimate website to respond to, they preferred not to respond in order to be on the safe side. Thus, the FPR for the Phishing Aware group is higher (worse) than the FPR for the Phishing Unaware group.

Hypothesis (3.6):

- **Null Hypothesis (3.6):** For all subjects with regard to their Phishing knowledge (Phishing Aware or Phishing Unaware) and regardless of their technical ability level and before having the treatments, there is no difference between the false negative rate (FNR) for the Phishing Aware people group and the FNR for the Phishing Unaware people group.
- **Statistical analysis method used:** Mann-Whitney test.

- **Experimental design used:** Between-subjects.
- **Result:** Reject.
- **Discussion:**

Level of Phishing Awareness	N	Mean	Std. Deviation
Unaware	36	.9028	.20069
Aware	12	.3750	.31079

Table 69: Descriptive statistics for FNRs' comparisons related to hypothesis 3.6

For all subjects who participated in all the experiments, the comparison is between two groups; the Phishing Aware group and the Phishing Unaware. The false negative rates (FNRs) shown in this hypothesis are based on the decisions taken before having the treatments by subjects who are considered 'Phishing Aware' and 'Phishing Unaware' people. The FNR for the Phishing Unaware subjects is 0.90, whereas it is about 0.38 for the Phishing Aware subjects (See Table 69). The Phishing Aware group has a lower (better) FNR than the Phishing Unaware group. The statistical difference between the two groups is significant; p=.000. As a result of this analysis, the null hypothesis 3.6 is rejected. This means that Phishing awareness has a significant effect on Phishing websites detection.

9.2.3.1. Theme Summary

Hypothesis	Result
Hypothesis 3.1	Rejected
Hypothesis 3.2	Rejected
Hypothesis 3.3	Rejected
Hypothesis 3.4	Rejected
Hypothesis 3.5	Accepted
Hypothesis 3.6	Rejected

Table 70: A summary of hypotheses' analysis results in assessing the effect of Phishing awareness and Phishing unawareness on Phishing detection

Table 70 shows a summary of the results of the six hypotheses discussed for assessing the effect of Phishing awareness and Phishing unawareness on Phishing

websites detection. There is significant effect for the Phishing Aware users in accurately detecting Phishing websites and this allows them to prevent Phishing attacks. In total, the decisions of Phishing Aware users are better than the decisions of Phishing Unaware users. This appears in the comparisons of the correct decisions rates (CDRs) of the two groups. The difference between the CDRs shows that there is a significant positive effect of Phishing awareness in comparison with Phishing unawareness. As a result of this, Phishing awareness has a significant positive effect on users' decisions in websites' legitimacy.

9.2.4. Anti-Phishing Knowledge Retention

The theme of anti-Phishing knowledge retention has two different aspects. Each aspect is discussed separately. They are as follows:

9.2.4.1. Aspect: Assessing the Retention of Anti-Phishing Knowledge within Each Individual Group

Hypothesis (4.1):

- **Null Hypothesis (4.1):** With regards to the post-treatment websites, there is no difference between the correct decisions rate (CDR) of the Old Approach group in the second experiment and their CDR in the first experiment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Experiment	N	Mean	Std. Deviation
CDR in the first experiment	6	.6250	.13693
CDR in the second experiment	6	.5833	.20412

Table 71: Descriptive statistics for CDRs' comparisons related to hypothesis 4.1

The correct decisions rates (CDRs) shown in this hypothesis are based on the decisions taken by the Old Approach group subjects after having their treatment (i.e. an anti-Phishing training email) in the first experiment and in the second experiment. The period between the two experiments varied from subject to subject. However, the average period was 16.7 days. Table 71 presents the two rates for the same subjects in the two experiments. The CDR in the first experiment is about 0.63, whereas it is 0.58 in the second experiment. There is no significant statistical difference between the two rates; p=.500. As a result of this analysis, the null hypothesis 4.1 is accepted.

The CDR is higher (better) in the first experiment than in the second experiment. This means that subjects performed better in terms of the proper judgment of websites' legitimacy at the time they received the Old Approach training. They did not maintain exactly the same performance when they were given the same tricks after approximately 16.7 days. Their performance in the two experiments did not differ significantly.

Hypothesis (4.2):

- **Null Hypothesis (4.2):** With regards to the post-treatment websites, there is no difference between the false positive rate (FPR) of the Old Approach group in the second experiment and their FPR in the first experiment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Experiment	N	Mean	Std. Deviation
FPR in the first experiment	6	.0833	.20412
FPR in the second experiment	6	.4167	.49160

Table 72: Descriptive statistics for FPRs comparisons related to hypothesis 4.2

The false positive rates (FPRs) shown in this hypothesis are based on the decisions taken by the Old Approach group subjects after having their treatment in the first experiment and in the second experiment. As presented in Table 72, the FPR in the first experiment is approximately 0.08, whereas it is 0.42 in the second experiment. The FPR

is higher (worse) in the second experiment than the rate in the first experiment. However, there is no significant statistical difference between the two groups; p=.125. As a result of this analysis, the null hypothesis 4.2 is accepted.

This means that subjects performed better in terms of properly judging legitimate websites at the time they received the Old Approach treatment. However, they did not exactly maintain the same performance when they were given the same tricks after approximately 16.7 days.

Hypothesis (4.3):

- **Null Hypothesis (4.3):** With regards to the post-treatment websites, there is no difference between the false negative rate (FNR) of the Old Approach group in the second experiment and their FNR in the first experiment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Experiment	N	Mean	Std. Deviation
FNR in the first experiment	6	.6667	.25820
FNR in the second experiment	6	.4167	.37639

Table 73: Descriptive statistics for FNRs' comparisons related to hypothesis 4.3

The false negative rates (FNRs) shown in this hypothesis are based on the decisions taken by the Old Approach group subjects after having their treatment in the first experiment and in the second experiment. As presented in Table 73, the FNR in the first experiment is approximately 0.67, whereas it is 0.42 in the second experiment. The FNR is lower (better) in the second experiment than the rate in the first experiment. However, there is no significant statistical difference between the two groups; p=.125. As a result of this analysis, the null hypothesis 4.3 is accepted.

This means that the subjects who had taken the Old Approach training performed better (but with no statistical difference) in terms of properly detecting Phishing websites after approximately 16.7 days.

Hypothesis (4.4):

- **Null Hypothesis (4.4):** With regards to the post-treatment websites, there is no difference between the correct decisions rate (CDR) of the New Approach group in the second experiment and their CDR in the first experiment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Experiment	N	Mean	Std. Deviation
CDR in the first experiment	6	.7500	.15811
CDR in the second experiment	6	.5417	.18819

Table 74: Descriptive statistics for CDRs' comparisons related to hypothesis 4.4

The correct decisions rates (CDRs) shown in this hypothesis are based on the decisions taken by the New Approach group subjects after having their treatment (i.e. anti-Phishing intervention) in the first experiment and in the second experiment. The period between the two experiments was 16.7 days as an average. As Table 74 shows, the CDR in the first experiment is about 0.75. The rate in the second experiment is 0.54. There is no significant statistical difference between the two groups; p=.156. As a result of this analysis, the null hypothesis 4.4 is accepted.

The CDR is higher (better) in the first experiment than the CDR in the second experiment. This means that subjects performed better in terms of the proper judgment of the legitimacy of websites at the time they received the New Approach training. However, they did not maintain exactly the same performance when they were given the same tricks after approximately 16.7 days.

Hypothesis (4.5):

- **Null Hypothesis (4.5):** With regards to the post-treatment websites, there is no difference between the false positive rate (FPR) of the New Approach group in the second experiment and their FPR in the first experiment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.
- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Experiment	N	Mean	Std. Deviation
FPR in the first experiment	6	.0833	.20412
FPR in the second experiment	6	.1667	.25820

Table 75: Descriptive statistics for FPRs' comparisons related to hypothesis 4.5

The false positive rates (FPRs) shown in this hypothesis are based on the decisions taken by the New Approach group subjects after having their treatment in the first experiment and in the second experiment. As presented in Table 75, the FPR in the first experiment is approximately 0.08, whereas it is 0.17 in the second experiment. There is no significant statistical difference between the two groups; p=.500. As a result of this analysis, the null hypothesis 4.5 is accepted.

The false positive rate FPR is higher (worse) in the second experiment than the FPR in the first experiment. This means that subjects performed better in terms of properly judging legitimate websites at the time they received the New Approach training. However, they did not exactly maintain the same performance when they were given the same tricks after approximately 16.7 days.

Hypothesis (4.6):

- **Null Hypothesis (4.6):** With regards to the post-treatment websites, there is no difference between the false negative rate (FNR) of the New Approach group in the second experiment and their FNR in the first experiment.
- **Statistical analysis method used:** Wilcoxon Signed Ranks test.

- **Experimental design used:** Within-subjects.
- **Result:** Accept.
- **Discussion:**

Experiment	N	Mean	Std. Deviation
FNR in the first experiment	6	.4167	.37639
FNR in the second experiment	6	.7500	.41833

Table 76: Descriptive statistics for FNRs' comparisons related to hypothesis 4.6

The false negative rates (FNRs) shown in this hypothesis are based on the decisions taken by the New Approach group subjects after having their treatment in the first experiment and in the second experiment. As presented in Table 76, the FNR in the first experiment is approximately 0.42, whereas it is 0.75 in the second experiment. There is no significant difference between the two groups; p=.156. As a result of this analysis, the null hypothesis 4.6 is accepted.

The FNR is higher (worse) in the second experiment than the FNR in the first experiment. This means that the subjects performed (but with no statistical difference) better in terms of properly detecting Phishing websites just after they took the New Approach than when they repeated the experiment after in approximately 16.7 days.

Aspect Discussion:

Hypothesis	Result
Hypothesis 4.1	Accepted
Hypothesis 4.2	Accepted
Hypothesis 4.3	Accepted
Hypothesis 4.4	Accepted
Hypothesis 4.5	Accepted
Hypothesis 4.6	Accepted

Table 77: A summary of hypotheses' analysis results in assessing anti-Phishing knowledge retention for users within each group

Approximately 16 days after conducting the first experiment, users of the New Approach and Old Approach were asked to perform a follow up experiment. The goal was

to see which approach's users retain their anti-Phishing knowledge better. Table 77 presents a summary of the results in assessing anti-Phishing knowledge retention in each group. There are no differences between the CDRs, FPRs and FNRs for the two experiments for each approach's users. This means that the users of both approaches retained their anti-Phishing knowledge after 16 days from their first training. More importantly, the CDRs at the time of training were slightly better (but with no statistical difference) than their decisions after 16 days in both approaches.

9.2.4.2. Aspect: Comparing the Retention of Anti-Phishing Knowledge between Groups

Hypothesis (4.7):

- **Null Hypothesis (4.7):** In the second experiment, there is no difference between the correct decisions rate (CDR) for the New Approach group and the CDR for the Old Approach group.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Experiment	N	Mean	Std. Deviation
CDR in the Old Approach group	6	.5225	.12259
CDR in the New Approach group	6	.5417	.12910

Table 78: Descriptive statistics for CDRs' comparisons related to hypothesis 4.7

The correct decisions rates (CDRs) shown in this hypothesis are based on the decisions taken by the two groups' subjects in the second experiment. The average period between the two experiments was 16.7 days. Table 78 presents the two rates for the Old Approach and New Approach groups. The CDR for the Old Approach group is about 0.52, whereas it is 0.54 for the New Approach group. The two rates are nearly the same. There is no significant statistical difference between the two groups; p=.526. As a result of this analysis, the null hypothesis 4.7 is accepted.

This means that subjects of both approaches performed nearly equally in terms of properly judging the legitimacy of websites after approximately 16.7 days.

Hypothesis (4.8):

- **Null Hypothesis (4.8):** In the second experiment, there is no difference between the false positive rate (FPR) for the New Approach group and the FPR for the Old Approach group.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Experiment	N	Mean	Std. Deviation
FPR in the Old Approach group	6	.4583	.45871
FPR in the New Approach group	6	.0833	.12910

Table 79: Descriptive statistics for FPRs' comparisons related to hypothesis 4.8

The false positive rates (FPRs) shown in this hypothesis are based on the decisions taken by the two groups' subjects in the second experiment. As presented in Table 79, the FPR for the Old Approach group is approximately 0.46, whereas it is 0.08 for the New Approach group. There is no significant statistical difference between the two groups; p=.089. As a result of this analysis, the null hypothesis 4.8 is accepted.

Hypothesis (4.9):

- **Null Hypothesis (4.9):** In the second experiment, there is no difference between the false negative rate (FNR) for the New Approach group and the FNR for the Old Approach group.
- **Statistical analysis method used:** Mann-Whitney test.
- **Experimental design used:** Between-subjects.
- **Result:** Accept.
- **Discussion:**

Experiment	N	Mean	Std. Deviation
FNR in the Old Approach group	6	.5000	.31623
FNR in the New Approach group	6	.8333	.30277

Table 80: Descriptive statistics for FNRs' comparisons related to hypothesis 4.9

The false negative rates (FNRs) shown in this hypothesis are based on the decisions taken by the two groups' subjects in the second experiment. The FNR for the Old Approach group is exactly 0.50 whereas it is nearly 0.83 in the New Approach group (See Table 80). There is no significant statistical difference between the two groups; p=.067. As a result of this analysis, the null hypothesis 4.9 is accepted.

Aspect Discussion:

Hypothesis	Result
Hypothesis 4.7	Accepted
Hypothesis 4.8	Accepted
Hypothesis 4.9	Accepted

Table 81: A summary of hypotheses' analysis results in assessing anti-Phishing knowledge retention for users between groups

Approximately 16 days after conducting the first experiment, the users of the New Approach and Old Approach were asked to perform a follow up experiment. The goal was to see which approach's users retained their anti-Phishing knowledge better. Table 81 shows a summary of the results in the aspect of assessing the retention of the anti-Phishing knowledge by the Old Approach and New Approach groups in the second experiment. There are no statistical differences between CDRs, FPRs and FNRs of both approaches' users in the second experiment. This means that the subjects of both approaches performed nearly equally in terms of properly judging legitimate and Phishing websites approximately 16.7 days after they had the two approaches.

9.2.4.3. Theme Summary

Hypothesis	Result
Hypothesis 4.1	Accepted
Hypothesis 4.2	Accepted
Hypothesis 4.3	Accepted
Hypothesis 4.4	Accepted
Hypothesis 4.5	Accepted
Hypothesis 4.6	Accepted
Hypothesis 4.7	Accepted
Hypothesis 4.8	Accepted
Hypothesis 4.9	Accepted

Table 82: A summary of hypotheses' analysis results in assessing anti-Phishing knowledge retention for users

Approximately 16 days after conducting the first experiment, users of both the New and Old Approaches were asked to perform a follow up experiment. The goal was to see which approach's users retained their anti-Phishing knowledge better. Two aspects were discussed. The first aspect assessed the anti-Phishing knowledge retention in each group between the first and second experiments. The other aspect assessed the anti-Phishing knowledge retention in the second experiment between the two groups. As summarized in Table 82, there are no differences between the rates (CDRs, FPRS and FNRs) in the first experiment and the rates in the second experiment in each group. There are also no statistical differences between the rates for both approaches' users in the second experiment. It is found that firstly, users of both approaches retained their anti-Phishing knowledge after 16 days from their first training. More importantly, the CDRs at the time of training were slightly better (with no statistical difference) than their decisions after 16 days. Secondly, the subjects of both approaches performed nearly equally in terms of properly judging the legitimacy of websites after approximately 16.7 days.

There are two facts resulting from the analysis of anti-Phishing knowledge retention. The first fact is that in the first experiment, the New Approach group is different significantly from the Old approach group with regards to correctly judging the legitimacy of websites (See Section 9.2.1.2). The second fact is that when the retention of anti-Phishing knowledge was assessed, there is no difference between the correct decision rate (CDR), the false positive rate (FPR) and the false negative rate (FNR) of the Old Approach group in the

second experiment and their rates in the first experiment. Similarly, the same results were found for the New Approach group (See Section 9.2.4.1).

Based on the facts mentioned above, when the participants of the two groups were evaluated in the second experiment, logically and in theory the New Approach group should judge the legitimacy of websites better than the Old Approach group. However, in practice there is no significant difference between the two groups (See Section 9.2.4.2). This seems to be inconsistent.

The reason behind this inconsistency is that the sample size in the first experiment is different from the second experiment. There were 24 subjects in the two groups in the first experiment whereas there were 12 subjects in the second experiment.

Therefore, the data of the 12 subjects, who participated in the second experiment, collected in the first experiment was analyzed. It is found that there are no significant differences between the CDRs, FPRs and FNRs of the New Approach group and the Old Approach group. More importantly, the statistical difference between the CDRs is p=.116. With regards to the statistical difference between FPRs, it is p=.773 and it is p=.119 between the two FNRs. As a result of this, it is clear that the other 12 participants participated in the two groups in the first experiment made the difference.

9.3. Summary

This chapter evaluated the research hypotheses. The hypotheses were related to four different research themes and were assessed by statistical analysis. The chapter presented the themes evaluating the New Approach, the effect of high and low technical abilities on Phishing detection, the effect of Phishing awareness and Phishing unawareness on Phishing detection and anti-Phishing knowledge retention. The main results achieved in each theme are as discussed below.

1. Evaluating the New Approach:

Evaluating the New Approach had three aspects. They are assessing users without treatments, assessing the New Approach in comparison with the Old Approach and assessing users before and after using the treatments. For assessing users without treatments, users in the three groups were nearly equal with regards to their decisions about legitimate and Phishing websites before using any treatment. There was no difference in their correct decisions' rate (CDR). There were also no differences in their false positive and false negative rates. The aim of assessing users without treatments was to make sure that there were no differences between users in the three groups (Control, New Approach and Old Approach) before having any treatment.

In the aspect of assessing the New Approach in comparison with the Old Approach, it was found that there was a significant positive effect of using the New Approach in comparison with the Old Approach. The users of the New Approach were significantly better in judging legitimate and Phishing websites. In detail, the New and the Old Approaches were equal on helping the subjects to judge the legitimate websites. However, there was a significant effect of using the New Approach in comparison with the Old Approach in helping subjects to judge the Phishing websites properly and this enabled them to detect the Phishing attacks. This means that the New Approach worked better than the Old Approach. This was demonstrated by the comparisons of the correct decisions' rate (CDR) of the two approaches.

In assessing users before and after using the treatments, in the Control group, there were no differences between the correct decisions rates (CDRs), false positive rates (FPRs) and false negative rates (FNRs) before and after having the treatment (in this instance, an ordinary email from work). In the Old Approach group, there was no difference between the FPRs and the FNRs. The CDRs also were not significantly different. The CDR is more important because it is indicative of users' decisions on the total of both legitimate and Phishing websites. Therefore, the Old Approach had no significant effect on users' decisions on the legitimacy of websites. In the New Approach group, there was no significant difference between the FPRs before and after having the treatment (i.e. an anti-Phishing intervention). However, there was a significant difference between the FNRs (for the rate after the treatment). More importantly, the CDRs were significantly different. The CDR

after having the treatment was better than the rate before the treatment. Therefore, the New Approach had a significant effect on users' decisions on the legitimacy of websites.

To sum up, the New Approach is better than the Old Approach in helping users to detect Phishing websites. The New Approach has a significant positive effect on users' decisions.

2. Effects of high and low technical abilities on Phishing detection:

There were three aspects in this evaluation. These were assessing the effect of the technical ability level among Phishing unaware users, assessing the effect of the technical ability level among Phishing aware users and assessing the effect of the technical ability level regardless of the Phishing knowledge (Phishing aware and unaware). With regards to the first aspect, the hypotheses' analysis evaluated users without having treatments in the three groups (Control, New Approach and Old Approach). The result is that there was no significant difference between the decisions rates of high and low technical ability users. Therefore, the level of the technical ability had no effect on Phishing detection or on recognizing legitimate websites among Phishing unaware people.

In terms of assessing the effect of the technical ability level among Phishing aware users, there was no difference between the high and low technical ability users in Phishing detection and prevention. There was also no difference between the two groups (high and low technical ability) in properly judging legitimate websites. Therefore, technical ability had no effect on the decisions of Phishing Aware users in Phishing detection and in recognizing legitimate websites.

Assessing the effect of technical ability level among both Phishing Aware and Phishing Unaware users showed that there was no difference between the two groups (high and low technical ability) in properly judging both Phishing and legitimate websites. Therefore, technical ability had no effect on the decisions of Phishing Aware and Phishing Unaware users in Phishing detection and in recognizing legitimate websites.

To sum up, the technical ability of users does not have an effect on their ability to detect Phishing websites.

3. Effect of Phishing awareness and Phishing unawareness on Phishing detection

The results of the six hypotheses discussed for assessing the effect of Phishing awareness and Phishing unawareness on Phishing websites detection showed that there was significant effect for the Phishing Aware users in accurately detecting Phishing websites and this allows them to prevent Phishing attacks. In total, the decisions of Phishing Aware users were better than the decisions of Phishing Unaware users. This appeared in the comparisons of the correct decisions rates (CDRs) of the two groups. The difference between the CDRs showed that there was a significant positive effect of Phishing awareness in comparison with Phishing unawareness.

To sum up, Phishing awareness has a significant positive effect on users' ability to detect Phishing websites. Phishing Aware users were better than Phishing Unaware users in detecting Phishing websites.

4. Anti-Phishing knowledge retention:

Approximately 16 days after conducting the first experiment, users of both the New and Old Approaches were asked to perform a follow up experiment. The goal was to see for which approach the users retained their anti-Phishing knowledge better. Two aspects were discussed. The first aspect assessed the anti-Phishing knowledge retention in each group between the first and second experiments. The other aspect assessed the anti-Phishing knowledge retention in the second experiment between the two groups. For the first aspect, there were no differences between the rates (CDRs, FPRS and FNRs) in the first experiment and the rates in the second experiment in each group. The results of the second aspect showed that there were also no statistical differences between the rates for both approaches' users in the second experiment.

To sum up, users retain the anti-Phishing knowledge given to them by both the New Approach and the Old Approach. They were slightly better (with no statistical difference) in detecting Phishing websites at the time they first used the approaches. However, users of both approaches were nearly the same in terms of properly detecting Phishing websites approximately 16 days after having their approaches.

10. Comparisons

10.1. Introduction

After evaluating the New Approach, the effects of technical ability and Phishing knowledge on Phishing detection and legitimate website recognition and anti-Phishing knowledge retention in the previous chapter, this chapter presents comparisons with some related anti-Phishing approaches by others. It looks at the similarities and differences between the evaluations in this work and the work of others. Evaluation issues such as participants' recruitment, groups, scenarios, emails and websites, anti-Phishing tips used and implementation are discussed. Comparisons of the results are shown.

10.2. Evaluation

10.2.1. Participants Recruitment

Kumaraguru et al. [Kumaraguru et al.07a] evaluated their approach using participants who were recruited based on their technical abilities (TA), using the criteria presented in Chapter 3. People were classified as 'experts' or 'non-experts' using pre-study screening questions. Technical ability was judged on whether the participants had changed preferences or settings in their web browser, created a web page, and helped someone fix a computer problem. The participant who said 'no' to at least two of the screening questions was considered as 'non-expert' and selected to take part in their experiments. Kumaraguru et al. had 30 participants distributed equally into three groups (10 participants each group). The groups were called the security notices group, the graphical training intervention group and

the comic strip training intervention group. Therefore, they had two different training interventions groups and a security notices (Old Approach) group. They did not have a control group which did not take any treatment.

In this thesis, the pre-study survey included questions about the Internet and email use, technical ability, web browser knowledge and knowledge of computer terms. The knowledge of computer related terms section had the question about Phishing knowledge. The questions about Phishing knowledge and participant's technical ability were the main concerns in the survey. A Phishing-Aware person is the one who defines Phishing correctly. Technical ability was judged based on the criteria used by Kumaraguru et al. [Kumaraguru et al.07a]. The expert and non-expert users, in terms of their technical abilities, are named 'high' and 'low' technical users respectively. High and low technical ability people were included in the experiments.

In the evaluation experiments of the New Approach (APTIPWD), there were three groups, Control, Old Approach and New Approach. All participants were 'Phishing Unaware' regardless of their technical ability level. There were 36 participants in the experiment. Each group had 12 participants divided into two subgroups, High and Low technical ability. Each subgroup had 6 participants. The Old Approach group was nearly the same as the security notices group in Kumaraguru et al.'s approach evaluation because their treatment was an anti-Phishing email. However, the number of anti-Phishing tips given to both groups is different and discussed in Section 10.3.

Kumaraguru et al. [Kumaraguru et al.07a] conducted two surveys, a pre-study survey and a post-study survey. The aim from having the pre-study survey was to select only 'non-experts' users to participate in the experiments. The post-study survey was to debrief the participants and ask them for feedback about their approach. In this research, two surveys were conducted. They were a pre-study survey and a pre-session survey. The pre-study survey was to classify participants into 'high technical ability' and 'low technical ability' and to select only Phishing Unaware people to participate in the experiments. The pre-session survey took place when the participants came to the experiment's location and it was administered just before the participants performed the experiment. Its aim was to check whether each participant was properly classified. This is because the participants' technical ability or Phishing knowledge could have changed in the period between the pre-study

survey and the experiments. This means that when the pre-study survey was done, participants who:

- had been classified as 'low' technical ability could be 'high' because their technical ability had improved,
- were classified as 'high' technical ability could be 'low' because they could have made a mistake in answering the technical ability question in the pre-study survey,
- were classified as 'Phishing Unaware' could be 'Aware' because they may have gained knowledge about Phishing from another source and
- were classified as 'Phishing Aware' could be 'Unaware' because they could have made a mistake in answering the Phishing knowledge question in the pre-study survey.

Table 83 shows a comparative summary of participant recruitment.

Research	Approach	Recruitment Surveys			Recruitment Criteria		Participants	Groups
		Pre-study	Pre-session	Post-study	TA	PK		
This research	Anti-Phishing approach for websites.	Yes	Yes	No	Yes	Yes	36	3
Kumaraguru et al. [Kumaraguru et al.07a]	Anti-Phishing approach for emails.	Yes	No	Yes	Yes	No	30	3

Table 83: Summary of participant recruitment comparison discussion

10.2.2. Effectiveness Ratios

Kumaraguru et al.'s [Kumaraguru et al.07a] approach used one effectiveness ratio for evaluation. This effectiveness ratio is False Negative Rate (FNR), which reflects the participants' decisions about Phishing websites. In contrast, three effectiveness ratios were

used in the New Approach's evaluation. They are Correct Decisions Rate (CDR), False Positive Rate (FPR) and False Negative Rate (FNR). The CDR indicates the participants' decisions against both legitimate and Phishing websites. The FPR shows the participants' decisions against legitimate websites.

Due to the fact that Kumaraguru et al.'s approach evaluation used just the ratio FNR, only the FNR results related to the New Approach are presented in Section 10.4.1. The CDR and FPR results are not presented.

10.2.3. Scenarios

The New Approach's evaluation experiments and those of Kumaraguru et al. [Kumaraguru et al.07a] used email and web role-play protocol. However, the New Approach's experiments used 14 emails and 9 legitimate and Phishing websites. Kumaraguru et al. used 19 emails but they did not specify the number of websites used. Their approach was focused on emails whereas the New Approach was focused on websites. This may clarify why they used a larger number of emails. Table 84 presents a summary of the scenario comparison discussion presented in this section.

Research	Approach	Scenario	
		# Emails	# Websites
This research	Anti-Phishing approach for websites.	14	9
Kumaraguru et al. [Kumaraguru et al.07a]	Anti-Phishing approach for emails.	19	Not reported

Table 84: Summary of scenario comparison discussion

10.2.4. Implementation

The evaluation experiments of both the New Approach and Kumaraguru et al.'s [Kumaraguru et al.07a] approach used identical copies of real emails and websites. The legitimate and Phishing websites were stored on local machines and run by Apache servers. However, Kumaraguru et al. created emails using *SquirrelMail*. *SquirrelMail* is a standards-based web mail package that is easily administered [SquMail]. They used it to gain control of the messages they sent to their participants. In the New Approach's experiments, fake emails were written and sent using Microsoft Outlook 2002 and were read using *Maktoob* email portal [Maktoob]. Emails were sent by using *Maktoob*'s MX Record as the outgoing mail or server.

10.3. Training

Kumaraguru et al.'s [Kumaraguru et al.07a] approach used many anti-Phishing tips to train users to detect Phishing emails. The tips are Phishing emails with a professional looking format and message content, they are urgent messages, they warn of an account status threat and they have links that do not match with the status bar. In contrast, in evaluating the New Approach, one anti-Phishing tip for detecting a Phishing website was used. It was used by the intervention given to the New Approach users and it was also sent to the Old Approach users by email. This anti-Phishing tip was evaluated as the most effective tip in the evaluation of users' tips for Phishing websites detection, discussed in Chapter 6.

Kumaraguru et al.'s [Kumaraguru et al.07a] approach used multimedia to present the anti-Phishing tips for detecting Phishing emails. The approach explains the anti-Phishing tips using screenshots for Phishing emails and comic strips. Kumaraguru et al. gave their participants anti-Phishing training twice. However, the anti-Phishing tip used in the New Approach was presented as plain text. The New Approach presented the anti-Phishing training only once.

The New Approach did not give training more than once and did not use multimedia or comic strips. This aimed to evaluate the effectiveness of the idea of training intervention in Phishing websites detection. Table 85 shows a summary of training comparison discussion.

Research	Approach	Clues	# Training	Multimedia
This research	Anti-Phishing approach for websites.	Suspicious URL	Single	No
Kumaraguru et al. [Kumaraguru et al.07a]	Anti-Phishing approach for emails.	1. Professional looking emails and messages content. 2. Urgent messages. 3. Account status threat. 4. Links does not match with status bar.	Double	Screenshots and Comic strip

Table 85: Summary of training comparison discussion

10.4. Results

10.4.1. Assessment Parts

There are three parts in assessing the New Approach. They are assessing users without treatments, assessing users after having the treatments and assessing users before and after the treatments. Kumaraguru et al.'s [Kumaraguru et al.07a] evaluation assessed users in just two parts. They assessed users after having treatments and assessed users before and after the treatments. Kumaraguru et al. did not statistically compare their groups before having any treatments. However, in the New Approach's evaluation, comparisons were carried out between the three groups before they had the treatments in order to make sure that there were no significant differences between the groups prior to the treatment.

Regarding assessing users without treatment, in the New Approach's evaluation, there were no significant differences between the FNRs for the three groups. The statistical difference between each group and others is not significant; p=1.000. It was shown that users were equal with regards to their decisions to Phishing websites before using treatments.

In terms of assessing the approaches after having the treatment, Kumaraguru et al. [Kumaraguru et al.07a] found that there was a significant difference between the notices group and the comic strips intervention group (p=.001). There was also significant difference in the effectiveness of the treatment between the graphical intervention group and the comic strip group (p=.001). However, the difference between the notices group and the graphical intervention group was not significant (p=.546). The mean scores across Phishing emails after the intervention were lowest for the comic strip group. In the New Approach's evaluation, there was a significant difference between the FNRs of the New Approach and the Old Approach groups (p=.012). There was also a significant effect of the New Approach in comparison with the Control group (p=.001). Furthermore, there was no statistical significant difference between the Old Approach group and the Control group (p=.107).

In assessing users before and after using the treatments, Kumaraguru et al. did not statistically compare the participants' decisions before and after their treatments within each group. However, they compared the decisions between their groups accrued before and after using the treatments. They found that there was a significant difference between the notices group and comic strip group (p=.001). Also there was a significant difference between the graphical intervention group and the comic strip group (p=.007). There was no significant difference between the notices group and the graphical intervention group. Regarding the New Approach's evaluation, statistical comparisons between the participants' decisions before and after their treatments within each group were conducted. In the New Approach group, there was a significant statistical difference between the FNR after the treatment and the FNR before the treatment (p=.001). However, there was no statistical difference between the FNR after the treatment and the FNR before the treatment in the Old Approach group (p=.063). There was also no statistical difference between the FNRs before and after the treatment (in this instance an ordinary email from work, which essentially was no treatment) in the Control group (p=.500).

10.4.2. Evaluation of the Effects of Technical Ability and Phishing Knowledge

In this research, the effects of technical ability and Phishing knowledge on Phishing websites' detection were evaluated. User experiments were conducted in the evaluation. Regarding the effects of technical ability, the low technical ability (LTA) and the high technical ability (HTA) participants were nearly equal in the correctness of their decisions on legitimate and Phishing websites. Therefore, the technical ability had no effect on the decisions of users in Phishing websites detection and in recognizing legitimate websites. In contrast, it was shown that there was a significant positive effect for Phishing knowledge on Phishing websites detection. Phishing Aware people were better than Phishing Unaware people on Phishing websites detection.

Because technical ability has no effect on the decisions of users in Phishing websites' detection, recruiting people based on their technical ability without knowing about their Phishing knowledge in order to conduct anti-Phishing experiments may produce biased results. This is because both low and high technical people may be Phishing Aware before participating in the evaluation experiments. People who know about Phishing may use their own Phishing knowledge rather than the anti-Phishing approaches' when they participate in an evaluation. Therefore, in evaluating an anti-Phishing approach, recruiting users based on their Phishing knowledge is better than recruiting them based on their technical ability.

10.4.3. Anti-Phishing Knowledge Retention

The evaluation was made of anti-Phishing knowledge retention for the users' of the New Approach (embedded) in comparison with the users' of the Old Approach of sending anti-Phishing tips by email (non-embedded). Two user experiments were conducted to evaluate the retention of the anti-Phishing knowledge. It was found that users of both approaches retained their anti-Phishing knowledge after 16 days from their first training. Users' decisions at the time of the training was slightly better (i.e. no statistical difference) than their decisions after about 16 days. Additionally, users of the two approaches performed

nearly equally in terms of properly judging legitimate and Phishing websites after about 16 days from experiencing the two approaches. With regards to the Kumaraguru et al.'s [Kumaraguru et al.07c] study, they compared the effectiveness of the training materials delivered via their approach (embedded or training multimedia intervention) and delivered via email messages (non-embedded). They found that participants in their approach group retained more knowledge than participants in a non-embedded training group. There was a significant difference between the two groups in identifying correctly the Phishing email.

The results related to the Kumaraguru et al.'s study in anti-Phishing knowledge retention are better than the results in the retention study in this research. This might be because of three reasons. The first reason is that Kumaraguru et al. used multimedia for presenting their anti-Phishing materials. Multimedia (screenshots and comic strips) has a positive effect on information retention [Large06]. In contrast, anti-Phishing materials were shown in plain text in this research.

The second reason is the difference between the periods between the first and second experiments in the two studies. The period in this research (mean= 16.7 days) is more than double the period in the Kumaraguru et al. study (mean= 7.2 days). The difference between the two periods might affect users' anti-Phishing knowledge retention because one of the factors that can affect people's knowledge retention is the time interval between training and practice [StothardNicholson01]. The longer the time between training and practice, the greater skill loss that people can have [ibid].

The third reason is that Kumaraguru et al. gave their participants training material twice whereas participants of the New Approach were given training material once (i.e. one intervention). Kumaraguru et al. state that the double training in a short time was helpful because some participants did not understand what was happening the first time the training information was shown but they read it carefully in the second time.

However, one advantage of the New Approach is that it keeps the anti-Phishing training as an ongoing process (See Figure 26 in Chapter 7). Every time users try to submit information to a Phishing website, they will be trained. Therefore, the New Approach has the capability to train users many times, which in turn improves their ability to detect Phishing websites.

Kumaraguru et al. [Kumaraguru et al.07c] used 42 participants in their anti-Phishing retention study. However, the research in this thesis used 12 participants who were asked to participate in the retention experiments (the second experiment). They were among the 36 participants who took part in the first experiment. The target number of participants in the second experiment was as many as could come back. However, participants with low technical ability (LTA) and high technical ability (HTA) were needed within each group in order to ensure equal chances. Therefore, 12 participants were available and participated in the second experiment. Table 86 presents the comparison of participants and the period between experiments discussed in the anti-Phishing knowledge retention section.

Research	Approach	Participants	Period between Experiments
This research	Anti-Phishing approach for websites	12	16.7 days
Kumaraguru et al. [Kumaraguru et al.07a]	Anti-Phishing approach for emails	42	7.2 days

Table 86: Summary of anti-Phishing knowledge retention comparison discussion

10.5. Comparison with another Approach

In August, 2008, the APWG and Carnegie Mellon CyLab launched the "*Phishing Education Landing Page Program*" (PELPP) [PEI08]. There is a similarity and differences between PELPP and the New Approach proposed in this research.

The similarity between them is that they consider helping people about Phishing websites detection during their normal use of the Internet. However, there are some differences between The Phishing Education Landing Page Program (PELPP) and the New Approach. They are as follows:

1. The Phishing Education Landing Page Program (PELPP) requires an involvement of external parties such as ISPs, registrars, and persons who have control of the Phishing page. This involvement is vital for the project to work since the external parties' need to redirect any Phishing URL to an anti-Phishing training webpage. Therefore, the PELPP requires amendments in the external parties' servers.
2. The New Approach does not require an involvement of other parties since it is based on its own proxy and blacklists. Therefore, it works by its own components.
3. The New Approach uses the most effective tips evaluated by the research in this thesis whereas the PELPP does not state the reason why they use the tips presented in their ant-Phishing training webpage (See Figure 17 in Chapter 3).
4. PELPP has been proposed but not been evaluated whereas the New Approach is evaluated and showed that it is more effective in helping users distinguish between legitimate and Phishing websites than the Old Approach of sending anti-Phishing tips by email.

10.6. Summary

In this chapter, comparisons between the evaluations carried out in this research and some related anti-Phishing approaches by others were presented. The chapter presented discussions on the similarities and differences on issues such as participants' recruitment, groups, scenarios, emails and websites, anti-Phishing tips used, implementation and results.

There were comparable issues such as participants' recruitment, effectiveness ratios, scenarios, implementation and training strategies. However, the results of evaluating the New Approach with the related studies were not comparable because the groups in the two studies were different. There were also differences in issues such as participants' knowledge (before participating in the experiments), the tips given to the participants in the experiments, the number of times that the intervention was given to the participants and the period length between the two phases of experiments.

11. Conclusions and Future Work

11.1. Introduction

The Internet has become a very important medium of communication recently. Many people go online and do a wide range of businesses. They can send emails, sell and buy goods, do different banking activities and even participate in political and social elections by casting a vote online.

Security for conducting businesses online is vital and critical. All security-critical applications (e.g. online banking login page) that are accessed using the Internet are at risk of Internet fraud. Once users go online, they are at risk from online fraud (also known as Internet fraud). The parties involved in any transaction never need to meet and the user may have no idea whether the goods or services exist. Due to this, the Internet is a good vehicle to defraud the users who would like to buy goods or services using it [Philippsohn01]. The application access keys could be stolen. Applications such as e-commerce, online banking, e-voting, email and so forth might be targets for fraudsters. Violating the security in these applications would result in severe consequences such as financial loss in area such as e-commerce and online banking.

Phishing attacks are forms of Internet fraud and have become a serious problem for Internet users. The problem is when a user receives a Phishing email. The user's intention may be "go to eBay" but the actual implementation of the hyperlink may be "go to a server in South Korea" [Wu06]. Users gain their understanding of interaction from the presentation or the way it appears on the screen. Some technical details of web pages and email messages are hidden and some of them are not understandable to most users. Thus, the user does not interpret the system clues or is unable to do so. This misunderstanding enables Phishing and makes it very hard to defend against. Due to the Phishing problem, anti-Phishing approaches

are required to mitigate it. There are anti-Phishing solutions that help in detecting and preventing Phishing attacks. The effectiveness of anti-Phishing approaches is always improved.

The effectiveness of existing online anti-Phishing tips to detect Phishing emails and websites have been evaluated [Kumaraguru et al.07b]. However, this effectiveness research did not consider the effectiveness of each individual tip.

People do not read anti-Phishing online training materials. Thus, Kumaraguru et al. [Kumaraguru et al.07a] considered helping people in detecting Phishing emails during their normal use of emails. However, Kumaraguru et al's approach does not consider helping people to detect Phishing websites. Phishing websites can be reached via various methods in addition to emails such as online advertisements and typing their web addresses in a web browser. Therefore, helping users to make correct decisions in distinguishing Phishing and legitimate websites during their normal use is required.

In the process of designing anti-Phishing approaches, user experiments were conducted to evaluate them. Several approaches were evaluated using participants who were recruited based on their technical abilities [Downs et al.06, Kumaraguru et al.07a, Kumaraguru et al.07b, Sheng et al.07]. Participants were classified into 'experts' and 'non-experts' users based on pre-study screening questions. Participants, who were classified as 'non-experts', were selected to participate in the experiments. Participants who were technically considered non-experts could know about Phishing and how to detect attacks before participating in the evaluation experiments. Having participants with Phishing knowledge in advance may provide biased results in anti-Phishing approaches' evaluation experiments. This is because people who know about Phishing before participating in the evaluation experiments may use their own Phishing knowledge rather than the anti-Phishing approaches that are being evaluated. Downs et al. [Downs et al.07] studied whether there are correlations between some web environment experiences and susceptibility to Phishing. They found that people who correctly answered the knowledge question about the definition of Phishing (i.e. Phishing Aware people) were significantly less likely to fall for Phishing emails. Low technical users may be Phishing Aware and high technical users may be Phishing Unaware. Therefore, an investigation on the effects of technical ability and Phishing knowledge on Phishing websites' detection is required. This would clarify whether

or not the previous screening questions for recruiting low technical users in evaluating anti-Phishing approaches are beneficial.

In this thesis, problems related to the anti-Phishing effectiveness for Phishing websites detection have been addressed. First of all, the effectiveness of the most common users' tips for detecting Phishing websites was evaluated individually. A novel effectiveness criteria was proposed and used to examine each single tip. Then, the tips were ranked accordingly based on an effectiveness score. The research found the most effective anti-Phishing tips that users can focus on to detect and prevent Phishing attacks. The effective tips also can be focused by anti-Phishing training approaches.

Secondly, the investigation that assesses using Phishing knowledge instead of technical ability in the screening questions to recruit participants was presented. User experiments were conducted to evaluate the effects of technical ability and Phishing knowledge. The results of the investigation showed that there is no effect of technical ability on Phishing website detection whereas there is a significant effect of Phishing awareness on Phishing website detection. Thus, recruiting people based on their technical ability without knowing their Phishing knowledge in order to conduct anti-Phishing experiments may produce biased results. Therefore, there is a need to make sure that the participants do not know about Phishing regardless of their technical ability level when they are evaluating the effectiveness of a new anti-Phishing approach.

This thesis also proposed a novel Anti-Phishing Approach that uses Training Intervention for Phishing Websites' Detection (APTIPWD). User experiments were conducted to evaluate the approach. The results showed that New Approach is more effective than the Old Approach of sending anti-Phishing tips by email in helping users distinguish between legitimate and Phishing websites.

This thesis also evaluated the anti-Phishing knowledge retention for users. User experiments were conducted. There were comparisons made between the retention of the users' of the New Approach and the retention of users of the Old approach of sending anti-Phishing tips by email. It was found that users of the Old and the New Approaches retain their anti-Phishing knowledge after 16 days from their first training. Users' decisions during the training are slightly better (i.e. no statistical difference) than their decisions after about

16 days. Additionally, users of the two approaches performed nearly equally in terms of properly judging legitimate and Phishing websites after about 16 days from having the two approaches.

11.2. Criteria for Success

A set of objectives entitled 'criteria for success' was set out in Chapter 1. This section addresses each criterion to find out to what degree the research has succeeded.

1. An evaluation of the anti-Phishing tips' effectiveness for Phishing websites detection.

An examination of the effectiveness of most common users' tips for detecting Phishing websites was presented in Chapter 6. Novel effectiveness criteria were proposed (See Section 6.2.2) and used to examine every tip and to rank it based on its effectiveness score.

It was found that there is no completely effective tip (with an effective score of 1). The most effective tip met three out of the four criteria and it had an effectiveness score of 0.75. It did not meet the criterion four. This is because the tip helps in finding the true URL of a page but it does not help in verifying whether or not the URL is related to a legitimate website. Thus, it possibly produces False Positive (FP) or False Negative (FN) results by using it alone. Using a search engine, such as Google, to verify the URL after using the tip can overcome this weakness.

2. Development of a more effective anti-Phishing approach and its evaluation.

A range of anti-Phishing approaches and their effectiveness have been already developed. This thesis reviewed them and presented them in Chapter 3. The chapter finished with a discussion on the limitations of anti-Phishing approaches shown in Section 3.8.

This thesis proposed a novel Anti-Phishing Approach that uses Training Intervention for Phishing Websites Detection (APTIPWD) described in Chapter 7. The New Approach

presents an intervening message to users who access Phishing websites and try to submit their information. The intervention message is triggered by anti-Phishing blacklists and uses the most effective anti-Phishing tip evaluated in Chapter 6.

By using this approach, users do not need to attend training courses and do not need to access online training materials. This is because the approach brings information to end-users and helps them immediately after they have made a mistake so that they can detect Phishing websites by themselves.

Due to the fact that the blacklists are dynamic and therefore are hard to control, evaluating the New Approach on the real Internet was difficult. A better solution was to evaluate it under experimental conditions. In order to evaluate the approach, all possible scenarios were simulated and described in Section 7.3 and the blacklists (dynamic components) were made fixed (See Section 7.4.2.1).

In order to evaluate the APTIPWD, a hypothesis was identified in Chapter 8 Section 8.2.1. Then, in Chapter 8 user experiments were designed. The recruitment of participants, the effectiveness ratios identified, the considerations on simulating real Phishing attacks and methodology (experiment story board) of the experiment were specified in Sections 8.3., 8.4, 8.5 and 8.6.2 respectively.

In evaluating the hypothesis shown in Section 8.2.1, other hypotheses were extracted and were shown in Section 9.2.1 in Chapter 9. The extracted hypotheses were evaluated and analyzed individually. The New Approach was compared with a control group and the Old Approach of sending anti-Phishing tips to users. The analysis had three aspects. Firstly, there was an assessment of users without taking any of the treatments. Secondly, there was an assessment for using the New Approach in comparison with the Old Approach and the Control group. Thirdly, there was an assessment of each individual group before and after having the treatments. Details for these analyses can be found in Sections 9.2.1.1, 9.2.1.2 and 9.2.1.3 respectively. To sum up, users in the three groups were nearly equal with regards to their decisions about legitimate and Phishing websites before having any treatment. After having the treatments, there was shown to be a significant positive effect of using the New Approach in comparison with the Old Approach. The New Approach was

successful and was better than the Old Approach in helping users properly judging legitimate and Phishing websites.

3. Success to identify factors that influence users decisions against Phishing websites.

The effects of technical ability and Phishing knowledge of users on Phishing websites' detection were discussed and shown in Criterion 3.1 and 3.2 respectively.

3.1. Effect of technical ability on Phishing websites detection.

The effect of the technical ability of users on Phishing websites' detection was discussed. User experiments were designed in Chapter 8 and then used. The effects of technical ability and the results were analyzed and discussed in Chapter 9.

The research hypothesis was identified in Chapter 8 Section 8.2.2.1 and then user experiments were designed. In evaluating the hypothesis shown in Section 8.2.2.1, other hypotheses were extracted and these were shown in Section 9.2.2 in Chapter 9. The extracted hypotheses were evaluated and analyzed individually. The low technical ability (LTA) people were compared with high technical ability (HTA) people on Phishing websites detection. The analysis had three aspects. Firstly, Section 9.2.2.1 presented a detailed assessment of the effect of technical ability level among Phishing Unaware people on Phishing website detection. Secondly, Section 9.2.2.2 presented in detail an assessment of the effect of technical ability level among Phishing Aware people. Thirdly, Section 9.2.2.3 presented an assessment of the effect of technical ability level regardless of Phishing knowledge (Unaware and Aware people). To sum up, it was found that technical ability had no effect on their decisions in Phishing website detection and in recognizing legitimate websites in the three aspects.

3.2. Effect of Phishing knowledge on Phishing websites detection.

The effect of Phishing knowledge for users on Phishing websites' detection was discussed. User experiments were designed in Chapter 8 and then used. The effects of Phishing knowledge on Phishing website detection and the results were analyzed and discussed in Chapter 9.

The research hypothesis was identified in Chapter 8 Section 8.2.2.2 and then user experiments were designed. The experiment methodology (story board) was presented in Section 8.6.4. In evaluating the hypothesis shown in Section 8.2.2.2, other hypotheses were extracted and these were shown in Chapter 9 Section 9.2.3. The extracted hypotheses were then evaluated and analyzed. The Phishing Unaware people were compared with Phishing Aware people on Phishing websites detection. It was found that there was a significant positive effect for the Phishing Aware users on detecting Phishing websites properly. The decisions of Phishing Aware users were better than the decisions of Phishing Unaware users. The conclusion was that Phishing awareness has a significant positive effect on users' decisions in Phishing website detection.

4. *An evaluation of the anti-Phishing knowledge retention when using the New Approach.*

The evaluation of the anti-Phishing knowledge retention by users who use the New Approach compared with the users' of the Old Approach of sending anti-Phishing tips by email was presented. User experiments were designed in Chapter 8 and then used, and the results were analyzed and discussed in Chapter 9.

The research hypothesis was identified in Chapter 8 Section 8.2.3. Two phases of user experiments were conducted to evaluate the retention of the anti-Phishing knowledge. The experiment methodology (story board) was presented in Section 8.6.3. In evaluating the hypothesis shown in Section 8.2.3, other hypotheses were extracted and shown in Section 9.2.4 in Chapter 9. The extracted hypotheses were evaluated and analyzed individually. The users of the New Approach group were compared with the users of the Old Approach group on Phishing websites detection. The analysis had two different aspects. Firstly, there was an assessment of the anti-Phishing knowledge retention in each group individually in both the first and the second phases of the experiments. Secondly, there was an assessment of the retention of the anti-Phishing knowledge between the Old Approach and the New Approach groups in the second phase of the experiments. Details of these analyses can be found in Sections 9.2.4.1 and 9.2.4.2 respectively. To sum up, it was found that users of both approaches retained their anti-Phishing knowledge after 16 days from their first training. Users' decisions during the training are slightly better (i.e. no statistical difference) than their decisions after about 16 days. Additionally, users in the two approaches performed

nearly equally in terms of properly identifying legitimate and Phishing websites after 16 days from experiencing the two approaches.

5. *Comparisons with other related studies.*

The work in this thesis was compared with the relevant work of other researchers in Chapter 10. The chapter presented discussions on the similarities and differences on methodological issues such as the recruitment of participants, groups, scenarios, emails and websites, anti-Phishing tips used and implementation (See Section 10.2). The results' comparison was presented in Section 10.4.

There were comparable issues such as participants' recruitment, effectiveness ratios, scenarios, implementation and training strategies. However, the results of evaluating the New Approach with the related studies were not comparable because the groups in the two studies were different. There are also differences in issues such as participants' knowledge (before participating in the experiments), the tips given to the participants in the experiments, the number of times that the training intervention was given to the participants and the period length between the two phases of experiments.

6. *A proof of concept implementation.*

In Chapter 7, a prototype proof of concept implementation of the Anti-Phishing Approach that uses Training Intervention for Phishing Websites Detection (APTIPWD) was presented. Section 7.4 discussed the design and the implementation of each component of the prototype. It was shown that the New Approach was doable and it could be implemented easily without writing a single line of a programming code and without undue disruption of the users system.

11.3. Future Work

Based on the research in this thesis, a number of possible future work directions can be identified. They are as follows:

1. After finding the most effective anti-Phishing tips for Phishing websites detection, this could be used in developing (or improving the previous) anti-Phishing approaches that are aimed at detecting Phishing websites.
2. The same effectiveness evaluation criteria for Phishing websites detection tips will be carried out on anti-Phishing tips for Phishing emails detection. Then, if the resulting tip is considered effective, it could be used in developing or improving any existing anti-Phishing approach that is aimed at detecting Phishing emails.
3. The evaluation of the effectiveness of the most common anti-Phishing tips for Phishing websites detection carried out in this thesis is subjective. Therefore, an objective evaluation (using user experiments) will be carried out in order to see if the results in the two evaluations change.
4. The possibility of using search engines automatically to verify the credibility and the legitimacy of a URL will be investigated.
5. Due to the fact that promising findings have been achieved regarding the use of the New Approach (APTIPWD), the approach will be implemented and applied to the real Internet using dynamic anti-Phishing blacklists that are updated continuously.
6. Due to the facts that (i) the New Approach used the most effective tip found by a part of this research, (ii) the tip helps to verify the true URL of a page and (iii) the structure of a URL is commonly based on English syntax, the New Approach evaluation experiments will be re-conducted using non-English speakers. This will investigate whether the URL could be verified improperly because of the users' language even if it belongs to a well-known website (False Positive).
7. In case that the URL could be verified improperly because of the users' language even if it belongs to a well-known website (False Positive), the criteria for evaluating the effectiveness of Phishing websites detection tips will be improved and then applied to the tips again to see whether or not changes in the ranking occur.
8. After finding that the previous screening questions for recruiting low technical users in evaluating anti-Phishing approaches are not beneficial, future research will attempt to re-conduct the experiments of previous researchers after recruiting people based on their

Phishing knowledge. The participants of the experiments will be only those who are considered 'Phishing Unaware' regardless of their technical ability level.

9. The definition of technical ability in this thesis is based on three technical skills (See Chapter 8). An investigation will be carried out to identify the factors that might define technical ability in a more accurate manner.
10. The effect of the New Approach on users' anti-Phishing retention in longer term will be evaluated.
11. The effect of multiple treatment sessions (e.g. double and triple) on the users' anti-Phishing retention using the New Approach will be evaluated.
12. All the experiments in the research will be re-conducted using bigger sample sizes.

11.4. Summary

The problems related to the anti-Phishing effectiveness for Phishing websites detection have been addressed in this thesis. First of all, the effectiveness of the most common users' tips for detecting Phishing websites individually was evaluated. Novel effectiveness criteria were proposed and used to examine every tip and to rank it based on its effectiveness score. The research found the most effective anti-Phishing tips that users can focus on to detect and prevent Phishing attacks. The effective tips also can be focused by anti-Phishing training approaches. Secondly, the investigation that used Phishing knowledge instead of technical ability in the screening questions to recruit participants was presented. The results of the investigation showed that there is no effect of technical ability on Phishing websites detection whereas there is a significant positive effect of Phishing awareness on Phishing website detection. Thus, there is a need to make sure that the participants do not know about Phishing regardless of their technical ability level in evaluating the effectiveness of a new anti-Phishing approach. Thirdly, a novel Anti-Phishing Approach that uses Training Intervention for Phishing Websites' Detection (APTIPWD) was proposed and evaluated by conducting user experiments. The results showed that the New Approach is more effective than the Old Approach of sending anti-Phishing tips by email in helping users properly distinguish legitimate and Phishing websites. A prototype proof of concept implementation of the New Approach was presented. It showed that the New Approach was viable and it

could be implemented easily without writing a single line of a programming code and without undue disruption of the users system. Finally, this thesis also evaluated the anti-Phishing knowledge retention of the New Approach's users. There were comparisons between the retention of the users' of the New Approach with the retention of the users of the Old Approach of sending anti-Phishing tips by email. It was found that users of the Old and the New Approaches retain their anti-Phishing knowledge after 16 days from their first training. Users' decisions at the time of using the approaches were slightly better (but with no statistical difference) than their decisions after about 16 days. Additionally, users in the two approaches performed nearly equally in terms of properly identifying legitimate and Phishing websites after about 16 days from experiencing the two approaches.

12. References

[Aladwani01] Aladwani, A., 2001. Online banking: a field study of drivers, development challenges, and expectations. *International Journal of Information Management*. 21 (3), pp. 213-225.

[Alfuraih02] Alfuraih, S. I., Sui, N. T. and and McLeod, D., 2002. Using Trusted Email to Prevent Credit Card Frauds in Multimedia Products. *World Wide Web*. 5 (3), pp. 245-256.

[AlnajimMunro08] Alnajim, A. and Munro, M., 2008. An Evaluation of Users' Tips Effectiveness for Phishing Websites Detection. Proceedings of the third IEEE International Conference on Digital Information Management ICDIM, London, IEEE Press, pp. 63-68.

[AlnajimMunro09a] Alnajim, A. and Munro, M., 2009. Effects of Technical Abilities and Phishing Knowledge on Phishing Websites Detection. Proceedings of the IASTED International Conference on Software Engineering (SE 2009), Innsbruck, Austria, ACTA Press, pp. 120-125.

[AlnajimMunro09b] Alnajim, A. and Munro, M., 2009. An Anti-Phishing Approach that Uses Training Intervention for Phishing Websites Detection. Proceedings of the 6th IEEE International Conference on Information Technology - New Generations (ITNG), Las Vegas, USA, IEEE Press, pp. 405-410.

[AlnajimMunro09c] Alnajim, A. and Munro, M., 2009. An Anti-Phishing Approach for Phishing Websites Detection. In: Pichappan, P., ed. *Handbook of Research on Threat Management and Information Security: Models for Countering Attacks, Breaches and Intrusions*. Pennsylvania USA: IGI Global, (To appear).

[AlnajimMunro09d] Alnajim, A. and Munro, M., 2009. Detecting Phishing Websites: On the Effectiveness of Users' Tips. *Journal of Information Assurance and Security (JIAS)*, ISSN 1554-1010, (To appear).

[AlnajimMunro09e] Alnajim, A. and Munro, M., 2009. An Evaluation of Users' Anti-Phishing Knowledge Retention. Proceedings of the International Conference on 2009 International Conference on Information Management and Engineering (ICIME 2009), Kuala Lumpur, Malaysia, IEEE Press, pp. 210-214.

[Anandpara et al.08] Anandpara, V., Dingman, A., Jakobsson, M., Liu, D., and Roinestad, H., 2008. Phishing IQ tests measure fear, not ability. In: Dietrich, S. and Dhamija, R., eds. *Financial Cryptography and Data Security*. New York, USA: Springer Berlin / Heidelberg, pp. 362-366.

[Anderson93] Anderson, J. R., 1993. Rules of the Mind. New Jersey, USA: Lawrence Erlbaum Associates.

[Anderson et al.96] Anderson, J. R., Reder L. M. and Simon, H. A, 1996. Situated learning and education. *Educational Researcher,* 25 (4), pp. 5–11.

[Anderson et al.00] Andresen, L. Boud, D. and Cohen, R., 2000. Experience-based Learning. *In*: Foley, G., ed. *Understanding Adult Education and Training*. Second Edition. Sydney: Allen & Unwin, pp. 225-239.

[ApacHttp] Apache. Apache HTTP Server Project (online). Available at: http://httpd.apache.org, last access on 1 December 2008.

[ApacHTTPVirtual] Apache, Apache Virtual Host documentation (online). Available at: http://httpd.apache.org/docs/2.0/vhosts, last access on 1 December 2008.

[APWG07a] Anti-Phishing Working Group. EBay- 'eBay Verify Accounts' (online). Available at: http://www.antiphishing.org/phishing_archive/04-18-05_eBay/04-18-05_eBay.html, last access on 26 March 2007.

[APWG07b] Anti-Phishing Working Group. Consumer Advice: How to Avoid Phishing Scams (online). Available at: http://www.antiphishing.org/consumer_recs.html, last access on 26 March 2007.

[APWG07c] Anti-Phishing Working Group Archive. Phishing Archive (online). Available at: http://www.antiphishing.org/phishing_archive/phishing_archive.html, last access on 26 March 2007.

[APWG08] Anti-Phishing Working Group, 2008. Phishing activity trends report (February) (online). Available at: http://www.apwg.com/reports/apwg_report_Q1_2008.pdf, last access on 26 June 2008.

[APACSa] APACS, Spot and Stop Card-Not-Present Fraud Report: Full Pack (online). Available at: http://www.cardwatch.org.uk, last access on 11 November 2006.

[APACSb] APACS, AVS/CSC – Effective CNP Fraud Prevention Report (online). Available at: http://www.cardwatch.org.uk, last access on 11 November 2006.

[APACSc] APACS, Types of card fraud (online), 2005. Available at: http://www.apacs.org.uk/payments_industry/payment_fraud_1_1.html, last access on 10 October 2006.

[CAB06] Citizens Advice Bureau, 2006. Fraud on the Internet (online). Available at: www.adviceguide.org.uk, last access on 19 March 2007.

[Chandrasekaran et al.06] Chandrasekaran, M., Chinchani, R. and Upadhyaya, S., 2006. PHONEY: Mimicking User Response to Detect Phishing Attacks. Proceedings of International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM'06). Washington DC, USA: IEEE Computer Society, pp. 668-672.

[Chou et al.04] Chou, N., Ledesma, R., Teraguchi, Y., Boneh, D. and Mitchell, J. C., 2004. Client Side Defense Against Web-based Identity Theft. Proceedings of 11th Annual

Network and Distributed System Security Symposium (NDSS'04). Available at: http://www.isoc.org/isoc/conferences/ndss/04/proceedings/Papers/Chou.pdf, last access on 14 April 2007.

[Claessens et al.02] Claessens, J., Dem, V., Cock, D. D., Preneel, B. and Vandewalle, J., 2002. On the Security of Today's Online Electronic Banking Systems. *Computers & Security*, 21(3), pp. 253-265.

[ClarkMayer02] Clark, R. C. and Mayer, R. E., 2002. E-Learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning. San Francisco, USA: Pfeiffer.

[Cook02] Cook, N., 2002. Card not present fraud. *Journal of Card Technology Today*, 14(7-8), pp. 11-13.

[Coffield et al.04] Coffield, F., Moseley, D., Hall, E., and Ecclestone, K., 2004. Learning styles and pedagogy in post-16 learning: A systematic and critical review (online). Learning and Skills Research Centre. Available at: http://www.lsda.org.uk/files/PDF/1543.pdf, last access on 3 October 2008.

[Cranor et al.06a] Cranor, L, Egelman, S., Hong, J. and Zhang, Y., 2006. Phinding Phish: an evaluation of anti-phishing toolbars. Technical report CMU-CyLab-06-018, Carnegie Mellon University, USA.

[Cranor et al.06b] Cranor, L. F., 2006. What do they "indicate?": evaluating security and privacy indicators. *Interactions*, SPECIAL ISSUE: HCI and security, 13(3), pp. 45 – 47.

[CyberSource08] CyberSource, 9th Annual Online Fraud Report (online). Edition: 2008. Available at: http://www.cybersource.com.

[CyberSource09] CyberSource, 5th Annual UK Online Fraud Report (online). Edition: 2009. Available at: http://www.cybersource.com.

[DaraGundemoni06] Dara, J. and Gundemoni, L., 2006. Credit Card Security and E-payment: Enquiry into Credit Card Fraud in E-payment. Thesis (Master). Lulea University of Technology, Sweden.

[Dhamija et al.06] Dhamija, R.,Tygar, J. D. and Hearst, M., 2006. Why phishing works. Proceedings of the SIGCHI conference on human factors in computing systems. New York, USA: ACM Press, pp. 581 - 590.

[Dodge et al.07] Dodge, R. C., Carver, C. and Ferguson, A. J., 2007. Phishing for user security awareness. *Computers & Security*, 26 (1), pp. 73-80.

[Downs et al.06] Downs, J. S., Holbrook, M.B. and Cranor, L. F., 2006. Decision strategies and susceptibility to phishing. Proceedings of the second symposium on usable privacy and security. New York, USA: ACM Press, pp. 79 – 90.

References

[Downs et al.07] Downs, J. S., Holbrook, M. and Cranor, L. F., 2007. Behavioral response to phishing risk. Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit. New York, USA: ACM Press, pp. 37 – 44.

[Emigh05] Emigh, A., 2005. Online Identity Theft: Phishing Technology, Chokepoints and Countermeasures (online). Radix Labs. Available at: http://www.antiphishing.org/Phishing-dhs-report.pdf, last access on 23 August 2008.

[Field05] Field, A., 2005. Discovering statistics using SPSS: (and sex, drugs and rock 'n' roll). 2nd ed. London: Sage.

[Friedlander et al.07] Friedlander, A., Mankin, A., Maughan, W. D and Crocker S. D., 2007. DNSSEC: a protocol toward securing the internet infrastructure. *Communications of the ACM*, 50 (2), pp. 44-50.

[GatautisNeverauskas05] Gatautis, R. and Neverauskas, B., 2005. E-commerce adoption in transition economies: SMEs perspectives in Lithuania. Proceedings of the 7th international conference on Electronic commerce. New York, USA: ACM Press, pp. 109-113.

[Getley78] Getley R., 1978. Notes on training methods. *Industrial and Commercial Training*, 10 (7), pp. 280 – 281.

[Harrison88] Harrison R., 1988. Training and Development. London: Institute of Personnel Management.

[Hassler01] Hassler, V., 2001. Security Fundamentals for E-Commerce. London: Artech House.

[HartCrisp91] Hart, L. B. and Crisp, M. G., 1991. Training Methods that Work: A Handbook for Trainers. USA: Thomson Crisp Learning.

[Health e-Tech] Health e-Technologies, Computer/Internet Experience and Skills Questionnaire (online). Available at: http://www.hetinitiative.org/UWashHart_Baseline%20Computer%20Skills%20Survey.pdf, last access on 28 December 2007.

[Hilley05] Hilley, S. (ed), 2005. Online banking - catch 22. *Journal of Computer Fraud & Security*, 2005 (4), pp. 1-2.

[IC3] Internet Crime Complaint Center "IC3" (online). Available at: http://www.ic3.gov, last access on 14 February 2007.

[IDTheft] Identity Theft IFCAG Website, What is identity theft? (online). Available at: http://www.identity-theft.org.uk/what-is-identity-theft.asp, last access 1 November 2008.

[Jackson01] Jackson, S., 2001. Editing computer hardware procedures for multimedia presentation. Proceedings of the 19th annual international conference on computer documentation, ACM Special Interest Group for Design of Communication. New York, USA: ACM Press, pp. 68 – 72.

[Jakobsson07] Jakobsson, M. 2007. The Human Factor in Phishing (online). Privacy & Security of Consumer Information '07. Available at:http://www.informatics.indiana.edu/markus/papers/aci.pdf, last access on 28 March 2007.

[JakobssonMyers] Jakobsson, M. and Myers, S., 2007. Phishing and countermeasures; understanding the increasing problem of electronic identity theft. New Jersey: Wiley.

[JakobssonRatkiewicz06] Jakobsson, M. and Ratkiewicz, J., 2006. Designing ethical phishing experiments: a study of (ROT13) rOnl query features. Proceedings of the 15th international conference on World Wide Web. New York, USA: ACM Press, pp. 513 - 522.

[Jammalamadaka et al.05] Jammalamadaka, R. C., Mehrotra, S. and Venkatasubramanian, N., 2005. Pvault: A Client Server System Providing Mobile Access to Personal Data. Proceedings of the ACM workshop on storage security and survivability, New York, USA: ACM Press, pp. 123 – 129.

[JiaWanlei04] Jia, W. and Zhou, W, 2004. Distributed Network Systems: From Concepts to Implementations. New York: Springer.

[Juang07] Juang, W., 2007. D-cash: A flexible pre-paid e-cash scheme for date-attachment. *Electronic Commerce Research and Applications*, 6 (1), pp. 74-80.

[KenneyReid86] Kenney, J. and Reid, M.A., 1986. Training interventions. London: Institute of Personnel Management.

[Khare99] Khare, R., 1999. Anatomy of a URL (and other Internet-Scale Namespaces, Part1). *IEEE Internet Computing*, (3) 5, pp. 78-81.

[KinnearGray04] Kinnear, P. R. and Gray, C. D., 2004. SPSS 12 Made Simple: Release 12.0. UK: Taylor & Francis.

[Kirkley et al.03] Kirkley, J. R., Kirkley, S. E., Myers, T. E., Lindsay, N. and Singer, M. J., 2003. Problem-based embedded training: An instructional methodology for embedded training using mixed and virtual reality technologies (online). Proceedings of Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Available at: http://www.iforces.org/downloads/problem-based.pdf, last access on 10 October 2008.

[Kolb84] Kolb, D. A., 1984. Experiential learning: Experience as the source of learning and development. New Jersey: Prentice-Hall.

[Kozlowski et al.01] Kozlowski, S. W. J., Toney R. J., Mullins M. E., Weissbein D. A., Brown K. J. and Bell B. S., 2001. Developing adaptability: A theory for the design of integrated-embedded training systems, *Advances in Human Performance and Cognitive Engineering Research*, 2001 (1), pp. 59-123.

[Kumaraguru et al.07a] Kumaraguru, P., Rhee, Y., Acquisti, A., Cranor, L. F., Hong, J. and Nunge, E., 2007. Protecting people from phishing: the design and evaluation of an embedded training email system. Proceedings of the SIGCHI conference on Human factors in computing systems. New York, USA: ACM Press, pp. 905 – 914.

[Kumaraguru et al.07b] Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. and Hong, J., 2007. Teaching jonny not to fall for phish. Technical report CMU-CyLab-07-003, Carnegie Mellon University, USA.

[Kumaraguru et al.07c] Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L. F. and Hong, J., 2007. Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. Proceedings of the anti-phishing working group's 2nd annual eCrime researchers summit. New York, USA: ACM Press, pp. 70 – 81.

[Large06] Large, A., 1996. Computer Animation in an Instructional Environment. *Library & Information Science Research*, 18 (1), pp. 3-23.

[Litan05] Litan, A., 2005. Increased Phishing and Online Attacks Cause Dip in Consumer Confidence (online), Gartner Group. Available at: http://gartner11.gartnerweb.com/DisplayDocument?doc_cd=129146, last access on 20 March 2007.

[Litan09] Litan, A., 2009. The War on Phishing Is Far From Over (online). Gartner Group. Available at: http://www.gartner.com/DisplayDocument?ref=g_search&id=927921, last access on 25 June 2009.

[LogicGr] Logic Group, Card Security Codes and Address Verification product sheet (online). Available at: http://www.the-logic-group.com/Downloads/csc-avs.pdf, last access on 11 November 2006.

[Maktoob] Maktoob mail portal (online). Available at: http://mail.maktoob.com/login.php, last access on 10 August 2008.

[MarshallTompsett05] Marshall, A. M. and Tompsett, B. C., 2005. Identity theft in an online world. *Journal of Computer Law & Security Report*, 21(2), pp. 128-137.

[Mason et al.03] Mason R. L., Gunst R. F. and Hess J. L., 2003. Statistical Design and Analysis of Experiments, with Applications to Engineering and Science. 2nd ed. New Jersey: Wiley-Interscience.

[MasterCard] MasterCard, MasterCard SecureCode: How It Works (online). Available at: http://www.mastercard.com/us/personal/en/cardholderservices/securecode/how_it_works.html, last access on 21 March 2007.

[McKenna05] McKenna, B., April 2005. 2.4 billion lost to hi-tech crime. *Journal of Computer Fraud & Security*, 2005 (4), p. 2.

[Microsofta] Microsoft Corporation, What is the Microsoft Phishing Filter and how does it help protect me? (online). Available at: https://phishingfilter.microsoft.com/faq.aspx, last access on 26 March 2007.

[Microsoftb] Microsoft Corporation, Microsoft Security for Home Computer Users Newsletter (online). Available at: http://www.microsoft.com/protect/secnews/default.mspx, last access on 16 March 2007.

[Montgomery05] Montgomery, D C., 2005. Design and analysis of experiments. 6th ed. New Jersey: Wiley.

[OED] Oxford English Dictionary (online). Available at: http://www.oed.com, last access on 23 December 2008.

[Oppliger00] Oppliger R. 2000. Security Technologies for the World Wide Web. London: Artech House.

[Orgill et al.04] Orgill, G. L., Romney G. W., Bailey M. G. and Orgill, P. M., 2004. The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems. Proceedings of the 5th conference on information technology education. New York, USA: ACM Press, pp.177 – 181.

[PEI08] The APWG Public Education Initiative (PEI). The Phishing Education Landing Page Program (online). Available at: http://education.apwg.org/r/about.html, last access on 24 October 2008.

[Pfleeger95] Pfleeger, S. L., 1995. Experimental design and analysis in software engineering. *Annals of Software Engineering*, 1(1), pp. 219-253.

[Philippsohn01] Philippsohn, S., 2001. Trends In Cybercrime — An Overview Of Current Financial Crimes On The Internet. *Computers & Security*, 20 (1), pp. 53-69.

[Poong et al.06] Poong, Y, Zaman, K. and Talha, M., 2006. E-commerce today and tomorrow: a truly generalized and active framework for the definition of electronic commerce. Proceedings of the 8th international conference on electronic commerce. New York, USA: ACM Press, pp. 553 – 557.

[ReadKleiner96] Read C.W. and Kleiner B.H, 1996. Which training methods are effective? *Management Development Review*, 9 (2), pp. 24-29.

[Reavley05] Reavley, N., 2005. Securing online banking. *Journal of Card Technology Today,* 17(10), pp. 12-13.

[RobilaRagucci06] Robila S. A. and Ragucci, J. W., 2006. Don't be a Phish: Steps in User Education. Proceedings of the 11th annual SIGCSE conference on innovation and technology in computer science education. New York, USA: ACM Press, pp. 237 – 241.

[Rotter80] Rotter, J. B., 1980. Interpersonal Trust, Trustworthiness, and Gullibility. *American Psychologies*, 35 (1), pp. 1-7.

[Santner et al.03] Santner, T. J., Williams, B. J. and Notz, W. I., 2003. The design and analysis of computer experiments. New York: Springer.

[Sheng et al.07] Sheng, S., Magnien, B., Kumaraguru, P., Acquisti, A., Cranor, L. F., Hong, J. and Nunge, E., 2007. Anti-Phishing Phil: the design and evaluation of a game that teaches people not to fall for phish. Proceedings of the 3rd symposium on usable privacy and security SOUPS '07. New York, USA: ACM Press, pp. 88 – 99.

[ShuHsiu02] Shu-Sheng, L.and Hsiu-Mei, H., 2002. How web technology can facilitate learning. *Information systems management.* 19 (1), pp. 56-61.

[Singh05] Singh, M. P., 2005. The Practical Handbook of Internet Computing. USA: Chapman & Hall/CRC Publisher.

[Sprent00] Sprent, P., 2000. Applied Nonparametric Statistical Methods. USA: CRC Press.

[SquMail] SquirrelMail, What is SquirrelMail? (online). Available at: http://www.squirrelmail.org/about, last access on 19 September 2008.

[StothardNicholson01] Stothard C. and Nicholson, R, 2001. Skill Acquisition and Retention in Training: DSTO Support to the Army Ammunition Study (online). DSTO Electronics and Surveillance Research Laboratory, DSTO-CR-0218. Available at: http://dspace.dsto.defence.gov.au/dspace/bitstream/1947/3401/1/DSTO-CR-0218%20PR.pdf, last access on 3 October 2008.

[Symantec04] Symantec, 2004. Mitigating Online Fraud: Customer Confidence, Brand Protection, and Loss Minimization: Report (online). Available at: http://www.antiphishing.org/sponsors_technical_papers/symantec_online_fraud.pdf, last access on 21 March 2007.

[Urdan05] Urdan T. C., 2005. Statistics in Plain English. 2nd ed. USA: Lawrence Erlbaum.

[Visa05] Verified by Visa (VbV), 2005. Merchant Implementation Guide (online). Available at: http://www.visaeurope.com/documents/vbv/verifiedbyvisa_3dsecure.pdf, last access on 4 July 2007.

[Weber09] Weber, T., 2009. Cybercrime threat rising sharply (online). Available at: http://news.bbc.co.uk/1/hi/business/davos/7862549.stm, last access on 2 February 2009.

[Wilson00] Wilson H. C., 2000. Emergency response preparedness: small group training, Part 2. *Disaster Prevention and Management*, 9 (3), pp. 180-199.

[Wu et al.06] Wu, M., Miller, R. C. and Garfinkel, S. L., 2006. Do security toolbars actually prevent phishing attacks? Proceedings of the SIGCHI conference on human factors in computing systems CHI '06. New York, USA: ACM Press, pp. 601 – 610.

[Wu06] Wu, M., 2006. Fighting Phishing at the User Interface. Thesis (PhD). Massachusetts Institute of Technology, USA.

[XiaoChen08] Xiao, Y. and Chen, H., 2008. Mobile Telemedicine: A Computing and Networking Perspective. USA: Auerbach Publications.

[Yamagishi et al.99] Yamagishi T., Kikuchi M. and Kosugi M., 1999. Trust, gullibility, and social intelligence. *Asian Journal of Social Psychology*, 2 (1), pp. 145–161.

[Young07] Young, T., 2007. PayPal acts to stamp out phishing attacks (online). Computing magazine 1 Feb 2007. Available at:

http://www.computing.co.uk/computing/news/2173907/paypal-stamp-phishing-attacks, last access on 27 March 2007.

[Zachary et al.99] Zachary, W., Cannon-Bowers, J., Bilazarian, P., Krecker, D., Lardieri, P. and Burns, J., 1999. The Advanced Embedded Training System (AETS): An Intelligent Embedded Tutoring System for Tactical Team Training. *International Journal of Artificial Intelligence in Education*, (1999) 10, pp. 257-277.

[Zhang et al.07] Zhang, Y., Hong, J. I. and Cranor, L. F., 2007. Cantina: a content-based approach to detecting phishing web sites. Proceedings of the 16th international conference on World Wide Web. New York, USA: ACM Press, pp. 639 – 648.

Appendix A

The appendix presents the full table (21 tips) of the evaluation of the users' tips effectiveness for Phishing websites detection.

#	Tip	Criteria				TE	TR
		1	2	3	4		
		v=0.25	v=0.25	v=0.25	v=0.25		
1	Do not use links to access a site.	N	N	NA	N	0	>6
2	Type in your browser the address of the site you intend to go or use a bookmark that you previously created.	N	Y	NA	Y	0.5	2
3	Do not give your personal contact or account information to a website that looks suspicious.	N	N	NA	N	0	>6
4	Make sure you are on a secure connection when entering sensitive information. Secure Web pages will have the text https: instead of http:	Y (73.8)	N	N	N	0.25	5
5	Click on the padlock to check that the seller is who they say they are and that their certificate is current and registered to the right address.	N	N	N	N	0	>6
6	Do not be fooled by a padlock that appears on the web page itself. It's easy for conmen to copy the image of a padlock. Look for one that is in the window frame of the browser itself.	N	N	N	N	0	>6
7	Before entering card details, look for MasterCard SecureCodeTM sign as an endorsement of retailers security.	N	N	N	N	0	>6
8	Before entering card details, look for the VeriSign Secured™ Seal.	N	N	N	N	0	>6
9	Use sites that carry the TrustUK logo.	N	N	N	N	0	>6
10	Look beyond the logo and do not give out your information before you check the privacy and security seals. Scammers often include actual logos and images of legitimate companies.	N	Y	Y	N	0.5	~2
11	A fake website may have this caristaritic: The website's address is different from what you are used to, perhaps there are extra characters or words in it or it uses a completely different name or no name at all, just numbers. Check the address in your browser's address bar after you arrive at a website.	Y (71.4)	N	N	N	0.25	6

12	Even though you are asked to enter private information there is NO padlock in the browser window or 'https://' at the beginning of the web address to signify that it is using a secure link and that the site is what it says it is.	Y (90.4)	N	N	N	0.25	4
13	A fake website may have this caristaritic: A request for personal information such as user name, password or other security details IN FULL, when you are normally only asked for SOME of them.	N	N	NA	N	0	> 6
14	In the case of spotting dodgy sites: Use your instincts and commonsense. If it smells bad, it's probably rotten.	N	N	NA	N	0	> 6
15	In the case of spotting dodgy sites: Avoid sites that hype investments, whether in shares or alleged rarities like old wine, whisky or property. Do your homework and always get professional advice before making investment decisions.	N	N	NA	N	0	> 6
16	In the case of spotting dodgy sites: Be wary of sites that promise easy profits.	N	N	NA	N	0	> 6
17	In the case of spotting dodgy sites: Do a web search to see if anyone has had any problems with a suspicious-looking website.	N	N	NA	N	0	> 6
18	In the case of spotting dodgy sites: Be wary of websites that are advertised in unsolicited emails from strangers.	N	N	NA	N	0	> 6
19	A fake website may have this characteristic: The website's address is different from what you are used to, perhaps there are extra characters or words in it or it uses a completely different name or no name at all, just numbers. Check the True URL. The true URL of the site can be seen in the page 'Properties'.	Y (100)	Y	Y	N	0.75	1
20	Read about phishing, social engineering, e-commerce fraud and identity theft. Much of this advice to individuals also applies to businesses.	N	N	NA	N	0	> 6
21	Be suspicious of deals that seem too good to be true. They usually are.	N	N	NA	N	0	> 6

Appendix B

The appendix presents the Scenario Information sheets given to participants in the evaluation experiments in order to explain to them the experiment simple scenarios. The Scenario Information sheet (1) was used in the first experiment whereas Scenario Information sheet (2) was used in the second and third experiments. They are very similar to each other. This appendix presents also the pre-study and pre-session surveys used.

Scenario Information Sheet (1)

THE USE OF WWW: HOW PEOPLE EFFECTIVELY MANAGE AND USE THE INTERNET AND EMAILS

Thank you for volunteering to participate in this study. You will participate in a short interaction with emails and websites as you do in your normal surfing. You will be playing the role of 'Dave Smith'. Dave Smith is an imaginary person who works in the marketing department in an IT company. Therefore, please follow the following scenario:

1. Imagine that you are Dave Smith.
2. OK Dave, check your email and deal with the emails in your inbox as you do usually.
3. Deal with emails in order starting from the top one.
4. All the IDs and passwords for you Dave to use in this study are written down below in this sheet.

That's it!

Dave Smith's IDs and passwords

PayPal
Email Address: davesmith2001@hotmail.com
PayPal Password: car1000

Amazon
Email Address: davesmith2001@hotmail.com
Password: car2001

eBay
User ID: dave88
Password: car555

Barklays
Surname: smith
Membership number: 20-1281554577

Lloyds

User ID: dave88

Password: car333

Halifax

Username: davesmith2001

Password: car1000

Citibank

Username: davesmith2001

Password: car111

Scenario Information Sheet (2)

THE USE OF WWW: HOW PEOPLE EFFECTIVELY MANAGE AND USE THE INTERNET AND EMAILS

Thank you for volunteering to participate in this study. You will participate in a short interaction with emails and websites as you do in your normal surfing. You will be playing the role of 'Dave Smith'. Dave Smith is an imaginary person who works in the marketing department in an IT company. Therefore, please follow the following scenario:

1. Imagine that you are Dave Smith.
2. OK Dave, check your email and deal with the emails in your inbox as you do usually.
3. Deal with emails in order starting from the top one.
4. All the IDs and passwords for you Dave to use in this study are written down below in this sheet.

That's it!

Dave Smith's IDs and passwords

Argos

Login Name: davesmith2001

Password: car1000

Comet

Email Address: davesmith2001@hotmail.com

Password: car1000

Capital One

Username: davesmith2001

Password: car555

Egg Online Bank

Account Number: 05/09/1980

Postcode: DH1 3LE

Mother's Maiden Name: Masary

Password: car2001

Abbey Bank

Personal ID: 23458679876

Passcode: car2001

Registration number: 20012

Co-operative Bank

Sort Code: 466787

Account Number: 23736892

PIN: car333

Natwest Bank

Customer Number: 0509807638963

Password: car111

Postcode: DH1 3LE

Pre-study Survey

THE USE OF WWW: HOW PEOPLE EFFECTIVELY MANAGE AND USE THE INTERNET AND EMAILS

A. Your Contacts

1. Your Name: *

2. Your E-mail: *

B. The Internet and e-mail Usage

2. How would you rate your current e-mail skills? *

- ◉ Very poor.
- ☐ Poor.
- ☐ Fair.
- ☐ Good.
- ☐ Very Good

4. Have you ever purchased anything on the web before? *

- ☐ Yes.
- ☐ No.

5. Have you had an active account with PayPal? *

- ☐ Yes.
- ☐ No.

6. Have you used the Internet to access your bank account? *

- ☐ Yes.
- ☐ No.

7. Have you ever used eBay to either purchase or sell anything? *

- ☐ Yes.
- ☐ No.

8. Have you changed preferences or settings in your web browser? *

- ☐ Yes.
- ☐ No.

9. Have you created a web page? *

- ☐ Yes.
- ☐ No.

10. Have you helped someone fix a computer problem? (e.g. software problem, browser problem, etc.) *

- ☐ Yes.
- ☐ No.

Look at to this image and please answer the related following two questions (12 & 13)?

12. Have you seen "this lock image" before? *

- ☐ Yes.
- ☐ No.

13. What does the lock image mean about a web site? *

- ☐ It means that you need a key or a password to enter the site.
- ☐ It means that a website is trustworthy.
- ☐ It means that any information you enter will be sent securely.
- ☐ It means that any information being displayed will be sent securely.

14. Please match the each term to its definition? *

	Something that protects your computer from unauthorized communication outside the network.	Something that watches your computer and sends that information over the Internet.	An Internet search engine	Something websites put on your computer so you do not have to type in the same information the next time you visit.	Something put on your computer without your permission, that changes the way your computer works.	An instant messaging client	Email or website trying to trick you into giving your sensitive information to thieves.	Email trying to sell you something.	I have seen this word before but I do not know what it means for computers.	I have never seen this word before
Virus	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Spyware	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Phishing	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Cookie	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Messenger	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Google	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Pre-session Survey

The same as the questions 8,9,10 and 14 in the pre-study survey were given to participants

Appendix C

This appendix shows the statistical methods and the results used for testing all hypotheses discussed in Chapter 9. A glossary is provided by the end of this appendix to explain the technical terms used in tables.

Hypothesis 1.1 [Kruskal-Wallis Test]

Ranks

	Group	N	Mean Rank
Pre_Correct_Rate	Control Group	12	19.46
	Old Approach Group	12	16.58
	New Approach Group	12	19.46
	Total	36	

Test Statistics(a,b)

	Pre_Correct_Rate
Chi-Square	2.593
df	2
Asymp. Sig.	.273
Exact Sig.	.464
Point Probability	.242

a Kruskal Wallis Test
b Grouping Variable: Group

Hypothesis 1.1 (Control vs. Old) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Pre_Correct_Rate	Control Group	12	13.46	161.50
	Old Approach Group	12	11.54	138.50
	Total	24		

Test Statistics(b)

	Pre_Correct_Rate
Mann-Whitney U	60.500
Wilcoxon W	138.500
Z	-1.384
Asymp. Sig. (2-tailed)	.166
Exact Sig. [2*(1-tailed Sig.)]	.514(a)
Exact Sig. (2-tailed)	.522
Exact Sig. (1-tailed)	.261
Point Probability	.261

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.1 (Control vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Pre_Correct_Rate	Control Group	12	12.50	150.00
	New Approach Group	12	12.50	150.00
	Total	24		

Test Statistics(b)

	Pre_Correct_Rate
Mann-Whitney U	72.000
Wilcoxon W	150.000
Z	.000
Asymp. Sig. (2-tailed)	1.000
Exact Sig. [2*(1-tailed Sig.)]	1.000(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.761
Point Probability	.522

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.1 (Old vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Pre_Correct_Rate	Old Approach Group	12	11.54	138.50
	New Approach Group	12	13.46	161.50
	Total	24		

Test Statistics(b)

	Pre_Correct_Rate
Mann-Whitney U	60.500
Wilcoxon W	138.500
Z	-1.384
Asymp. Sig. (2-tailed)	.166
Exact Sig. [2*(1-tailed Sig.)]	.514(a)
Exact Sig. (2-tailed)	.522
Exact Sig. (1-tailed)	.261
Point Probability	.261

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.2 [Kruskal-Wallis Test]

Ranks

	Group	N	Mean Rank
Pre_FPR	Control Group	12	17.00
	Old Approach Group	12	21.50
	New Approach Group	12	17.00
	Total	36	

Test Statistics(a,b)

	Pre_FPR
Chi-Square	3.500
df	2
Asymp. Sig.	.174
Exact Sig.	.316
Point Probability	.184

a Kruskal Wallis Test
b Grouping Variable: Group

Hypothesis 1.2 (Control vs. Old) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Pre_FPR	Control Group	12	11.00	132.00
	Old Approach Group	12	14.00	168.00
	Total	24		

Test Statistics(b)

	Pre_FPR
Mann-Whitney U	54.000
Wilcoxon W	132.000
Z	-1.476
Asymp. Sig. (2-tailed)	.140
Exact Sig. [2*(1-tailed Sig.)]	.319(a)
Exact Sig. (2-tailed)	.317
Exact Sig. (1-tailed)	.158
Point Probability	.140

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.2 (Control vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Pre_FPR	Control Group	12	12.50	150.00
	New Approach Group	12	12.50	150.00
	Total	24		

Test Statistics(b)

	Pre_FPR
Mann-Whitney U	72.000
Wilcoxon W	150.000
Z	.000
Asymp. Sig. (2-tailed)	1.000
Exact Sig. [2*(1-tailed Sig.)]	1.000(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.761
Point Probability	.522

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.2 (Old vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Pre_FPR	Old Approach Group	12	14.00	168.00
	New Approach Group	12	11.00	132.00
	Total	24		

Test Statistics(b)

	Pre_FPR
Mann-Whitney U	54.000
Wilcoxon W	132.000
Z	-1.476
Asymp. Sig. (2-tailed)	.140
Exact Sig. [2*(1-tailed Sig.)]	.319(a)
Exact Sig. (2-tailed)	.317
Exact Sig. (1-tailed)	.158
Point Probability	.140

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.3 [Kruskal-Wallis Test]

Ranks

	Group	N	Mean Rank
Pre_FNR	Control Group	12	19.00
	Old Approach Group	12	17.50
	New Approach Group	12	19.00
	Total	36	

Test Statistics(a,b)

	Pre_FNR
Chi-Square	.345
df	2
Asymp. Sig.	.842
Exact Sig.	1.000
Point Probability	.344

a Kruskal Wallis Test
b Grouping Variable: Group

Hypothesis 1.3 (Control vs. Old) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Pre_FNR	Control Group	12	13.00	156.00
	Old Approach Group	12	12.00	144.00
	Total	24		

Test Statistics(b)

	Pre_FNR
Mann-Whitney U	66.000
Wilcoxon W	144.000
Z	-.492
Asymp. Sig. (2-tailed)	.623
Exact Sig. [2*(1-tailed Sig.)]	.755(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.342

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.3 (Control vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Pre_FNR	Control Group	12	12.50	150.00
	New Approach Group	12	12.50	150.00
	Total	24		

Test Statistics(b)

	Pre_FNR
Mann-Whitney U	72.000
Wilcoxon W	150.000
Z	.000
Asymp. Sig. (2-tailed)	1.000
Exact Sig. [2*(1-tailed Sig.)]	1.000(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.705
Point Probability	.410

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.3 (Old vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Pre_FNR	Old Approach Group	12	12.00	144.00
	New Approach Group	12	13.00	156.00
	Total	24		

Test Statistics(b)

	Pre_FNR
Mann-Whitney U	66.000
Wilcoxon W	144.000
Z	-.492
Asymp. Sig. (2-tailed)	.623
Exact Sig. [2*(1-tailed Sig.)]	.755(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.342

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.4 [Kruskal-Wallis Test]

Ranks

	Group	N	Mean Rank
Post_Correct_Rate	Control Group	12	14.13
	Old Approach Group	12	14.50
	New Approach Group	12	26.88
	Total	36	

Test Statistics(a,b)

	Post_Correct_Rate
Chi-Square	13.591
df	2
Asymp. Sig.	.001
Exact Sig.	.001
Point Probability	.000

a Kruskal Wallis Test
b Grouping Variable: Group

Hypothesis 1.4 (Control vs. Old) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_Correct_Rate	Control Group	12	12.46	149.50
	Old Approach Group	12	12.54	150.50
	Total	24		

Test Statistics (b)

	Post_Correct_Rate
Mann-Whitney U	71.500
Wilcoxon W	149.500
Z	-.035
Asymp. Sig. (2-tailed)	.972
Exact Sig. [2*(1-tailed Sig.)]	.977(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.127

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.4 (Control vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_Correct_Rate	Control Group	12	8.17	98.00
	New Approach Group	12	16.83	202.00
	Total	24		

Test Statistics(b)

	Post_Correct_Rate
Mann-Whitney U	20.000
Wilcoxon W	98.000
Z	-3.256
Asymp. Sig. (2-tailed)	.001
Exact Sig. [2*(1-tailed Sig.)]	.002(a)
Exact Sig. (2-tailed)	.002
Exact Sig. (1-tailed)	.001
Point Probability	.001

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.4 (Old vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_Correct_Rate	Old Approach Group	12	8.46	101.50
	New Approach Group	12	16.54	198.50
	Total	24		

Test Statistics(b)

	Post_Correct_Rate
Mann-Whitney U	23.500
Wilcoxon W	101.500
Z	-2.995
Asymp. Sig. (2-tailed)	.003
Exact Sig. [2*(1-tailed Sig.)]	.004(a)
Exact Sig. (2-tailed)	.004
Exact Sig. (1-tailed)	.002
Point Probability	.002

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.5 [Kruskal-Wallis Test]

Ranks

	Group	N	Mean Rank
Post_FPR	Control Group	12	17.00
	Old Approach Group	12	21.50
	New Approach Group	12	17.00
	Total	36	

Test Statistics(a,b)

	Post_FPR
Chi-Square	2.188
df	2
Asymp. Sig.	.335
Exact Sig.	.490
Point Probability	.206

a Kruskal Wallis Test
b Grouping Variable: Group

Hypothesis 1.5 (Control vs. Old) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_FPR	Control Group	12	11.00	132.00
	Old Approach Group	12	14.00	168.00
	Total	24		

Test Statistics(b)

	Post_FPR
Mann-Whitney U	54.000
Wilcoxon W	132.000
Z	-1.238
Asymp. Sig. (2-tailed)	.216
Exact Sig. [2*(1-tailed Sig.)]	.319(a)
Exact Sig. (2-tailed)	.400
Exact Sig. (1-tailed)	.200
Point Probability	.155

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.5 (Control vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_FPR	Control Group	12	12.50	150.00
	New Approach Group	12	12.50	150.00
	Total	24		

Test Statistics(b)

	Post_FPR
Mann-Whitney U	72.000
Wilcoxon W	150.000
Z	.000
Asymp. Sig. (2-tailed)	1.000
Exact Sig. [2*(1-tailed Sig.)]	1.000(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.680
Point Probability	.360

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.5 (Old vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_FPR	Old Approach Group	12	14.00	168.00
	New Approach Group	12	11.00	132.00
	Total	24		

Test Statistics(b)

	Post_FPR
Mann-Whitney U	54.000
Wilcoxon W	132.000
Z	-1.238
Asymp. Sig. (2-tailed)	.216
Exact Sig. [2*(1-tailed Sig.)]	.319(a)
Exact Sig. (2-tailed)	.400
Exact Sig. (1-tailed)	.200
Point Probability	.155

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.6 [Kruskal-Wallis Test]

Ranks

	Group	N	Mean Rank
Post_FNR	Control Group	12	24.25
	Old Approach Group	12	19.96
	New Approach Group	12	11.29
	Total	36	

Test Statistics(a,b)

	Post_FNR
Chi-Square	11.245
df	2
Asymp. Sig.	.004
Exact Sig.	.002
Point Probability	.000

a Kruskal Wallis Test

b Grouping Variable: Group

Hypothesis 1.6 (Control vs. Old) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_FNR	Control Group	12	14.21	170.50
	Old Approach Group	12	10.79	129.50
	Total	24		

Test Statistics(b)

	Post_FNR
Mann-Whitney U	51.500
Wilcoxon W	129.500
Z	-1.367
Asymp. Sig. (2-tailed)	.172
Exact Sig. [2*(1-tailed Sig.)]	.242(a)
Exact Sig. (2-tailed)	.214
Exact Sig. (1-tailed)	.107
Point Probability	.027

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.6 (Control vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_FNR	Control Group	12	16.54	198.50
	New Approach Group	12	8.46	101.50
	Total	24		

Test Statistics(b)

	Post_FNR
Mann-Whitney U	23.500
Wilcoxon W	101.500
Z	-3.007
Asymp. Sig. (2-tailed)	.003
Exact Sig. [2*(1-tailed Sig.)]	.004(a)
Exact Sig. (2-tailed)	.002
Exact Sig. (1-tailed)	.001
Point Probability	.001

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.6 (Old vs. New) [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_FNR	Old Approach Group	12	15.67	188.00
	New Approach Group	12	9.33	112.00
	Total	24		

Test Statistics(b)

	Post_FNR
Mann-Whitney U	34.000
Wilcoxon W	112.000
Z	-2.480
Asymp. Sig. (2-tailed)	.013
Exact Sig. [2*(1-tailed Sig.)]	.028(a)
Exact Sig. (2-tailed)	.023
Exact Sig. (1-tailed)	.012
Point Probability	.008

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 1.7 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Pre_Correct_Rate - Post_Correct_Rate	Negative Ranks	2(a)	2.50	5.00
	Positive Ranks	2(b)	2.50	5.00
	Ties	8(c)		
	Total	12		

a Pre_Correct_Rate < Post_Correct_Rate
b Pre_Correct_Rate > Post_Correct_Rate
c Pre_Correct_Rate = Post_Correct_Rate

Test Statistics(b)

	Pre_Correct_Rate - Post_Correct_Rate
Z	.000(a)
Asymp. Sig. (2-tailed)	1.000
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.688
Point Probability	.375

a The sum of negative ranks equals the sum of positive ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 1.8 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Pre_FPR - Post_FPR	Negative Ranks	2(a)	1.50	3.00
	Positive Ranks	0(b)	.00	.00
	Ties	10(c)		
	Total	12		

a Pre_FPR < Post_FPR
b Pre_FPR > Post_FPR
c Pre_FPR = Post_FPR

Test Statistics(b)

	Pre_FPR - Post_FPR
Z	-1.414(a)
Asymp. Sig. (2-tailed)	.157
Exact Sig. (2-tailed)	.500
Exact Sig. (1-tailed)	.250
Point Probability	.250

a Based on positive ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 1.9 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Pre_FNR - Post_FNR	Negative Ranks	0(a)	.00	.00
	Positive Ranks	2(b)	1.50	3.00
	Ties	10(c)		
	Total	12		

a Pre_FNR < Post_FNR
b Pre_FNR > Post_FNR
c Pre_FNR = Post_FNR

Test Statistics(b)

	Pre_FNR - Post_FNR
Z	-1.414(a)
Asymp. Sig. (2-tailed)	.157
Exact Sig. (2-tailed)	.500
Exact Sig. (1-tailed)	.250
Point Probability	.250

a Based on negative ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 1.10 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Pre_Correct_Rate - Post_Correct_Rate	Negative Ranks	3(a)	2.67	8.00
	Positive Ranks	1(b)	2.00	2.00
	Ties	8(c)		
	Total	12		

a Pre_Correct_Rate < Post_Correct_Rate
b Pre_Correct_Rate > Post_Correct_Rate
c Pre_Correct_Rate = Post_Correct_Rate

Test Statistics(b)

	Pre_Correct_Rate - Post_Correct_Rate
Z	-1.134(a)
Asymp. Sig. (2-tailed)	.257
Exact Sig. (2-tailed)	.500
Exact Sig. (1-tailed)	.250
Point Probability	.188

a Based on positive ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 1.11 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Pre_FPR - Post_FPR	Negative Ranks	4(a)	3.50	14.00
	Positive Ranks	2(b)	3.50	7.00
	Ties	6(c)		
	Total	12		

a Pre_FPR < Post_FPR
b Pre_FPR > Post_FPR
c Pre_FPR = Post_FPR

Test Statistics(b)

	Pre_FPR - Post_FPR
Z	-.816(a)
Asymp. Sig. (2-tailed)	.414
Exact Sig. (2-tailed)	.688
Exact Sig. (1-tailed)	.344
Point Probability	.234

a Based on positive ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 1.12 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Pre_FNR - Post_FNR	Negative Ranks	0(a)	.00	.00
	Positive Ranks	4(b)	2.50	10.00
	Ties	8(c)		
	Total	12		

a Pre_FNR < Post_FNR
b Pre_FNR > Post_FNR
c Pre_FNR = Post_FNR

Test Statistics(b)

	Pre_FNR - Post_FNR
Z	-1.890(a)
Asymp. Sig. (2-tailed)	.059
Exact Sig. (2-tailed)	.125
Exact Sig. (1-tailed)	.063
Point Probability	.063

a Based on negative ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 1.13 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Pre_Correct_Rate - Post_Correct_Rate	Negative Ranks	9(a)	5.00	45.00
	Positive Ranks	0(b)	.00	.00
	Ties	3(c)		
	Total	12		

a Pre_Correct_Rate < Post_Correct_Rate

b Pre_Correct_Rate > Post_Correct_Rate
c Pre_Correct_Rate = Post_Correct_Rate

Test Statistics(b)

	Pre_Correct_Rate - Post_Correct_Rate
Z	-2.762(a)
Asymp. Sig. (2-tailed)	.006
Exact Sig. (2-tailed)	.004
Exact Sig. (1-tailed)	.002
Point Probability	.002

a Based on positive ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 1.14 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Pre_FPR - Post_FPR	Negative Ranks	3(a)	2.50	7.50
	Positive Ranks	1(b)	2.50	2.50
	Ties	8(c)		
	Total	12		

a Pre_FPR < Post_FPR
b Pre_FPR > Post_FPR
c Pre_FPR = Post_FPR

Test Statistics(b)

	Pre_FPR - Post_FPR
Z	-1.000(a)
Asymp. Sig. (2-tailed)	.317
Exact Sig. (2-tailed)	.625
Exact Sig. (1-tailed)	.313
Point Probability	.250

a Based on positive ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 1.15 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Pre_FNR - Post_FNR	Negative Ranks	0(a)	.00	.00
	Positive Ranks	10(b)	5.50	55.00
	Ties	2(c)		
	Total	12		

a Pre_FNR < Post_FNR
b Pre_FNR > Post_FNR
c Pre_FNR = Post_FNR

Test Statistics(b)

	Pre_FNR - Post_FNR
Z	-2.889(a)
Asymp. Sig. (2-tailed)	.004
Exact Sig. (2-tailed)	.002
Exact Sig. (1-tailed)	.001
Point Probability	.001

a Based on negative ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 2.1 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Total_Correct_Rate	Low	6	5.67	34.00
	High	6	7.33	44.00
	Total	12		

Test Statistics(b)

	Total_Correct_Rate
Mann-Whitney U	13.000
Wilcoxon W	34.000
Z	-.962
Asymp. Sig. (2-tailed)	.336
Exact Sig. [2*(1-tailed Sig.)]	.485(a)
Exact Sig. (2-tailed)	.636
Exact Sig. (1-tailed)	.318
Point Probability	.227

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.2 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Total_FPR	Low	6	6.83	41.00
	High	6	6.17	37.00
	Total	12		

Test Statistics(b)

	Total_FPR
Mann-Whitney U	16.000
Wilcoxon W	37.000
Z	-.422
Asymp. Sig. (2-tailed)	.673
Exact Sig. [2*(1-tailed Sig.)]	.818(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.136

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.3 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Total_FNR	Low	6	7.00	42.00
	High	6	6.00	36.00
	Total	12		

Test Statistics(b)

	Total_FNR
Mann-Whitney U	15.000
Wilcoxon W	36.000
Z	-.631
Asymp. Sig. (2-tailed)	.528
Exact Sig. [2*(1-tailed Sig.)]	.699(a)
Exact Sig. (2-tailed)	.727
Exact Sig. (1-tailed)	.364
Point Probability	.136

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.4 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_Correct_Rate	Low	6	7.00	42.00
	High	6	6.00	36.00
	Total	12		

Test Statistics(b)

	Pre_Correct_Rate
Mann-Whitney U	15.000
Wilcoxon W	36.000
Z	-1.000
Asymp. Sig. (2-tailed)	.317
Exact Sig. [2*(1-tailed Sig.)]	.699(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.500

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.5 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_FPR	Low	6	5.50	33.00
	High	6	7.50	45.00
	Total	12		

Test Statistics(b)

	Pre_FPR
Mann-Whitney U	12.000
Wilcoxon W	33.000
Z	-1.173
Asymp. Sig. (2-tailed)	.241
Exact Sig. [2*(1-tailed Sig.)]	.394(a)
Exact Sig. (2-tailed)	.545
Exact Sig. (1-tailed)	.273
Point Probability	.242

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.6 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_FNR	Low	6	7.00	42.00
	High	6	6.00	36.00
	Total	12		

Test Statistics(b)

	Pre_FNR
Mann-Whitney U	15.000
Wilcoxon W	36.000
Z	-.638
Asymp. Sig. (2-tailed)	.523
Exact Sig. [2*(1-tailed Sig.)]	.699(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.409

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.7 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_Correct_Rate	Low	6	7.00	42.00
	High	6	6.00	36.00
	Total	12		

Test Statistics(b)

	Pre_Correct_Rate
Mann-Whitney U	15.000
Wilcoxon W	36.000
Z	-1.000
Asymp. Sig. (2-tailed)	.317
Exact Sig. [2*(1-tailed Sig.)]	.699(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.500

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.8 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_FPR	Low	6	7.00	42.00
	High	6	6.00	36.00
	Total	12		

Test Statistics(b)

	Pre_FPR
Mann-Whitney U	15.000
Wilcoxon W	36.000
Z	-1.000
Asymp. Sig. (2-tailed)	.317
Exact Sig. [2*(1-tailed Sig.)]	.699(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.500

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.9 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_FNR	Low	6	5.50	33.00
	High	6	7.50	45.00
	Total	12		

Test Statistics(b)

	Pre_FNR
Mann-Whitney U	12.000
Wilcoxon W	33.000
Z	-1.483
Asymp. Sig. (2-tailed)	.138
Exact Sig. [2*(1-tailed Sig.)]	.394(a)
Exact Sig. (2-tailed)	.455
Exact Sig. (1-tailed)	.227
Point Probability	.227

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.10 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_Correct_Rate	Low	18	19.94	359.00
	High	18	17.06	307.00
	Total	36		

Test Statistics(b)

	Pre_Correct_Rate
Mann-Whitney U	136.000
Wilcoxon W	307.000
Z	-1.716
Asymp. Sig. (2-tailed)	.086
Exact Sig. [2*(1-tailed Sig.)]	.424(a)
Exact Sig. (2-tailed)	.257
Exact Sig. (1-tailed)	.129
Point Probability	.129

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.11 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_FPR	Low	18	17.50	315.00
	High	18	19.50	351.00
	Total	36		

Test Statistics(b)

	Pre_FPR
Mann-Whitney U	144.000
Wilcoxon W	315.000
Z	-.882
Asymp. Sig. (2-tailed)	.378
Exact Sig. [2*(1-tailed Sig.)]	.584(a)
Exact Sig. (2-tailed)	.658
Exact Sig. (1-tailed)	.329
Point Probability	.240

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.12 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_FNR	Low	18	18.00	324.00
	High	18	19.00	342.00
	Total	36		

Test Statistics(b)

	Pre_FNR
Mann-Whitney U	153.000
Wilcoxon W	324.000
Z	-.415
Asymp. Sig. (2-tailed)	.678
Exact Sig. [2*(1-tailed Sig.)]	.791(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.299

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.13 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Toral_Correct_Rate	Low	6	6.92	41.50
	High	6	6.08	36.50
	Total	12		

Test Statistics(b)

	Toral_Correct_Rate
Mann-Whitney U	15.500
Wilcoxon W	36.500
Z	-.527
Asymp. Sig. (2-tailed)	.598
Exact Sig. [2*(1-tailed Sig.)]	.699(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.273

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.14 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Total_FPR	Low	6	5.67	34.00
	High	6	7.33	44.00
	Total	12		

Test Statistics(b)

	Total_FPR
Mann-Whitney U	13.000
Wilcoxon W	34.000
Z	-.874
Asymp. Sig. (2-tailed)	.382
Exact Sig. [2*(1-tailed Sig.)]	.485(a)
Exact Sig. (2-tailed)	.494
Exact Sig. (1-tailed)	.247
Point Probability	.130

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.15 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Total_FNR	Low	6	6.58	39.50
	High	6	6.42	38.50
	Total	12		

Test Statistics(b)

	Total_FNR
Mann-Whitney U	17.500
Wilcoxon W	38.500
Z	-.086
Asymp. Sig. (2-tailed)	.932
Exact Sig. [2*(1-tailed Sig.)]	.937(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.100

a Not corrected for ties.

b Grouping Variable: Level of Technical Ability

Hypothesis 2.16 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_Correct_Rate	Low	24	25.85	620.50
	High	24	23.15	555.50
	Total	48		

Test Statistics(a)

	Pre_Correct_Rate
Mann-Whitney U	255.500
Wilcoxon W	555.500
Z	-.867
Asymp. Sig. (2-tailed)	.386
Exact Sig. (2-tailed)	.541
Exact Sig. (1-tailed)	.270
Point Probability	.111

a Grouping Variable: Level of Technical Ability

Hypothesis 2.17 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_FPR	Low	24	23.00	552.00
	High	24	26.00	624.00
	Total	48		

Test Statistics(a)

	Pre_FPR
Mann-Whitney U	252.000
Wilcoxon W	552.000
Z	-1.019
Asymp. Sig. (2-tailed)	.308
Exact Sig. (2-tailed)	.494
Exact Sig. (1-tailed)	.247
Point Probability	.163

a Grouping Variable: Level of Technical Ability

Hypothesis 2.18 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Pre_FNR	Low	24	24.50	588.00
	High	24	24.50	588.00
	Total	48		

Test Statistics(a)

	Pre_FNR
Mann-Whitney U	288.000
Wilcoxon W	588.000
Z	.000
Asymp. Sig. (2-tailed)	1.000
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.550
Point Probability	.099

a Grouping Variable: Level of Technical Ability

Hypothesis 2.19 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Toral_Correct_Rate	Low	12	12.21	146.50
	High	12	12.79	153.50
	Total	24		

Test Statistics(b)

	Toral_Correct_Rate
Mann-Whitney U	68.500
Wilcoxon W	146.500
Z	-.216
Asymp. Sig. (2-tailed)	.829
Exact Sig. [2*(1-tailed Sig.)]	.843(a)
Exact Sig. (2-tailed)	.830
Exact Sig. (1-tailed)	.415
Point Probability	.035

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.20 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Total_FPR	Low	12	11.83	142.00
	High	12	13.17	158.00
	Total	24		

Test Statistics(b)

	Total_FPR
Mann-Whitney U	64.000
Wilcoxon W	142.000
Z	-.515
Asymp. Sig. (2-tailed)	.606
Exact Sig. [2*(1-tailed Sig.)]	.671(a)
Exact Sig. (2-tailed)	.744
Exact Sig. (1-tailed)	.372
Point Probability	.133

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 2.21 [Mann-Whitney Test]

Ranks

	Level of Technical Ability	N	Mean Rank	Sum of Ranks
Total_FNR	Low	12	13.21	158.50
	High	12	11.79	141.50
	Total	24		

Test Statistics(b)

	Total_FNR
Mann-Whitney U	63.500
Wilcoxon W	141.500
Z	-.517
Asymp. Sig. (2-tailed)	.605
Exact Sig. [2*(1-tailed Sig.)]	.630(a)
Exact Sig. (2-tailed)	.626
Exact Sig. (1-tailed)	.313
Point Probability	.016

a Not corrected for ties.
b Grouping Variable: Level of Technical Ability

Hypothesis 3.1 [Mann-Whitney Test]

Ranks

	Level of Phishing *Awareness*	*N*	*Mean Rank*	Sum of Ranks
Toral_Correct_Rate	Unaware	12	*7.67*	*92.00*
	Aware	12	17.33	*208.00*
	Total	24		

Test Statistics(b)

	Toral_Correct_Rate
Mann-Whitney U	14.000
Wilcoxon W	92.000
Z	-3.581
Asymp. Sig. (2-tailed)	.000
Exact Sig. [2*(1-tailed Sig.)]	.000(a)
Exact Sig. (2-tailed)	.000
Exact Sig. (1-tailed)	.000
Point Probability	.000

a Not corrected for ties.
b Grouping Variable: Level of Phishing Awareness

Hypothesis 3.2 [Mann-Whitney Test]

Ranks

	Level of Phishing Awareness	N	Mean Rank	Sum of Ranks
Total_FPR	Unaware	12	10.08	121.00
	Aware	12	14.92	179.00
	Total	24		

Test Statistics(b)

	Total_FPR
Mann-Whitney U	43.000
Wilcoxon W	121.000
Z	-1.868
Asymp. Sig. (2-tailed)	.062
Exact Sig. [2*(1-tailed Sig.)]	.101(a)
Exact Sig. (2-tailed)	.086
Exact Sig. (1-tailed)	.043
Point Probability	.022

a Not corrected for ties.
b Grouping Variable: Level of Phishing Awareness

Hypothesis 3.3 [Mann-Whitney Test]

Ranks

	Level of Phishing Awareness	N	Mean Rank	Sum of Ranks
Total_FNR	Unaware	12	17.17	206.00
	Aware	12	7.83	94.00
	Total	24		

Test Statistics(b)

	Total_FNR
Mann-Whitney U	16.000
Wilcoxon W	94.000
Z	-3.406
Asymp. Sig. (2-tailed)	.001
Exact Sig. [2*(1-tailed Sig.)]	.001(a)
Exact Sig. (2-tailed)	.000
Exact Sig. (1-tailed)	.000
Point Probability	.000

a Not corrected for ties.
b Grouping Variable: Level of Phishing Awareness

Hypothesis 3.4 [Mann-Whitney Test]

Ranks

	Level of Phishing Awareness	N	Mean Rank	Sum of Ranks
Pre_Correct_Rate	*Unaware*	*36*	*19.81*	*713.00*
	Aware	*12*	*38.58*	*463.00*
	Total	48		

Test Statistics(a)

	Pre_Correct_Rate
Mann-Whitney U	47.000
Wilcoxon W	713.000
Z	-5.208
Asymp. Sig. (2-tailed)	.000
Exact Sig. (2-tailed)	.000
Exact Sig. (1-tailed)	.000
Point Probability	.000

a Grouping Variable: Level of Phishing Awareness

Hypothesis 3.5 [Mann-Whitney Test]

Ranks

	Level of Phishing Awareness	N	Mean Rank	Sum of Ranks
Pre_FPR	Unaware	36	23.00	828.00
	Aware	12	29.00	348.00
	Total	48		

Test Statistics(a)

	Pre_FPR
Mann-Whitney U	162.000
Wilcoxon W	828.000
Z	-1.766
Asymp. Sig. (2-tailed)	.077
Exact Sig. (2-tailed)	.113
Exact Sig. (1-tailed)	.086
Point Probability	.068

a Grouping Variable: Level of Phishing Awareness

Hypothesis 3.6 [Mann-Whitney Test]

Ranks

	Level of Phishing Awareness	N	Mean Rank	Sum of Ranks
Pre_FNR	Unaware	36	29.22	1052.00
	Aware	12	10.33	124.00
	Total	48		

Test Statistics(a)

	Pre_FNR
Mann-Whitney U	46.000
Wilcoxon W	124.000
Z	-4.735
Asymp. Sig. (2-tailed)	.000
Exact Sig. (2-tailed)	.000
Exact Sig. (1-tailed)	.000
Point Probability	.000

a Grouping Variable: Level of Phishing Awareness

Hypothesis 4.1 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Retention_Post_Corr_Rate - Post_Correct_Rate	Negative Ranks	1(a)	2.00	2.00
	Positive Ranks	1(b)	1.00	1.00
	Ties	4(c)		
	Total	6		

a Retention_Post_Corr_Rate < Post_Correct_Rate
b Retention_Post_Corr_Rate > Post_Correct_Rate
c Retention_Post_Corr_Rate = Post_Correct_Rate

Test Statistics(b)

	Retention_Post_Corr_Rate - Post_Correct_Rate
Z	-.447(a)
Asymp. Sig. (2-tailed)	.655
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.250

a Based on positive ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 4.2 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Retention_Post_FPR - Post_FPR	Negative Ranks	0(a)	.00	.00
	Positive Ranks	3(b)	2.00	6.00
	Ties	3(c)		
	Total	6		

a Retention_Post_FPR < Post_FPR

b Retention_Post_FPR > Post_FPR
c Retention_Post_FPR = Post_FPR

Test Statistics(b)

	Retention_Post_FPR - Post_FPR
Z	-1.633(a)
Asymp. Sig. (2-tailed)	.102
Exact Sig. (2-tailed)	.250
Exact Sig. (1-tailed)	.125
Point Probability	.125

a Based on negative ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 4.3 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Retention_Post_FNR - Post_FNR	Negative Ranks	3(a)	2.00	6.00
	Positive Ranks	0(b)	.00	.00
	Ties	3(c)		
	Total	6		

a Retention_Post_FNR < Post_FNR
b Retention_Post_FNR > Post_FNR
c Retention_Post_FNR = Post_FNR

Test Statistics(b)

	Retention_Post_FNR - Post_FNR
Z	-1.732(a)
Asymp. Sig. (2-tailed)	.083
Exact Sig. (2-tailed)	.250
Exact Sig. (1-tailed)	.125
Point Probability	.125

a Based on positive ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 4.4 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Retention_Post_Corr_Rate - Post_Correct_Rate	Negative Ranks	4(a)	3.13	12.50
	Positive Ranks	1(b)	2.50	2.50
	Ties	1(c)		
	Total	6		

a Retention_Post_Corr_Rate < Post_Correct_Rate
b Retention_Post_Corr_Rate > Post_Correct_Rate
c Retention_Post_Corr_Rate = Post_Correct_Rate

Test Statistics(b)

	Retention_Post_Corr_Rate - Post_Correct_Rate
Z	-1.414(a)
Asymp. Sig. (2-tailed)	.157
Exact Sig. (2-tailed)	.313
Exact Sig. (1-tailed)	.156
Point Probability	.125

a Based on positive ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 4.5 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Retention_Post_FPR - Post_FPR	Negative Ranks	0(a)	.00	.00
	Positive Ranks	1(b)	1.00	1.00
	Ties	5(c)		
	Total	6		

a Retention_Post_FPR < Post_FPR
b Retention_Post_FPR > Post_FPR
c Retention_Post_FPR = Post_FPR

Test Statistics(b)

	Retention_Post_FPR - Post_FPR
Z	-1.000(a)
Asymp. Sig. (2-tailed)	.317
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.500
Point Probability	.500

a Based on negative ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 4.6 [Wilcoxon Signed Ranks Test]

Ranks

		N	Mean Rank	Sum of Ranks
Retention_Post_FNR - Post_FNR	Negative Ranks	1(a)	2.50	2.50
	Positive Ranks	4(b)	3.13	12.50
	Ties	1(c)		
	Total	6		

a Retention_Post_FNR < Post_FNR
b Retention_Post_FNR > Post_FNR
c Retention_Post_FNR = Post_FNR

Test Statistics(b)

	Retention_Post_FNR - Post_FNR
Z	-1.414(a)
Asymp. Sig. (2-tailed)	.157
Exact Sig. (2-tailed)	.313
Exact Sig. (1-tailed)	.156
Point Probability	.125

a Based on negative ranks.
b Wilcoxon Signed Ranks Test

Hypothesis 4.7 [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Retention_Total_Corr_Rate	First-Old Approach Group	6	6.50	39.00
	First-New Approach Group	6	6.50	39.00
	Total	12		

Test Statistics(b)

	Retention_Total_Corr_Rate
Mann-Whitney U	18.000
Wilcoxon W	39.000
Z	.000
Asymp. Sig. (2-tailed)	1.000
Exact Sig. [2*(1-tailed Sig.)]	1.000(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.526
Point Probability	.052

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 4.8 [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Retention_Total_FPR	First-Old Approach Group	6	8.00	48.00
	First-New Approach Group	6	5.00	30.00
	Total	12		

Test Statistics(b)

	Retention_Total_FPR
Mann-Whitney U	9.000
Wilcoxon W	30.000
Z	-1.554
Asymp. Sig. (2-tailed)	.120
Exact Sig. [2*(1-tailed Sig.)]	.180(a)
Exact Sig. (2-tailed)	.177
Exact Sig. (1-tailed)	.089
Point Probability	.049

a Not corrected for ties.
b Grouping Variable: Group

Hypothesis 4.9 [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Retention_Total_FNR	First-Old Approach Group	6	4.83	29.00
	First-New Approach Group	6	8.17	49.00
	Total	12		

Test Statistics(b)

	Retention_Total_FNR
Mann-Whitney U	8.000
Wilcoxon W	29.000
Z	-1.696
Asymp. Sig. (2-tailed)	.090
Exact Sig. [2*(1-tailed Sig.)]	.132(a)
Exact Sig. (2-tailed)	.134
Exact Sig. (1-tailed)	.067
Point Probability	.043

a Not corrected for ties.
b Grouping Variable: Group

Retention Inconsistency Analysis

CDR [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_Correct_Rate	First-Old Approach Group	6	5.00	30.00
	First-New Approach Group	6	8.00	48.00
	Total	12		

Test Statistics(b)

	Post_Correct_Rate
Mann-Whitney U	9.000
Wilcoxon W	30.000
Z	-1.573
Asymp. Sig. (2-tailed)	.116
Exact Sig. [2*(1-tailed Sig.)]	.180(a)
Exact Sig. (2-tailed)	.232
Exact Sig. (1-tailed)	.116
Point Probability	.087

a Not corrected for ties.
b Grouping Variable: Group

FPR [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_FPR	First-Old Approach Group	6	6.50	39.00
	First-New Approach Group	6	6.50	39.00
	Total	12		

Test Statistics(b)

	Post_FPR
Mann-Whitney U	18.000
Wilcoxon W	39.000
Z	.000
Asymp. Sig. (2-tailed)	1.000
Exact Sig. [2*(1-tailed Sig.)]	1.000(a)
Exact Sig. (2-tailed)	1.000
Exact Sig. (1-tailed)	.773
Point Probability	.545

a Not corrected for ties.
b Grouping Variable: Group

FNR [Mann-Whitney Test]

Ranks

	Group	N	Mean Rank	Sum of Ranks
Post_FNR	First-Old Approach Group	6	8.00	48.00
	First-New Approach Group	6	5.00	30.00
	Total	12		

Test Statistics(b)

	Post_FNR
Mann-Whitney U	9.000
Wilcoxon W	30.000
Z	-1.563
Asymp. Sig. (2-tailed)	.118
Exact Sig. [2*(1-tailed Sig.)]	.180(a)
Exact Sig. (2-tailed)	.238
Exact Sig. (1-tailed)	.119
Point Probability	.097

a Not corrected for ties.
b Grouping Variable: Group

Glossary

LoTA Level of Technical Ability.

LoPA Level of Phishing Awareness.

Post_FPR False Positive Rate Based on Post-Treatments Websites.

Post_FNR False Negative Rate Based on Post-Treatments Websites

Post_CDR Correct Decision Rate Based on Post-Treatments Websites.

Pre_FPR False Positive Rate Based on Pre-Treatments Websites.

Pre_FNR False Negative Rate Based on Pre-Treatments Websites

Pre_CDR Correct Decision Rate Based on Pre-Treatments Websites.

Total_FPR False Positive Rate Based on Post- and Pre-Treatments Websites.

Total _FNR False Negative Rate Based on Post- and Pre-Treatments Websites.

Total _CDR Correct Decision Rate Based on Post- and Pre-Treatments Websites.

Retention_Post_FPR False Positive Rate Based on Post-Treatments Websites in Retention Experiments.

Retention_Post_FNR False Negative Rate Based on Post-Treatments Websites in Retention Experiments.

Retention_Post_CDR Correct Decision Rate Based on Post-Treatments Websites in Retention Experiments.

Retention_Pre_FPR False Positive Rate Based on Pre-Treatments Websites in Retention Experiments.

Retention_Pre_FNR False Negative Rate Based on Pre-Treatments Websites in Retention Experiments.

Retention_Pre_CDR Correct Decision Rate Based on Pre-Treatments Websites in Retention Experiments.

Retention_Total_FPR False Positive Rate Based on Post- and Pre-Treatments Websites in Retention Experiments.

Retention_Total _FNR False Negative Rate Based on Post- and Pre-Treatments Websites in Retention Experiments.

Retention_Total _CDR Correct Decision Rate Based on Post- and Pre-Treatments Websites in Retention Experiments.