INTEGRATING PERSUASIVE PRINCIPLES IN PHISHING EMAILS

NURUL AKBAR

UNIVERSITY OF TWENTE.

Supervisors:
Prof.Dr.P.H. Hartel
E.E.H. Lastdrager MSc.

SERVICES, CYBERSECURITY AND SAFETY RESEARCHGROUP
Faculty of Electrical Engineering, Mathematics and Computer Science
University of Twente
August 2014

Nurul Akbar: *Integrating persuasive principles in phishing emails,* Master thesis, © August 2014

SUPERVISORS:
Prof.Dr.P.H. Hartel
Put name here
E.E.H. Lastdrager MSc.

LOCATION:
Enschede

The life of this world is only the enjoyment of deception.

— Quran 3:185

ABSTRACT

Valuable information, such as login credentials and personal sensitive information, can be acquired by exploiting vulnerabilities within the user's understanding of a system, and particularly a lack of in-depth understanding of the user interface.

As the barrier to abuse system vulnerabilities has been improved significantly with time, attacking users' psyche has rapidly become a more efficient and effective alternative. The usage of email as a media electronic of communication has been exploited by phishers to deliver their attacks. The success of a phishing attack through distributed emails is determined by the response of the unsuspecting recipients. Persuasion techniques often used to obscure users in making decisions to response or not.

To protect users from phishing attacks in the email level, system designers and security professionals need to understand how phishers use persuasion techniques in phishing emails. In this thesis we present an improved understanding of persuasion techniques in phishing emails.

Our research is aimed to understand the characteristics of phishing emails considering persuasion techniques in the real world analysis which has not been done yet. The analysis consists of finding relationships between persuasion techniques and generic structural properties of phishing emails.

Our technological powers increase,
but the side effects and potential hazards also escalate.

— Arthur C, Clarke

ACKNOWLEDGMENTS

First and foremost, I would like use this opportunity to thank both of my supervisors, Prof. Dr. Pieter Hartel and Elmer Lastdrager MSc. I am thankful for their unrelenting guidance and support that have made this research possible. They have exceeded beyond the expected duties as supervisors and help me to make this master thesis possible. They have provided me with invaluably constructive thoughts and critical feedback. I am sincerely grateful to them for sharing their truthful and illuminating views on a number of issues related to the research. It was Elmer's vision to integrate phishing emails corpus with cialdini's principles as the core of my research. He has assisted me in obtaining the data as the it is confidential and sensitive.

I would like to thank PhD candidates in SCS group for letting me plucking their ideas when I did a brief presentation at the beginning of my research. Geert Jan for letting me work in the lab. Suse and Bertine for being able to lend me the key when no one else in the lab. I would also like to express my appreciation and gratitude to Drs. Jan Schut for providing me valuable advise and direction throughout my study in University of Twente.

A special thanks to Eyla who has give me incentives to strive towards my goal and for being there in difficult times. Gaurav and Vignesh who have assisted me by giving feedbacks on my writing. All my friends, Aldi, Saud, and all the others who supported me either directly and indirectly during my master studies. Without all their supports, accomplishing my studies would not have been possible. I would like to thank my family members and relatives who have supported financially and emotionally throughout my entire master education. Words cannot express how grateful I am to my mother and father in spite of all the difficult times, I thank you all for letting me cherish my dream. Lastly, I thank God almighty for answering my prayers.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AOL American Online

URL Uniform Resource Locator

IP Internet Protocol

TLD Top Level Domain

1 INTRODUCTION

With the advancement of information technology in modern generation, the evolution of digital era has become more mature in the sense of effectiveness and easiness for societies. They can sell and buy goods, conduct banking activities and even participate in political activities such as election through online. Trusted entities such as financial institutions generally offer their products and services to the public through the Internet. Furthermore, modern technology has greatly impacted our society in different ways, such as the way we communicate with each other. Nowadays, we are no longer need to use a computer to send an email, we can just use our smartphone, which we carry every day in our pockets, with internet connectivity to send an email. As a result, human society has been utilizing technology means such as emails, websites, online payment system, social networks to achieve their tasks efficiently, affordable and more relevant. However, the advancement in information and communication technology has been a double-edged sword. As the internet increasingly become more accessible, people tend to share more about themselves and as a consequence, it becomes easier to get personal information about someone on the Internet. Cyber criminals see this opportunity as a way to manipulate consumers and exploit their confidential information such as usernames, passwords, bank account information, credit card or social security numbers. Personalized information about someone such as email addresses, phone numbers, birth dates, relationships or work place might be obtained from the internet. Consequently, cyber criminals can compose an attack in a personalized way to persuade intended victims to grant their malicious requests.

One particular type of cyber crimes is called phishing. Many possible incentives that drive phishing attacks such as illicit corporate espionage, political power, and the most common incentive of phishing attacks is financial benefits. The attacker generally masquerades legitimate institutions to trick users into disclosing personal, financial or computer account information [32]. The attacker can then use this information for criminal activities such as identity theft or fraud. To manipulate unsuspecting victims, the attacker often uses emails and websites as the techniques to execute the attacks [32][14]. The practice of utilizing emails and websites is indeed useful as communication media, however, they can also accommodate deceptive attacks such as phishing as a form of social engineering and deception [32][2][14][33][30]. Social engineering involves the techniques used to

deceive people in order to get a compliance or a respond by specific actions which will disclose their sensitive information, such as replying to a fake email or clicking a link within an email [54]. Moreover, phishers often use persuasion techniques to manipulate potential victims to engage in certain emotions such as excitement and fear as well as interpersonal relationship, such as trust and commitments to divert users' attention [79]. Such persuasive influence might be delivered through phone calls, text messages, private messages or emails as ways to distract recipients' decisions.

1.1 PROBLEM STATEMENT

Countermeasures against phishing attacks in the email level can be technical or non technical approach. One of technical approaches to detect phishing email is achieved by discriminating phishing and original email based on its structural properties using machine learning technique [4]. Furthermore, One of non technical approaches to defense against phishing attacks is to make people aware of the threats. Security awareness in regards of phishing attacks might be achieved by embedded training methods that teach people about phishing during their normal use of email [38].

Moreover, the common characteristics of a phishing email can be distinguished by its structural properties such as misleading hyperlinks and misleading header information [82, 83]. However, to make a phishing email efficient, its content requires the intended victim to urgently act upon it, for example, an email that informs about account termination if the recipient does not respond or perform an action within a limited time. In order to obtain a compliance from the recipient in a phishing email, persuasion is used as underlying techniques to get a positive response from an intended victim [79].

The success of a phishing attack through distributed emails is determined by the response of the unsuspecting recipients. User decisions to click a link or open an attachment in an email might be influenced by how strong a phisher can persuade the intended victim. However, current literatures shows an absence in conducting a real world analysis of phishing email characterization by integrating persuasion techniques. This characterization can show to what extent the persuasion techniques are used in phishing emails. Our research fills the void as a milestone towards countermeasure against phishing attacks with an insight of psychological aspect.

1.2 RESEARCH GOAL

The main goal of this research is to characterize phishing email properties considering persuasive principles by finding the association between generic properties and persuasive principles. These generic

properties consist of phishing email structural properties or features based on the literature survey findings. Each of these properties and each of the persuasive principle will be introduced as a variable in our methodology. We will look for frequency and relationship involving these variables. This relationship can be used to show a different perspective of phishing email characteristics considering the persuasive elements within its content. The integration of persuasive principles in phishing emails also can be used to generate a new method in an automatic way of detecting phishing emails as one of the primary delivery techniques of phishing attacks.

1.3 RESEARCH QUESTIONS

To be able to meet the goal, we formulated two main research questions as follows:

- RQ1: What are the characteristics of reported phishing emails?
- RQ2: To what extent are persuasive principles used in phishing emails?

Several aspects of phishing email characteristics and hypotheses related to the research questions will be addressed in detail in Chapter 4.

1.4 STRUCTURES

This research project is structured as follows:

Chapter 2 describes background and literature reviews about phishing in general. The subsections include; a general understanding of what is phishing in terms of history and definition, an overview of its damage in terms of money, an exploration of its modus operandi based on phishing stages or phases, a basic understanding of bad neighborhood and phishing, general phishing countermeasures and lastly the human factor in phishing.

Experimental exploration with methods and results will be described in chapter 3. It consists of an experiment on the existence of bad neighborhood and finding differences of phishing website vs its original website. The methods and the results of the individual experiment will be presented in the respective sections.

In chapter 4 presents the rationale of our main research questions and hypotheses. It includes what aspect to be considered to answer the characteristics of phishing emails in the dataset and the motivation of our hypotheses to support our research questions.

In chapter 5, we will discuss our main data analysis and results. It includes the details of research methodology that we conducted as well as the results of our analysis.

Lastly, in chapter 6 we will present our discussion and conclusion of the research project, how the research questions are answered along with the recommendations, limitations and how these limitations could become the basis of further research.

2

BACKGROUND & LITERATURE REVIEW

In order to meet our goal, necessary knowledge on phishing in general is required. This chapter introduces general understanding of phishing, an exploration of its damage in term of money, the overview of its modus operandi, a brief explanation on types of phishing, an understanding of bad neighborhoods on phishing, general phishing countermeasures and human factor in phishing attacks.

2.1 WHAT IS PHISHING?

While the Internet has brought convenience to many people for exchanging information, it also provides opportunities to carry malicious behavior such as online fraud on massive scale with a little cost to the attackers. The attackers can manipulate the Internet users instead of computer system (hardware or software) that significantly increase the barriers of technological crime impact. Such human centered attacks could be done by social engineering. According to Jakobsson, et al. phishing is a form of social engineering that aims to retrieve credential from online users by mimicking trustworthy and legitimate institutions [32]. Phishing has a similar basic principle as 'fishing' in the physical world. Instead of fish, online users are lured by authentic looking communication and hooked by authentic looking websites. Not only that, online users also may be lured by respond to a phishing email, either replying or clicking an obfuscated link within its content. There are diverse definitions of phishing in our literature reviews, therefore, we would like to discuss about its universal definition in later section. However, one of phishing definitions according to Oxford dictionary:

> "A fraudulent practice of sending emails purporting to be from reputable companies in order to induce individuals to reveal personal information, such as passwords and credit card numbers, online" [15].

Several studies suggest that phishing is a form of online criminal activity by using social engineering techniques [32][79][30][4]. An individual or a group who uses this technique is called *Phisher(s)*. After successfully obtained a sensitive information from the victim, phishers use this information to access victim's financial accounts or committing credit card frauds. However, to formalize the damage of

phishing in term of money is a challenging task. We will briefly explore the cost of phishing attacks in the later section.

Furthermore, the technique or modus operandi of phishing may vary, but the most common technique of phishing attacks done by using fraudulent emails and websites [33]. A fraudulent website is designed in such a way that it may be identical to its legitimate target. While it may be true, phishing website also could be completely different with its target as there is no level of identicalness. With this in mind, preliminary analysis on what changed or added in phishing website would be conducted in the later section. In the following subsections, we will introduce how phishing was originally came about and how current literatures formally define phishing.

2.1.1 *The History*

The first time the term "phishing" was published by the American Online (AOL) UseNet Newsgroup on January 2, 1996 and was started to expand in 2004 [65]. Since then, we considered phishing development in cyberspace has been flourishing by phishers to make profit. Total losses due to phishing in 2004 reached more than U.S. $ 2 billion, it was involving more than 15,000 sites that become victims [22]. We will try to discuss about direct and indirect cost at present days in the later section. Evidently, Jakobsson, et al. [32] mentioned that in the early years of 90's (according to [65] it was around 1995) many hackers would create bogus AOL user accounts with automatically generated fraudulent credit card information. Their intention to give this fake credit card information was to simply pass the validity tests performed by AOL. By the time the tests were passed, AOL was thinking that these accounts were legitimate and resulted to activate them. Consequently, these hackers could freely access AOL resources until AOL tried to actually bill the credit card. AOL realized that these accounts were using invalid billing information, thus deactivated the account.

While creating false AOL user accounts with fake credit card information was not exactly phishing attacks, but AOL's effort to counter against the attacks was leading to development of phishing. This countermeasure includes directly verifying the legitimacy of credit card information and the associated billing identity, forced hackers to pursue alternative way [32]. Hackers were masquerading as AOL's employees asking to other users for credit card information through AOL instant messenger and email system [65]. Jakobsson et al. suggest that phishing attacks were originating from this incident [32]. Since such attack has not been done before, many of users have been victimized by then. Eventually, AOL enforced warning system to the most of its customers to be vigilant when it comes to sensitive information [65]. At the present day, phishing attacks might not only being

motivated by financial gain but also political reason, and they have been emerging not only aim to AOL users, but also any online users. Consequently, large number of legitimate institutions such as PayPal and eBay are being spoofed.

2.1.2 *The universal definition*

Before we begin to understand deeper about how and why phishing attack works, we will briefly explore common phishing definition. Currently, there is no consensus definition, since almost in every research papers, academic textbook or journals has its own definition of phishing [32, 33, 73, 9, 63, 31, 14]. Phishing is also constantly evolving, so it might be very challenging to define its universal terminology. There is not so much study that specifically addresses the standard of phishing definition. However, one research conducted by Lastdrager [41] addressed to achieve consensual definition of phishing. Before we comply with one consensual phishing terminology, we will take a look at various phishing definitions from other sources:

> "Phishing is the act of sending a forged e-mail (using a bulk mailer) to a recipient, falsely mimicking a legitimate establishment in an attempt to scam the recipient into divulging private information such as credit card numbers or bank account passwords" [33]

> "Phishing is a form of Internet scam in which the attackers try to trick consumers into divulging sensitive personal information. The techniques usually involve fraudulent e-mail and websites that impersonate both legitimate e-mail and websites" [73]

> "Phishing is an attack in which victims are lured by official looking email to a fraudulent website that appears to be that of a legitimate service provider" [9]

> "In phishing, an automated form of social engineering, criminals use the internet to fraudulently extract sensitive information from business and individuals, often by impersonating legitimate websites" [63]

It is noteworthy that the definition described by James, et al, Tally, et al, and Clayton, et al. [33, 73, 9] specifies that the phishers only use email as a communication channel to trick potential victims. While it might be true because using email would greatly cost effective, but we believe that phishing is not only characterized by one particular technological mean, as phishers can also use any other electronic communication to trick potential victims (i.e private message on online social network). This definition is also similar to dictionary libraries [15, 16, 75] that mention email as a medium communication between phishers and users.

We believe that standard definition of phishing should be applicable in most of phishing concept that are presently defined. Consequently, the high level of abstraction and is required to build common definition on phishing. We have convinced that the formal definition of phishing should not focus on the technology that is being used but rather on the technique how the deception is being conducted, the method of an "act" if you will. Therefore, We follow the definition of phishing by Lastdrager [41] which stated:

> *"Phishing is a scalable act of deception whereby impersonation is used to obtain information from a target"*

According to Lastdrager [41], to achieve this universal definition, a systematic review of literature up to August 2013 was conducted along with manual peer review, which resulted in 113 distinct definitions to be analyzed. We thereby agree with Lastdrager [41] that this definition addresses all the essential elements of phishing and we will adopt it as universally accepted terminology throughout our research.

2.2 THE COSTS OF PHISHING ATTACKS

It is a challenging task to find a real cost from phishing attacks in term of money or direct cost. This due to financial damage for bank is only known by banks and most institutions do not share this information with the public. Evidently, Jakobsson et al. argue that phishing economy is consistent with black market economy and does not advertise its successes [32]. On this section, a brief explanation of direct and indirect cost on phishing attack will be illustrated based on literature reviews.

According to Jakobsson et al., direct cost is depicted by the value of money or goods that are directly stolen through phishing attack [32]. While indirect cost is the costs that do not represent the money nor goods that are actually stolen, but it is the costs has to be paid by the people who handle these attacks [32] (i.e. time, money and resources spent to reset people password).

As we mentioned earlier, the difficulty of assessing the damage on phishing attacks is caused by banks and institutions that keep this information themselves and the unwillingness of many users to share to acknowledge that they have been victimized by phishing attacks. This happens because of fear of humiliation, financial loses, or legal liability [32]. Evidently, studies estimate the damage ranging from $61 million [27] to $3 billion per year [52] of direct losses to victims in the US only [28][58]. In addition, the Gartner Group claimed to estimate of $1.2 billion direct losses of phishing attack to US banks and credit card companies for the year 2004 [43]. By the 2007, it escalated to more than $3 billion loss [53]. The estimation also performed by TRUSTe

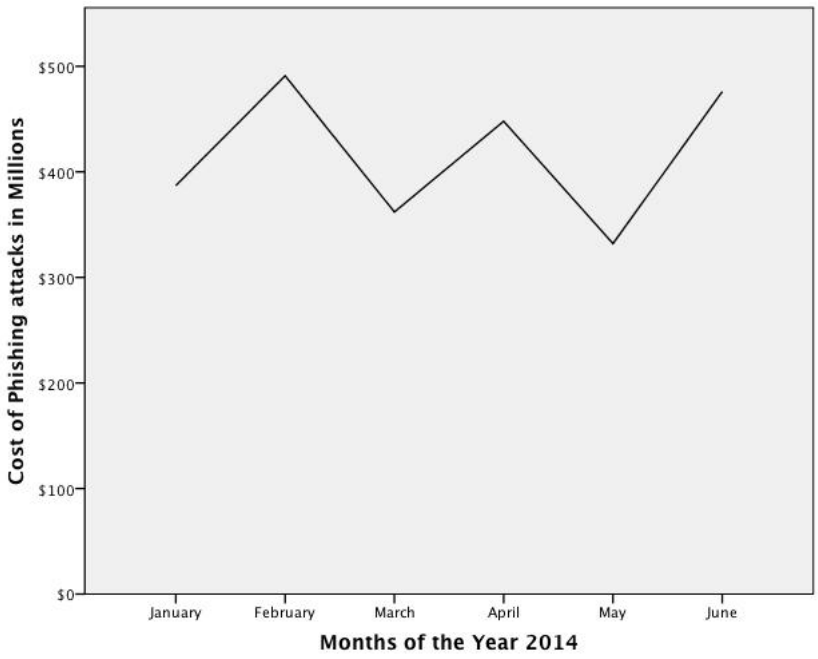


Figure 1: Global phishing cost from January 2014 until June 2014 [23]

and Ponemon Institute that stated the cost of phishing attack was up to $500 millions losses in the US for the same year [1]. Recently, RSA FraudAction gives monthly reports of how much the global phishing cost and we compiled them together to make a line graph in Figure 1 [23]. However, they do not specify whether this damage is direct cost or indirect cost. In their study, we can see that there are fluctuation of losses in term of money. Furthermore, this total cost is indeed in accordance with the total cost ranging from $61 million to $3billion per year [27, 52, 28, 58].

With this in mind, we might ask how phishing attack is carried out? is there any stages involved? The next chapter we will review its modus operandi in term of phishing stages or phases.

2.3 MODUS OPERANDI

As we mentioned earlier, Phishing attack is a subset of identity theft. The modus operandi might be carried out firstly by creating a fake website that spoofs legitimate website such as financial website, either identical or not identical as long as the phishers get responds from unsuspected victims. After that, the phishers will try to trick the potential victim to submit important information such as usernames, passwords, PINs, etc. through a fake website that they have created or through email reply from victims. With the information obtained, they will try to steal money from their victims if target institution is

1 http://www.theregister.co.uk/2004/09/29/phishing_survey/

a bank. Phishers employ variety of techniques to trick potential victims to access their fraudulent website. One of the typical ways is by sending illicit email in a large scale claiming to be from legitimate institution. In the email content, they usually imitate an official-looking logo, using good business language style and often also forge the email headers to make it look like originating from legitimate institution. For example, the content of the email is to inform the user that the bank is changing its IT infrastructure, and request urgently that the customer should update their data with the consequence of loosing their money if the action does not take place. When the user click the link that was on the email message, they will be redirected to a fraudulent website, which will prompt the victim to fill in the details of their information. While there are various techniques of phishing attack, we will address the common phases of phishing that we analyzed by literature survey by several studies and also we will address our own phishing phases. These phases are compiled in Table 1.

Table 1: Compilation of phishing phases

J. Hong [28]
1. Potential victims receive a phish 2. The victim may take a suggested action in the message 3. The phisher monetizes the stolen information
Frauenstein, et al. [24]
1. Planning 2. Email Design 3. Fabricated story 4. Threatening tone/Consequences 5. Spoofed website
Wetzel [77]
1. Planning 2. Setup 3. Attack 4. Collection 5. Fraud 6. Post-attack
Tally, et al. [73]
1. The attacker obtains E-mail addresses for the intended victims 2. The attacker generates an E-mail that appears legitimate 3. The attacker sends the E-mail to the intended victims in a way that appears legitimate and obscures the true source 4. The recipient opens a malicious attachment, completes a form, or visits a web site 5. Harvest and exploitation
Emigh [21]

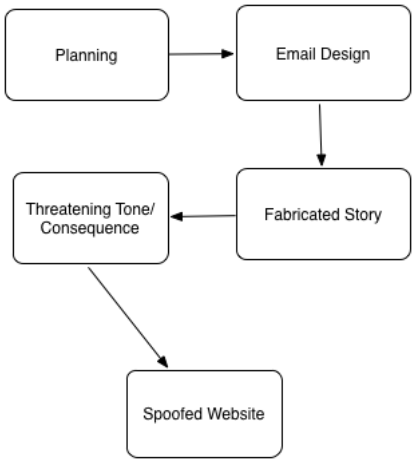


Figure 2: Phishing processes based on Frauenstein[24]

1. A malicious payload arrives through some propagation vector 2. The user takes an action that makes him or her vulnerable to an information compromise 3. The user is prompted for confidential information, either by a remote web site or locally by a Web Trojan 4. The user compromises confidential information 5. The confidential information is transmitted from a phishing server to the phisher 6. The confidential information is used to impersonate the user 7. The phisher engages in fraud using the compromised information
Nero et al. [59]
1. Preparation 2. Delivery of the Lure 3. Taking the Bait 4. Request for Confidential Information 5. Submission of Information 6. Collection of Data 7. Impersonation 8. Financial Gain

Based on the example scenario explained earlier, phishing attacks may consist of several phases. J. Hong [28] argued that there are three major phases. while Frauenstein, et al. [24] suggested that there are five main processes are used to perform phishing attacks based on the perspective of the attacker.

As we illustrated in Figure 2, on the first process is called *Planning*, a phisher usually would do some reconnaissance on how would the attack is executed and what information would be obtained from the victim. On the second process, the phisher would think about the design of the email. This email is desired by the phisher to look as legit as possible to potential victim. For this purpose, target institutions logo, trademark, symbol, etc. are used to make the content look official to the victim. The author called this process as *Email Design*. Figure 3 illustrates the example of fake ING bank logo in a phishing

Figure 3: Example of fake ING logo in phishing email

email to create "legitimate" feel[2]. On the third process, the phisher *fabricates* a story to make potential victim think that email is important. To achieve users attention, phisher might build up a story about system upgrade, account hijacked or security enhancement so that the victim would feel obliged to be informed. Evidently, this technique corresponds with Cialdini [8] that suggests there are six principles to persuade people to comply with a request. On the fourth process, a phisher usually include *threatening tone* or explain the urgency and consequences if the potential victim chooses not to take action desired by the phisher (for example; account removal, account blocked, etc.). Consequently, users may fear of their account being deleted. The last process involved with fraudulent website that has been created by the phisher. Users may falsely believe to the message given in the email and may click a Uniform Resource Locator (URL) that is embedded in the email. Subsequently, the URL would redirect users to a *spoofed website* which may prompt users' sensitive information. Furthermore, the website might be created to be as similar as possible to the target institution's website, so that potential victim may still believe that it is authentic. We will explain more on Cialdini's six basic tendencies of human behavior in generating positive response to persuasion [8] in a later section.

Considering that phishing attack is a process, Wetzel [77] suggested a taxonomy to make sense of the complex nature of the problem by mapping out a common attacks lifecycle, and a possible set of activities attackers engage in within each phase. The taxonomy is illustrated in Figure 4. We speculated that Wetzel's taxonomy is not analogous with Frauenstein's main phishing processes [24]. The difference is that Frauenstein et al. only focus in the design of the attack while Wetzel has added several phases like *Collection*, *Fraud* and *Post-attack*, therefore, Wetzel taxonomy is more holistic in term of phishing.

2 http://www.martijn-onderwater.nl/wp-content/uploads/2010/03/ing-phishing.jpg

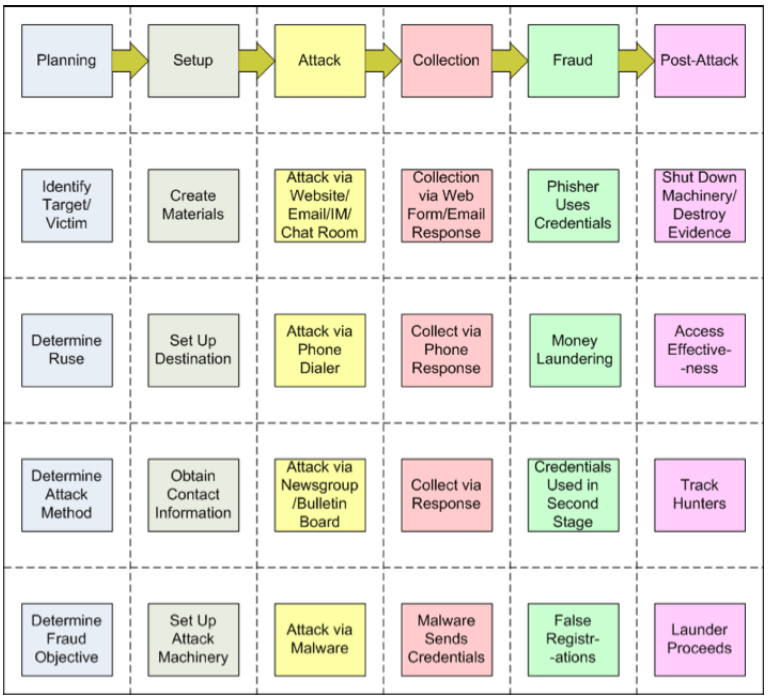


Figure 4: Phishing attack taxonomy and lifecycle[77]

As we listed Wetzel's taxonomy in Table 1, we explain more of the taxonomy as follows:

1. *Planning*: Preparation carried out by the phisher before continue to the next phase. Example activities include identifying targets and victims, determine the method of the attack, etc.

2. *Setup*: After the target, victim and the method are known, the phisher would craft a platform where the victim's information could be transmitted and stored, for example: fraudulent website/email.

3. *Attack*: Phisher distributes their fraudulent platform so that it can be delivered to the potential victims with fabricated stories.

4. *Collection*: Phisher collects valuable information via response from the victims

5. *Fraud*: Phisher abuses victim's information by impersonates the identity of the victim to the target. For example, A has gained B's personal information to access C so that A can pose as B to access C.

6. *Post-attack*: After the phisher gained profit from the attack and abuse phases, a phisher would not want to be noticed or detected by authority. Thus, phisher might need to destroy evidence of the activities that he/she previously were executed.

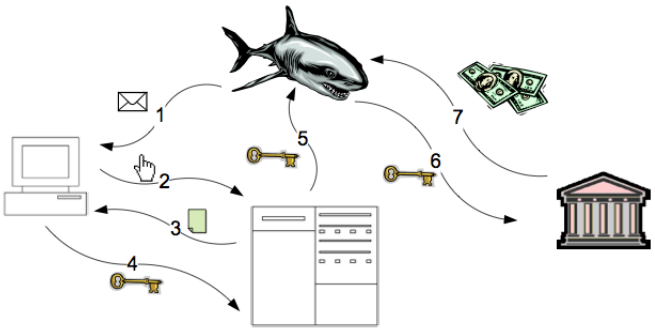


Figure 5: Flow of information in phishing attack [21]

As shown in Table 1. Tally, et al. suggest that there are several phases involved in phishing attack based on the attacker's point of view [73]. The first phase, it represents the planning, as we understand the attacker collects the email address of unsuspecting victims. The second phase, considering that it is related to creating a fake email that appears legitimate, this phase can be viewed as design phase. On the third phase, we consider this as delivery and attack phases as it involves the attacker sends the fake email to the unintended victims and obfuscated the true source. The fourth phase represents attack phase as it involves with the recipient complies with the attacker's request(s). Lastly the fifth phase, it represents the fraud phase, as it related to haversting and exploiting victim's resources by the attacker. Additionally, the phases described by Tally, et al. [73] are comparable with the information flow explained by Emigh[21] illustrated in Figure 5 and explained in Table 1. Phishing attack steps that executed by the phisher are also being addressed by Nero, et al [59]. In their study, a successful phishing attack involves several phases which can bee seen and compare in Table 1.

Based on our analysis by looking at the pattern of other phases from various sources, there is a major similarity between them. Therefore, we would like to define and design our own phase that are integrated with three key components suggested by Jakobsson, et al. [32]. These key components are include *the lure*, *the hook* and *the catch*. As we designed in Figure 6, we synthesized these three components with our phases based on the attacker point of view as follows:

- The lure
1. Phishers prepare the attack
2. Deliver initial payload to potential victim
3. Victim taking the bait
- The hook
4. Prompt for confidential information
5. Disclosed confidential information
6. Collect stolen information
- The catch
7. Impersonates victim

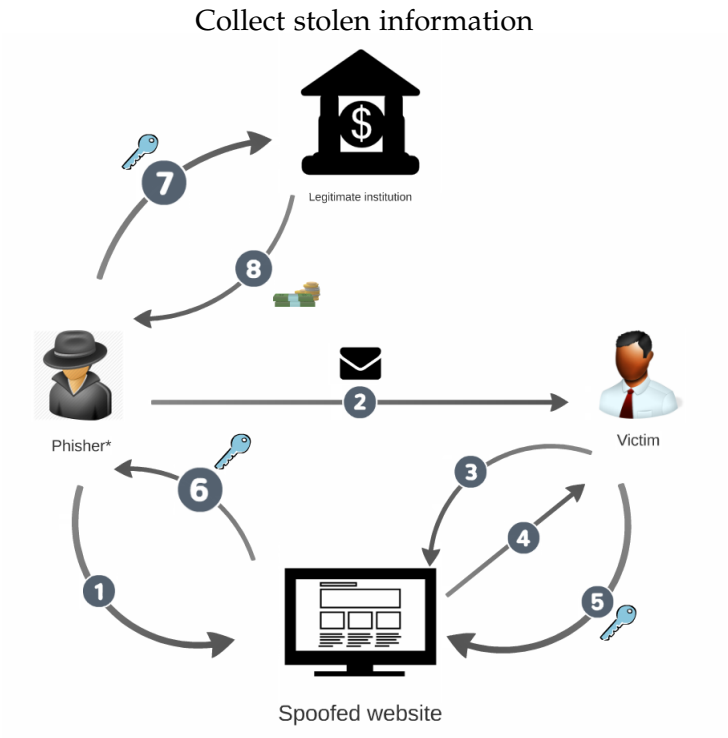


Figure 6: Information flow phishing attack

8. Received pay out from the bank

It is important to know that in the phase 3, there are different scenarios such as; victim might be redirected to a spoofed website, victim may comply to reply the email, victim may comply to open an attachment(s) or victim may comply to call by phone. However, in Figure 6, we have only illustrated the phases if the bait was using a spoofed website as a method.

We have reviewed various phases in phishing attack and from the review, we have constructed our own phases. In the next section, a brief introduction in respect to the types of phishing will be described. We believe that the general understanding of phishing types will help our main analysis to characterize phishing email properties.

2.4 TYPES OF PHISHING

In January 2014, 8300 patients data are being compromised in medical company in the US [25]. The data includes names, addresses, date of birth and phone numbers were being stolen. Other than demographic information, clinical information associated with this data was also stolen, including social security numbers. In the April 2014, phishers have successfully stolen US$163,000 from US public school based on Michigan [3]. It has been said that the email prompted to transfer money is coming from the finance director of the school. In March 2014, Symantec has discovered phishing attack aimed at

Google drive users [68]. The attack was carried firstly with incoming email asking for opening document hosted at Google docs. Users that have clicked on the link are taken to fraudulent Google login page prompted Google users credentials. Interestingly, the URL seems very convincing because it hosted on Google secure servers. We hypothesized that even more phishing incidents on financial area as well, but sometimes the news is kept hidden due to creditability reason. With this in mind, we believe fake websites might be hosted in the network which has more phishing domain than other networks. Subsequently, in the next section, we will discuss bad neighborhoods on phishing.

One may ask, what type of phishing are these? What are the general types of phishing relevant to our research? Evidently, based on the cost of phishing attacks in Section 2.2, the threat of phishing attacks is still alarming and might be evolving in the future with more sophisticated technique of attacks. For this reason, it might be useful to provide a brief insight on popular variants of phishing that currently exist. We will briefly explain the types of phishing which are the most relevant to our research based on Jakobsson, et al. [32]. These types of phishing is strongly related to the phishing definition that we used, considering phishing is based on the act of deception by the phishers.

2.4.1 *Phishing based on visual similarities*

Since all phishing is based on deception and social engineering, there is a phishing scenario based on visual similarities. Typical scenario of phishing based on visual similarities is to send a large amount of illicit emails containing call to action asking recipients to click embedded links [32]. These variations include cousin domain attack. For example, legitimate PayPal website addressed as www.paypal.com, this cousin domain attacks confuse potential victims to believe that www.paypal-security.com is a subdivision of the legitimate website due to identical looking addresses. Similarly, homograph attacks create a confusion using similar characters to its addresses. For example, www.paypal.com and www.paypa1.com, both addresses look the same but on the second link, it uses “1” instead of “l”.

Moreover, phishers may embed a login page directly to the email content. This suggests the elimination of the need of end-users to click on a link and phishers do not have to manage an active fraudulent website. IP addresses are often used instead of human readable hostname to redirect potential victim to phishing website and JavaScript is used to take over address bar of a browser to make potential victims believe that they are communicating with the legitimate institution. We will also see few examples of malicious JavaScript on our preliminary analysis section.

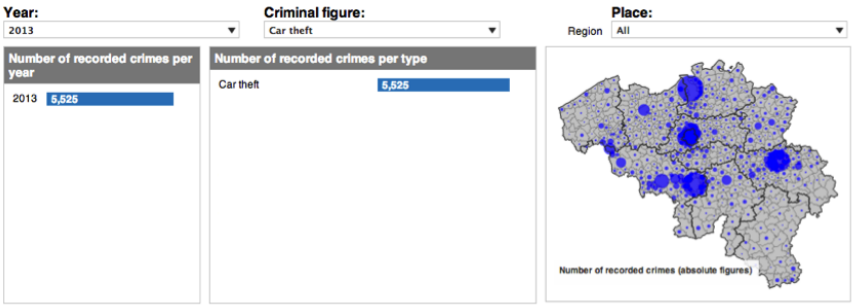


Figure 7: Belgium police record on car theft incidents in 2013

Another type of deceptive phishing scheme is rock-phish attacks. They held responsible for half a number of reported incidents worldwide in 2005 [55]. These attacks evade email filters by utilizes random text and GIF images which contain the actual message. Rock phish attacks also utilize a toolkit that capable to manage several fraudulent websites in a single domain. Sometimes, deceptive phishing schemes lead to installation of malware when users visit fraudulent website and we will describe malware based phishing scheme in the next section.

2.4.2 *Malware-based phishing*

Generally, malware based phishing refers to any type of phishing which involves installing malicious piece of software onto users' personal computer [32]. Subsequently, this malware is used to gather confidential information from victims instead of spoofing legitimate websites. This type of phishing incorporates malwares such as keyloggers/screenloggers, web Trojans and hosts file poisoning.

2.5 BAD NEIGHBORHOODS AND PHISHING

In the physical world, there are parts of certain area that have higher crime rates than others (e.g.. Bronx in the US) which called hotspots. Evidently, it is statistically more likely that a crime will occur compared to other locations [57]. To better illustrate this analogy, the police department in Belgium [13] has put up statistical information regarding crime rates in the country. Figure 7 shows an example in 2013; there were up to 5525 car theft incidents recorded and we can see there are certain areas that have higher probability that a car got stolen. For example, Antwerp had 415 incidents whereas Berlare had only 1 car theft incident [12]. The data is not necessarily based on cities, it can be based on the residential area within a city. This holds true that much higher crime rates in a concentrated location compared to any other locations. It is called bad neighborhood.

To reduce the crime rates in a bad neighborhood, it makes sense that the authority should put more enforcement in this location. Moreover, the citizen should avoid this location as much as they can if they want to feel much safer. Evidently, Moura, et al. suggest that the existence of bad neighborhood phenomenon also occurs in the Internet world called "Internet bad neighborhoods" [57]. There are certain networks of Internet infrastructure that contain more malicious activities that other networks. For our preliminary analysis, we will adopt formal definition of Internet bad neighborhoods or Internet Badhoods by [57] which states:

> *"Internet bad neighborhood is a set of IP addresses clustered according to an aggregation criterion in which a number of IP addresses perform a certain malicious activity over a specified period of time"*

Several studies have suggested that the source of the Internet Badhoods tend to be concentrated in certain portions of Internet Protocol (IP) address space [67, 10, 5]. Moura, et al. suggest that Internet Badhoods do not always only based on network prefixes level (e.g. /24, /32, etc..) but it can be aggregated into several levels (ISPs, Countries) [57]. Moreover, Internet Badhoods may vary depending on which application exploited. While spam is most likely distributed all around the world, however, phishing Badhoods are most likely concentrated in developed countries (e.g. US) [57]. This suggests that phishing sites are required to have more reliable hosts in term of availability, while spams are not. We will see on Chapter 3 that consists of our preliminary analysis on phishing Badhoods. In the next section, we will study on the general phishing countermeasures in term of phishing detection and prevention.

2.6 CURRENT COUNTERMEASURES

There are various types of phishing countermeasures that implemented in different levels. Purkait has conducted an extensive research in reviewing these countermeasures which are available up until 2012 and their effectiveness [66]. He suggests that there is a classification of phishing countermeasures in separate groups and according to Purkait [66], these groups are listed as follow:

- Stop phishing at the email level
- Security and password management toolbars
- Restriction list
- Visually differentiate the phishing site
- Two facto and multi channel authentication

- Takedown, transaction anomaly detection, log files
- Anti phishing training
- Legal solution

In addition, Parmar, et al. suggests that phishing detection can be classified into two types; user training approach and software classification approach [62]. He illustrated a diagram and a table that summarizes phishing detection as countermeasures in a broad view [62]. They also argued the advantages and disadvantages of each category [62]. However, as our research mainly focuses in synthesizing phishing email with cialdini's six principles of persuasion [8], we will briefly discuss the most relevant phishing countermeasures such as restriction list group (i.e. Phishtank), machine learning approach (web-based phishing), properties or features in a phishing email, and anti phishing training group (i.e PhishGuru). In the last section of this chapter, we will explore the human factor in phishing attacks, how phishing email is engineered to gain recipient's trust in order to get a response from the unsuspecting victims.

2.6.1 *Phishing detection*

In this subsection, we will conduct a literature review which related to phishtank as restriction list and machine learning approach to detect spoofed website as phishing detection.

2.6.1.1 *Phishtank*

One of the common approaches to detect phishing attacks is the implementation of restriction list. As the name suggest, it prevents users to visit fraudulent websites. One of the efforts to achieve restriction list, is to derive phishing URLs from Phishtank. Phishtank is a blacklisting company specifically for phishing URLs and it is a free community web based where users can report, verify and track phishing URLs [61]. Phishtank stores phishing URLs in its database and is widely available for use by other companies for creating restriction list. Some of the big companies that are using Phishtank's data includes; Yahoo Mail, McAfee, APWG, Web Of Trust, Kaspersky, Opera and Avira. In this section, we will discuss how the current literatures have to do with phish data provided by Phishtank. The first step to achieve the list of relevant literatures regarding phishtank is by keyword search in Scopus online library. By putting "Phishtank" as a keyword search, it results in 12 literatures. The next step, we read the all the abstracts and conclusions of the resulting keyword search and we decided 11 literatures that are relevant to our research. Lastly, Table 2 summarizes the papers selected and its relevancy with Phishtank

Table 2: Summary phishtank studies

Paper title	First author	Country	Relevancy with phishtank
Evaluating the wisdom of crowds in assessing phishing website [56]	Tyler Moore	United Kingdom	Examine the structure and outcomes of user participation in Phishtank. The authors find that Phishtank is dominated by the most active users, and that participation follows a power law distribution and this makes it particularly susceptible to manipulation.
Re-evaluating the wisdom of crowds in assessing web Security [7]	Pern Hui Chia	Norway	Examine the wisdom of crowds on web of trust that has similarity with Phishtank as a user based system.
Automatic detection of phishing target from phishing webpage [44]	Gang Liu	China	Phishtank database is used to test the phishing target identification accuracy of their method.
A method for the automated detection of phishing websites through both site characteristics and image analysis [78]	Joshua S. White	New york, US.	Phishtank database is used to perform additional validation of their method. They also collect data from twitter using twitter's API to find malicious tweets containing phishing URLs
Intelligent phishing detection and protection scheme for online transaction [1]	P.A. Barraclough	Newcastle, United Kingdom	Phishtank features is used as one of the input of neuro fuzzy technique to detect phishing website. The study suggested 72 features from Phishtank by exploring journal papers and 200 phishing website.
Towards preventing QR code based attacks on android phone using security warning [81]	Huiping Yao	New Mexico, US.	Phishtank API is used for lookup whether the given QR containing phishing URL in the Phishtank database.
A SVM based technique to detect phishing URLs [29]	Huajun Huang	China	Phishtank database is used as validation resulting 99% accuracy by SVM method, plus the top ten brand names in Phishtank archive is used as features in SVM method.

Socio technological phishing prevention [26]	Gaurav Gupta	Australia	Analyze the Phishtank verifiers (individual/organization) to be used as anti phishing model.
An evaluation of lightweight classification methods for identifying malicious URLs [20]	Shaun Egan	Grahamstown, South Africa	Indicating that lightweight classification methods achieves an accuracy of 93% to 96% when trained data from Phishtank.
Phi.sh/$oCiaL: The phishing landscape through short URLs [6]	Sidharth Chhabra	Delhi, India	Phishtank database is used to analyze suspected phish that is done through short URLs.
Discovering phishing target based on semantic link network [76]	Liu Wenyin	Hong Kong	Phishtank database is used as test dataset to verify their proposed method (Semantic Link Network)

From our literature survey, we know that Phishtank is crowd-sourced platform to manage phishing URLs. For that reason Moore, et al. aims to evaluate the wisdom of crowds platform accommodated by Phishtank [56]. Moore, et al. suggest that the user participation is distributed according to power law. It uses to model data which frequency of an event varies as a power of some attribute of that event [40]. Power law also applies to a system when large is rare and small is common [3]. For example, in the case of individual wealth in a country, 80% of the all wealth is controlled by 20% of population in a country. It makes sense that in Phishtank's verification system, a single highly active user's action can greatly impact the system's overall accuracy. Table 3 summarizes the comparison performed by [56] between Phishtank and closed proprietary anti-phishing feeds[4]. Moreover, there are some ways to disrupt Phishtank verification system; submitting invalid reports accusing legitimate website, voting legitimate website as phish, and voting illegitimate website as not phish. While all the scenarios described are for the phishers' benefit, the last scenario is more direct and the first two actions rather subtle intention to undermine Phishtank credibility.

To put it briefly, the lesson of crowd sourced anti-phishing technology such as Phishtank is that the distribution of user participation matters. It means that if a few high value participants do something wrong, it can greatly impact overall system [56]. Also, there is a high probability that bad users could also extensively participate in submitting or verifying URLs in Phishtank.

3 http://kottke.org/03/02/weblogs-and-power-laws
4 The author conceals the identity of the closed proprietary company

Phishtank	Proprietary
10924 URLs	13318 URLs
8296 URLs after removing duplication	8730 URLs after removing duplication
Shares 5711 URLs in common 3019 Unique to the company feeds while 2585 only appeared in Phishtank	
586 rock-phish domains	1003 rock phish domains
459 rock phish domains found in Phishtank	544 rock phish domains not found in Phishtank
Saw the submission first	11 minutes later appear on the feed
16 hours later after its submission for verification (voting based)	8 second to verified after it appears
Rock phish appear after 12 hours appeared in the proprietary feed and were not verified for another 12 hours	

Table 3: Comparison summary [56]

2.6.1.2 *Machine learning approach in detecting spoofed website*

The fundamental of phishing detection system would be to distinguish between phishing websites and the legitimate ones. As we previously discussed, the aim of phishing attack is to gather confidential information from potential victims. To do this, phishers often prompt for this information through fraudulent websites and masquerade as legitimate institutions. It does not make sense if phishers created them in a way very distinctive with its target. It may raise suspicions with result of unsuccessful attack. To put it another way, while it might be true, we speculated that most of the phishing websites are mostly identical with its legitimate websites as target to reduce suspiciousness from potential victim.

In contrast of one of blacklisting technique we saw in Phishtank that heavily depend on human verification, researchers make use of machine learning based technique to automatically distinguish between phishing and legitimate either websites or email. Basically, machine-learning system is a platform that can learn from previous data and predict future data with its classification, in this case, phishing and legitimate. In order for this machine to learn from data, there should be some kind of inputs to classify the data, it is called features or characteristics.

Furthermore, there are also several learning algorithms to classify the data, such as, logistic regression, random forest, neural networks

URL	www.naturenilai.com/form2/paypal/webscr.php?cmd=_login	
Auto-Selected	name=www, name=naturenilai, tld=com, dir=form2, dir=paypal file=webscr, ext=php, arg=cmd, arg=login	
Obfuscation-Resistant	URL	len=54, n_dot=3, blacklist=1
	Domain Name	len=19, IP=0, port=0, n_token=3, n_hyphen=0, max_len=11
	Directory	len=14, n_subdir=2, max_len=6, max_dot=0, max_delim=0
	File Name	len=10, n_dot=1, n_delim=0
	Argument	len=11, n_var=1, max_len=6, max_delim=1

Figure 8: Example lexical features [42]

and support vector machine. However, as this particular topic is out of scope of our research, we will not discuss about the learning algorithm that is currently implemented. We will only introduce three features that are used in machine learning based detection.

There are vast amount of features to utilize machine learning to detect phishing attack. Literatures are selected by keyword search such as "phishing + detection + machine learning". We analyze three features: lexical feature, host-based feature and site popularity feature. Each of these features will be introduced briefly as follows.

- Lexical features

Lexical features (URL based features) are based on the analysis of URL structure without any external information. Ma, et al. suggest that the structure URL of phishing may "look" different to experts [48]. These features include how many dots exist, the length, how deep the path traversal do the URL has or if there any sensitive words present in a URL. For example the URLs https://www.paypal.com and http://www.paypal.com.example.com/ or http://login.example.com/www.paypal.com/, we can see that the domain paypal.com positioned differently, with the first one being the benign URL. Figure 8 shows an example analysis of lexical features in a phishing URL [42].

Lexical features analysis may have performance advantage and reduces overhead in term of processing and latency, since it only tells the machine to learn URL structure. 90% accuracy is achieved when utilizing lexical features combined with external features such as WHOIS data [42]. Egan, et al. conducted an evaluation of lightweight classification that includes lexical features and host based features in its model [20]. The study found that the classification based on these features resulted in extremely high accuracy and low overhead. Table 4 lists the existing lexical features that are currently implemented by two different studies [80, 45]. However, Xiang, et al.[80] pointed out that URLs structure could be manipulated with little cost, causing the features to fail. For example, attackers could simply remove embedded domain and sensitive words to make their phishing URLs look legitimate. Embedded domain feature examines whether a domain or a hostname is present in the path segment [80], for exam-

Haotian Liu, et al. [45]	**Guang Xiang, et al. [80]**
- Length of hostname Length of entire URL - Number of dots - Top-level domain - Domain token count - Path token count - Average domain token length of all dataset - Average path token length of dataset - Longest domain token length of dataset - Longest path token length of dataset - Brand name presence - IP address presence - Security sensitive word presence	- Embedded domain - IP address presence - Number of dots - Suspicious URL - Number of sensitive words - Out of position top level domain (TLD)

Table 4: Existing lexical features [45, 80]

ple, http://www.example.net/pathto/www.paypal.com. Suspicious URL feature examine whether the URL has "@" or "-", the present of "@" is examined in a URL because when the symbol "@" is used, the string to the left will be discarded. Furthermore, according to [80], not many legitimate websites use "-" in their URLs. There are also plenty of legitimate domains presented only with IP address and contains more dots. Nevertheless, lexical analysis would be suitable features to use for first phase analysis in a large data [20].

- Host based features

Since phishers often hosted phishing websites in less reputable hosting services and registrars, host-based features are needed to observe on the external sources (WHOIS information, domain information, etc.). A study suggests host-based features have the ability to describe where phishing websites are hosted, who owns them and how they are managed [48]. Table 5 shows the host-based features from three studies that are currently used in machine learning phishing detection. These studies are selected only for example comparison.

Each of these features does matter for phishing detection. However, as our main objective is synthesizing cialdini's principle with phishing emails, we will not describe each of these features in detail. It is noteworthy that some of the features are subset of another feature, for instance, autonomous system number (ASN), IP country and number of registration information are derived from WHOIS information. Nevertheless, we will only explain few of them that we assume the most crucial.

Justin Ma, et al.[48, 47]	**Haotian Liu, et al. [46][45]**	**Guang Xiang, et al. [80]**
- WHOIS data	- Autonomous system number	- Age of Domain
- IP address information	- IP country	
- Connection speed	- Number of registration information	
- Domain name properties	- Number of resolved IPs	
	- Domain contains valid PTR record	
	- Redirect to new site	
	- All IPs are consistent	

Table 5: Host-based features [48, 47, 45, 80]

1. WHOIS information: Since phishing websites and hacked domains are often created at relatively young age, this information could provide the registration date, update date and expiration date. Domain ownership would also be included; therefore, a set of malicious websites with the same individual could be identified.

2. IP address information: Justin Ma, et al. used this information for identify whether or not an IP address is in blacklist [47, 48]. Besides the corresponding IP address, it provides records like nameservers and mail exchange servers. This allows the classifier to be able to flag other IP addresses within the same IP prefix and ASN.

3. Domain name properties: these include time to live (TTL) of DNS associated with a hostname. PTR record (reverse DNS lookup) of a domain could also be derived whether it is valid or not.

- Site popularity features

Site popularity could be an indicator whether a website is phishy or not. It makes sense if a phishing website has much less traffic or popularity than a legitimate website. According to [80], some of the features indicated in Table 6 are well performed when incorporated with machine learning system.

1. Page in top search results: this feature originally used by [83] to find whether or not a website shows up on the top N search result. If it is not the case, the website could be flagged as phishy since phishing websites have less chance of being crawled [80]. We believe this feature is similar to Number of external links feature since both of them are implying the same technique.

Guang Xiang, et al. [80]	Haotian Liu, et al. [45]
- Page in top search results - PageRank - Page in top results when searching copyright company name and domain - Page in top results when searching copyright company name and hostname	- Number of external links - Real traffic rank - Domain in reputable sites list

Table 6: Site popularity features [80, 45]

2. PageRank: this technique is originally introduced by Google to map which websites are popular and which are not, based on the value from 0 to 10. According to [80], the intuitive rationale of this feature is that phishing websites are often have very low PageRank due to their ephemeral nature and very low incoming links that are redirected to them. This feature similar to Real traffic rank feature employed by [45] where such feature can be acquired from alexa.com.

3. Page in top results when searching copyright company name and domain/hostname features are complement features of Page in top search results feature with just different queries. Moreover, we believe they are also similar to Domain in reputable sites list feature since they are determining the reputation of a website. The first two features can be identified by querying google.com [80] and the latter feature can be obtained from amazon.com [45].

2.6.1.3 *Stop phishing at email level*

In order to stop phishing at email level, phishing email properties or features should be investigated. Chandrasekaran, et al. and Drake, et al [4, 19] specify the structure of phishing emails properties as follows:

1. Spoofing of online banks and retailers. Impersonation of legitimate institutions may created in the email level. Phishers may design a fake email to resemble the reputable company to gain users trust.

2. Link in the text is different from the destination. A link(s) contained in the email message usually appears different than the actual link destination. This portrays URL obfuscation and this method used to trick users to believe that the email is legitimate.

3. Using IP addresses instead of URLs. Sometimes phishers may hide the link in the message by presenting it as IP address instead of URL.

4. Generalization in addressing recipients. As phishing emails are distributed by large number of recipients, the email often is not personalized, unlike the legitimate email that address its recipient by personalized information such as the last four digits of account information.

5. Usage of well-defined situational contexts to lure victims. Situational contexts such as false urgency and threat are a common method to influence the decision making of the recipients.

Moreover, Ma, et al. experimented with seven properties to consider in a phishing emails consist of the total number of links, total numbers of invisible links, whether the link that appears in the message is different than the actual destination, the existence of forms, whether scripts exist within an email, total appearance of blacklisted words in the body and the total appearance of blacklisted words in the subject [49]. Based on this survey, we established phishing email properties as variables in order to classify our data in Section 5.1.3.

2.6.2 *Phishing prevention*

Phishing attacks aim to by-pass technological countermeasures by manipulating users' trust and can lead to monetary losses. Therefore, human factors take a big part on the phishing taxonomy, especially in the organizational environment. Human factor in phishing taxonomy comprised of education, training and awareness [24]. Figure 9 illustrates where human factor takes part on phishing threats [24]. User's awareness of phishing has been explored by several studies [33, 24, 21, 37, 34, 17] as preventive measure against phishing attack. According to ISO/IEC 27002 [24][11], it has been shown that information security awareness is important and it has been critical success factors to mitigate security vulnerabilities that attack user's trust. One approach to hopefully prevent phishing attack was by implementing anti phishing warning/indicator. Dhamija, et al suggest that users often ignore security indicators thus makes them ineffective [14]. Even if users notice the security indicators, they often do not understand what they represent.

Moreover, the inconsistency of positioning on different browsers makes them much difficult to identify phishing [36]. Evidently, Schechter, et al. pointed out that 53% of their study participants were still attempting to provide their confidential information, even after their task was interrupted by strong security warning [69]. Therefore, these suggest that an effective phishing education must be added as a com-

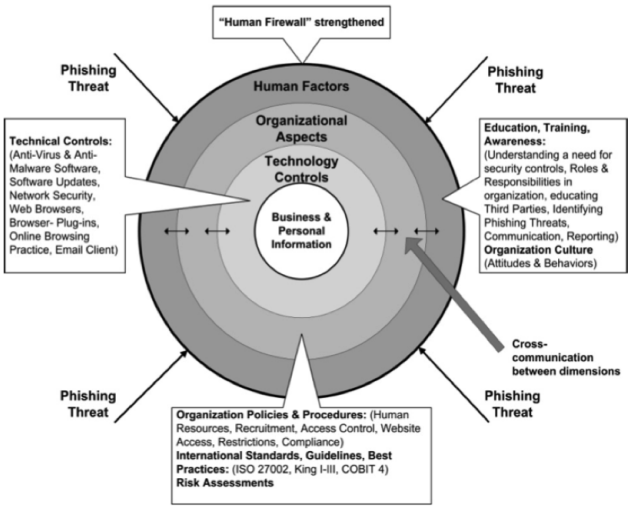


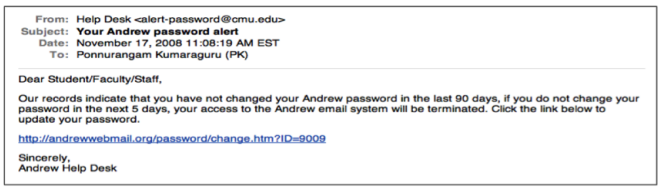
Figure 9: Holistic anti-phishing framework [24]

plementary strategy to complete technical anti-phishing measure as a strong remedy to detect phishing websites or emails.

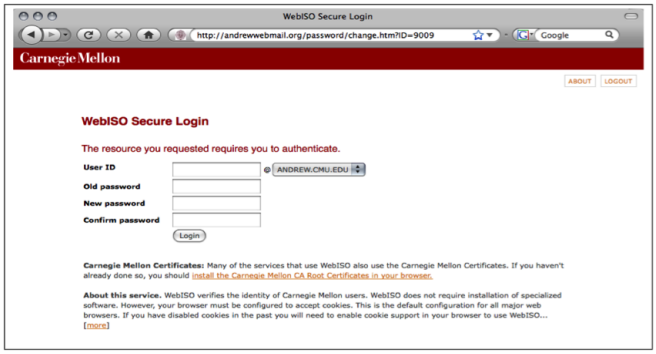
Phishing education for online users often by instructing not to click links in an email, ensure that SSL is present and to verify that the domain name is correct before giving information, and other similar education. This traditional practice evidently has not always effective [21]. One may ask what makes phishing education effective? A study suggests that in order online users to be aware of phishing threats, is to really engage them to so that they understand how vulnerable they are [50]. To do this, simulated phishing attacks often performed internally in an organization. Figure 10 shows a simulated phishing email and website carried out by Kumaraguru, et al. from PhishGuru [39]. As a result, this scenario puts them in the ultimate teachable moment if they fall for these attacks.

Phishguru is a security training system operated by Wombat security technology that teaches users not to be deceived by phishing attempts by simulation of phishing attacks[70]. They claimed Phishguru provides more effective training than traditional training as it is designed to be more engaging. Figure 11 illustrates how embedded phishing training was presented by PhishGuru.

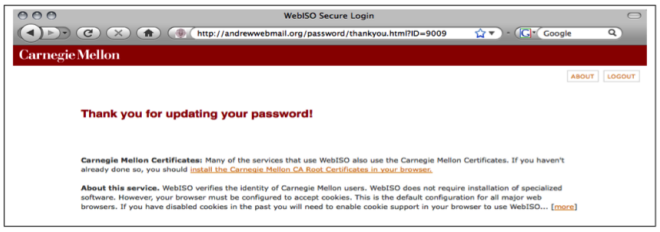
Kumaraguru, et al. investigates the effectiveness of embedded training methodology in a real world situation [39]. Evidently, they indicated that even after 28 days after training, users trained by PhishGuru were less likely to click the link presented in the simulated phishing email than those who were not trained. They also find that users who trained twice were less likely to give information to simulated fraudulent website than users who were trained once. Moreover, they argue that the training does not decrease the users' willingness to click on the links from legitimate emails; it means that less likely a trained



(a) simulated phishing email [39]



(b) simulated phishing website [39]



(c) simulated phishing message [39]

Figure 10: Simulated phishing attack [39]

Figure 11: Embedded phishing training [39]

user did a false positive when he or she requested to give information from true legitimate emails [39]. This suggests that user training strategy as an effective phishing education in order to improve phishing awareness especially in organizational environment.

2.7 HUMAN FACTOR

Phishing attacks generally aim to manipulate end users to comply phisher's request. Such manipulation in phishing attacks is achieved by social engineering. This means that human element is tightly involved with phishing. But how do phishers compose such deception? How come online users are gullible to these attacks?

Kevin Mitnick, who was obtaining millions of dollars by performing social engineering technique, is plausibly the best known person who had used social engineering technique to carry out his attacks. His book that titled "The art of deception: Controlling the Human Element of Security" [54] has defined social engineering as follows:

> "Using influence and persuasion to deceive people by convincing them that the attacker is someone he is not, or by manipulation. As a result, the social engineer is able to take advantage of people to obtain information, or to persuade them to perform an action item, with or without the use of technology."
>
> From his definition we can learn that people are the main target of the attack, specifies some of the important tools used by the attackers, such as influence and persuasion.

Cialdini suggests that there are six basic principles of persuasion [8], that is, the technique of making people grant to one's request. These principles include; *reciprocation, consistency, social proof, likeability, authority and scarcity*. Reciprocation is the norm that obligates individuals to repay in kind what they have received, return the favor or adjustment to smaller request [8]. Consistency is a public commitment

where people confirmed to commit publicly in the decision they have made [79][8]. Social proof is when people follow the behavior of their peer group, role models or important others because it is generally "fashionable" [79]. Stajano, et al. suggest people will let their guard down when everybody around them appears to share the same risk [71]. Likeability is when people giving their trust to the people they find attractive or credible [79, 8], when trust is achieved, compliance to grant a request may take place. While it is our human nature not to question authority, it can be used to cause fear, where people obey commands to avoid negative consequences such as losing a privilege or losing something valuable, fear of punishment, humiliation or condemnation [8, 79]. Stajano, et al suggest that scarcity is related to time principle, that is, when we are under time pressure to make important choice, we tend to have less reasoning to make decision [71]. We will use these principles as our foundation in synthesizing phishing email corpus with human factor.

Human as the “weakest link” in computer security has been exists and exploited for ages. And yet, security designers blame on users and whine “the system I designed would be secure, if only users were less gullible” [71]. Stajano, et al. stated that “a wise security designer would seek a robust solution which acknowledge the existence of these vulnerabilities as unavoidable consequence of human nature and actively build countermeasures that prevent this exploitation” [71]. With this in mind, the exploration of persuasion principles is congruent with our research goal. Cialdini’s six persuasion principles will be the foundation in our research.

3 EXPERIMENTAL EXPLORATION

Before we go into our main analysis on phishing email dataset, we would like to conduct a basic experimental exploration on phishtank and phishload datasets to know more about phishing bad neighborhood and phishing vs original websites. Basic investigation on both datasets will be conducted as initial experiments prior our main analysis to address the existence of phishing bad neighborhood and finding how many phishing webpages that maintains the "look" of legitimacy on their target.

3.1 PHISHING BAD NEIGHBORHOOD IN PHISHTANK

3.1.1 *Methods*

To examine whether phishing bad neighbors exist in phishing context, we established what tools we could use for our analysis. These tools include:

- www.yougetsignal.com
- www.majesticseo.com/research/neighbourhood-checker.php
- http://www.ip-address.org/reverse-lookup/reverse-ip.php
- http://www.my-ip-neighbors.com/
- ip-lookup.net
- Mac OS X network utility

We used these tools because they are freely available and capable of doing reverse IP lookup to find bad neighborhood in a particular IP address. We determined that the results given by these tools were sufficient for our basic analysis. We selected 10 valid and online phish URLs from phishtank in sequence on the date of 17 March 2014 and add 21 URLs more in sequence which were valid and online on the date of 24 March 2014.

After preparing 10 valid phishing URL we identified its Top Level Domain (TLD) and Top level IP address. The next step, we evaluated from which country are they from and counted how many characters it has. Next, we identified how many domains are hosted within a neighbors using reverse IP lookup. Lastly, we selected 10 random neighbors within its domain to be scrutinized whether they are harmful or legitimate.

3.1.2 *Results*

To fit the table onto the page and avoid a very huge and long table, we put the URLs separately in Table 31 of the appendix section.

Table 7: Phishtank URL analysis

URL ID	**TLD**[1]	**TL-IP**[2]	**N**[3]	**Random 10 neighbors**	**Country**	**URL Length**
1	munduslc.com	184.154.233.9	384	3 could not be resolved, 7 legitimate	Chicago, US	163
2	daff-inc.com	203.190.54.3	196	2 could not resolve, 8 legitimate	Jakarta, Indonesia	162
3	cntsiam.com	27.254.67.185	25	1 malware warning, 9 legitimate websites	Bangkok, Thailand	17
4	hockeyfollonica.com	194.184.71.7	297	9 404 error, 1 could not be resolved	Italy	165
5	douban.co.uk	193.61.190.231	18	4 legitimate, 1 phishing warning (douban.co.uk itself), 5 errors	UK	123
6	appsgeo.org	195.154.168.222	52	2 reported as phishing (including appsgeo.org), 1 reported untrustworthy (from WOT plugin), 6 legitimate, 1 error	France	164
7	fatihdabak.com	31.169.91.37	97	7 legitimate, 1 hacked, 1 phishing (WOT), 1 error	Turkey	176
8	edirnewebtasarimi.com	95.173.184.22	95	2 hacked, 2 error, 6 legitimate	Turkey	89
9	clientel-pl.com	209.17.116.6	362	10 legitimate	US	115
10	totalwhiteboard.com.au	203.84.238.17	224	10 legitimate	Australia	161
11	altervista.org	216.127.94.127	222	2 redirect to en.altervista.org (free host provider), 6 legitimate, 2 could not be resolved	US	66

1 Top Level Domain

2 Top Level IP address

3 No. of domain found

12	mytrickworld.com	192.254.71.149	102	1 could not be resolved, 2 empty pages, 7 legitimate	US	244
13	toughbook.cl	190.153.181.184	14	7 legitimate, 1 error, 2 request time out	Chile	162
14	doctorsantis.cl	200.63.97.50	426	2 could not be resolved, 2 404 error, 1 hacked, 5 legitimate	Chile	154
15	ankarabayanmodel.com	37.247.101.252	19	1 could not be resolved, 1 Suspended account, 1 expired domain, 6 legitimate, 1 poor reputation (WOT)	Turkey	105
16	theripe.tv	108.162.199.188	378	1 could not be resolved, 1 Under maintenance (warez), 8 Legitimate	US	185
17	yoursyours.com	204.77.0.196	18	1 warning (yoursyours.com, 1 could not be resolved, 8 legitimate	US	127
18	vidavallarta.com.mx	69.162.95.178	137	7 legitimate, 1 error (bandwidth limit exceeded), 2 could not be resolved	US	207
19	alvaroestrella.com	67.222.145.50	422	3 could not be resolved, 7 legitimate	US	117
20	no domain name	178.210.162.252	25	10 legitimate	Turkey	67
21	310bxgg.com	103.27.127.142	X[4]	hacked domain http://310bxgg.com/	Hong Kong	108
22	affordablebestwebsitehosting.com	173.214.178.24	25	4 suspected phishing (WOT warning), 1 suspended account, 1 could not be resolved, 4 legitimate	US	150
23	alwaysplotting.com	208.97.149.71	25	9 legitimate, 1 error (403 forbidden)	US	138
24	kcn.ru	193.232.252.56	X	X	Russia	124

4 X= No known domains hosted on this IP

25	cleanwheel.net	173.243.123.58	9	3 could not be resolved, 2 domain expired, 1 timeout, 1 error establishing a database connection, 2 legitimate	US	58
26	avatur.net	62.99.79.26	25	1 legitimate, 2 phishing warning, 6 access denied, 1 could not be resolved	Spain	146
27	vicfer.mx	198.27.68.106	X[5]. However, noticed that the last 5 URLs are from the same IP (198.27.68.106) So that it makes them neighbours		Canada	80
28	latitud-x.com.mx	198.27.68.106			Canada	104
29	carycar.com.mx	198.27.68.106			Canada	85
30	gentilee.com.ar	198.27.68.106			Canada	82
31	uniredmx.com	198.27.68.106			Canada	83

Based on our analysis we know that we have 31 phish URLs. We specify two categories that indicate less and higher probability to have a bad neighborhood. If one domain has 50% or more suspicious domains, it characterized to have higher probability of a bad neighborhood. In contrary, if one domain has less than 50% suspicious domains, it characterized to have less probability of a bad neighborhood. The data suggests that 3 of them are hacked domains (10%) while 11 domains (35%) have the chances of harmful instances. It makes 17 domains (55%) have less probability of a bad neighborhoods. In conclusion, while there are some chances of bad neighborhood, there are much more of legitimate domains attributed from a phishing domain as well.

5 X = No known domains hosted on this IP

3.2 FINDING DIFFERENCES OF PHISHING VS ORIGINAL IN PHISHLOAD DATABASE

3.2.1 *Methods*

To find differences between phishing webpage and its legitimate one, we utilized the differences of their HTML source codes. However, Phishtank does not have HTML source code of their phishing webpage, therefore, we decided to shift our analysis to Phishload database [51] which is collected from Phishtank as well. The phishload test database is a set of visited phishing websites and original websites that have been visited and scanned for testing purposes. Over a period of several weeks current phishing links from Phishtank have been collected and then been visited by a controlled instance of Firefox browsers. Derived from readme.txt files of phishload database, it has been created on 22 March 2013. It has 11215 URLs with 1185 non phishing URLs and 3718 assigned phishing pages [51]. Moreover, the tables of the database is already explained by Maurer on his website [51].

The methods to carry out this preliminary analysis are as follows:

1. Select all parent = NULL and site = paypal
    - SQL command: SELECT * FROM ‘websites‘ WHERE ‘name‘ = ’paypal’; does not result in anything/resulting in zero rows
    - SELECT * FROM ‘websites‘ WHERE ‘parent‘ IS NULL AND ‘urlBasedomain‘ = ’paypal.com’; result = 1 row
2. Save the ID and the field of HTML content
    - ID = 34 and HTML content are saved => paypal original.html
3. Find 20 phish websites where parent = ID
    - SELECT * FROM ‘websites‘ WHERE ‘parent‘= 34; Resulting in 1315 rows
    - SELECT * FROM ‘websites‘ WHERE ‘parent‘= 34 LIMIT 20; Resulting in 20 rows Export to SQL file => 20 phish paypal.sql

4. Compare legitimate paypal and the 20 fake paypal websites

All of the 20 examinations were done manually by finding the differences between the original web page against individual phishing web page stored in Phishload database.

3.2.2 *Results*

Through the HTML source code analysis, we may find malicious irregularities that ask victims to input their financial information or login credentials. Ludl et al. have defined page properties to characterize a web page before it can be analyzed for indication that might reveal it as a phishing site [46], of course in this experiment, all of the web pages are indicated as phishes according to Phishload. These properties are described as follows:

- Forms: Phishing website aims to trick users to input their sensitive information. Consequently, fake website needs some kind of forms as interface to contain those information. Web Forms may provide a good indicator to distinguish phishing and original.
- Input fields: Original web pages may have web forms with its input fields for users to input necessary information. Similarly, phishing web pages may have the same input fields as the original web page. However, the difference lies in where these information go to.
- Links: General properties of a web page is in its link structure. link(s) structure portrayed by not only links to other webpages, but also the link(s) which embedded in an image within a page. Ludl et al. argued that many phishing web pages contain links to the site they spoof, and evidently, often contain original elements from the web page they targeted [46].
- Script tags: Another good indicator to distinguish a phishing web page is to find out whether it has rouge JavaScript or not. There is a possibility that JavaScript may be used to by-passing Anti-phish so that it will not be detected as a phish [35, 46].

To illustrate, HTML source of the phish website row 1 was gathered and evaluated. It is clear that it is a phish website because inside the HTML source it has function called zzzz() that contains variables that manage the input of victim personal information.

```
function zzzz() {
var tes=true;
var first=document.fox.first_name.value;
var lasto=document.fox.last_name.value;
var doxa=document.fox.dob_a.value;
var doxb=document.fox.dob_b.value;
var doxc=document.fox.dob_c.value;
var numbex=document.fox.cc_number.value;
var email=document.fox.email.value;
var address=document.fox.address1.value;
var zipos=document.fox.zip.value;
var villss=document.fox.city.value;
var phonos=document.fox.phone.value;
var fvv=document.fox.cvv2_number.value;
}
```

We also found an explicit hack comment on the phish row 2:

```
<script type="text/javascript"> // This is an ugly hack until there is a reliable ondomready function if(typeof PAYPAL != 'undefined'){PAYPAL.core.Navigation.init(); }</script>
```

Based on our manual analysis, we found out that most of the HTML sources are tampered while maintaining the looks of legitimacy. The alterations consist of irregular scripts, irregular links, suspicious forms and inputs. Only two HTML sources that are completely different from the original source. 4 HTML sources are NULL so overall analysis will be N = 16. In conclusion, 87.5% of HTML sources are tampered while maintaining the appearance of original target and 12.5% are completely different from the original. A detailed of differences table of 20 phishing webpages vs original webpage will be shown in Table 32.

4

RESEARCH QUESTIONS AND HYPOTHESES

This chapter addresses the rationale of our main research questions and hypotheses to meet our research goal. We aim to answer these research questions by the analysis of data collected from a security organization based in Netherlands. First off, we wanted to know the characteristics of phishing email based on structural properties in our corpus.

RQ1: What are the characteristics of the reported phishing emails?

The characteristics of the reported phishing emails determined by the following parameters:

- How often phishing email include an attachment(s) and what specific attachment is the most frequent.
- Prevalent instructions
- Content characteristics
- The most targeted institutions
- The reasons that are frequently being used
- Persuasion principles characteristics
- Relationship between generic properties

To find out these characteristics, variable establishment of structural properties will be addressed in subsection 5.1.3.

Secondly, we wanted to know to what extent the involvement of persuasive principles are used in phishing emails and how relevant are they to the generic phishing email properties.

RQ2: To what extent the persuasive principles are used in phishing emails?

We established 16 hypotheses to indicate the relationship between generic properties and relevancy of persuasive principle to these properties. H8, H9, H10, H13, H14, H15 will partly answer RQ1 in respect to the relationship between generic properties and the rest will answer RQ2. We synthesize cialdini's principles with our dataset. In order to conduct the integration, we established our decision making to classify which persuasive elements that are exist in a phishing email. This process will be explained in subsection 5.1.5.

In our coding of cialdini's principles and phishing email dataset, we identified phishing emails with fake logos and signatures that may mistakenly regard them as legitimate by average internet users. For example in the context of phishing email, signature such as "Copyright 2013 PayPal, Inc. All rights reserved" or "Administrator Team" and Amazon logo were used to show the "aura of legitimacy". In the real world society, telemarketers and seller has been using authoritative element to increase the chance of potential consumer's compliance [74]. It means that they have to provide information in a confident way. Consumers will have their doubt if sellers unsure and nervous when they offer their product and services to consumers. This principle has been one of the strategies in a social engineering attack to acquire action and response from a target [60].

It is makes sense if government has the authority to compose laws and regulations and to control its citizens. Government sector includes court and police department also authorize to execute penalties if any wrongdoing happens within their jurisdiction. However, government may not have to be likeable to enforce their rules and regulation. Similarly, an administrator who control his network environment may behave in a similar fashion as government. Hence, in our dataset we hypothesize that

---

H1: There will be a significant association between Government sector and authority principle

H2: Phishing emails which targeting Administrator will likely to have authority principle

---

Similar to authority principle that may trigger reactance, scarce items and shortage may produce immediate compliance from people. In essence, people will react when their freedom is restricted about valuable matter when they think they are capable to make a choice among different options [64]. For example in phishing email context, an email from Royal Bank inform us that we have not been logged into our online banking account for a quite some time, as a security measure, they must suspend our online account and if we would like to continue to use the online banking facility, we have been asked to click the URL provided. Potential victim may perceives their online banking account as their valuable matter to access facility and information about their savings. Consequently, potential vicim may react to the request because of their account could be scarce and restricted. In the real world example, a hard worker bank customer who perceives money is a scarce item may immediately react when his bank inform him that he is in danger of losing his savings due to "security breach". We therefore hypothesize that

---

H3: There will be a significant correlation between Financial sector and scarcity principle

---

As we describe in our decision making consideration section, people tend to trust those they like. In a context of persuasion, perpetrators may find it more difficult to portray physical attractiveness, instead they are relying on emails, websites and phone calling [18]. To exhibit charm or charisma to the potential victims, perpetrators may gain their trust by establishing friendly emails, affectionate websites and soothing voice over the phone. In the phishing email context, Amazon praises our existence in an appealing fashion and extremely values our account security so that no one can break it. Based on this scenario, E-commerce/Retails sector may applied likeability principles to gain potential customers. We therefore hypothesize that

---

H4: Phishing emails which targeting E-Commerce/Retails will likely to have a significant relationship with likeability principle

---

Tajfel, et al. argued that people often form their own perception based on their relationship with others in a certain social circles [72]. This lead to affection of something when significant others have something to do with it. Social proof is one of the social engineering attacks based on the behavioral modeling and conformance [79]For example, we tend to comply to a request when a social networking site asks us to visit a website or recommends something and mention that others have been visiting the website as well. Thus, we hypothesize that

---

H5: Phishing emails which targeting Social networks will likely to have signification association with social proof principle

---

As we describe in our decision making consideration section, authority has something to do with "aura of legitimacy". This principle may lead to suggest the limitation on something that we deemed valuable. For example, a perpetrator masquerades as an authority and dressed as police officer halted us on the road, the perpetrator may tell us that we did something wrong and he will held our driving license if we do not pay him the fine. In the phishing email context, an email masquerades as "System Administrator" may tell us that we exceeded our mailbox quota, so the administrator must freeze our email account and we could re-activate it by clicking the URL provided in the email. Based on this scenario, we know that it has authority principle and also has scarcity principle. Therefore, we hypothesize that

---

H6: There will be a significant relationship between authority principle and scarcity principle

---

We often stumbled a group of people requesting to donate some of our money to the unfortunate people. Evidently, they would use physical attractiveness and kind words to get our commitment to support those people. Once they have got our commitment, they start asking for donation and we tend to grant their request and give some of our money to show that we are committed. Phishing email could be similar, for example, Paypal appreciates our membership on their system and PayPal kindly notifies us that in the membership term of agreement, they would performing annual membership confirmation from its customers. Based on this scenario, we know that the email has likeability principle and also has consistency principle. We would like to know if it is the case with phishing email in our dataset. Therefore, we hypothesize that

---

H7: The occurrence of likeability in a phish will impact the occurrence of consistency

---

We think it make sense if a fraudster tries to make his fake product as genuine as possible and hide the fabricated element of his product. There are also fraudster that did not make his product as identical as the legitimate product. In the phishing email context, we perceives fake product as URL in the email, phishers do not necessarily obfuscates the real URL with something else. Logically, such phishers do not aim to make a high quality of bogus email, rather they aim to take chances in getting potential victims that are very careless. This leads to our hypothesis that say

---

H8: Phishing emails that include URL will likely to be obfuscated

---

It is conspicuous from our knowledge if a sales agent tries to sell us a product, it would be followed by the request element to buy the product as well. However, it will not make sense if he tries to sell his product but he requests to buy another company's product. In other words, if we have something to sell, we do not just display our product without asking people's attention to look at our product. For example in phishing email context, phishers may include URL or attachment in the body of the email and also they may request unsuspecting victim to click the URL or to open the attachment. This leads us to two hypotheses which state

---

H9: Phishing emails that include URL will likely to request to click the URL

H10: Phishing emails that include attachment will likely to request to open the attachment

---

We sometimes find it suspicious if a person dressed as police officer that does not have a badge carried with him, unless he is a fake police officer. Consequently, a fake police officer may use a fake badge to build up even more "aura of legitimacy". Evidently, Cialdini suggests the increment of passerby who have stop and stare at the sky by 350 percent with suit and tie instead of casual dress [8]. Hence, we correlate that a person who wears police uniform and a fake badge in the real world context as authority principle and the presence of image in the phishing mail context. Another example, an email that masquerades Apple company, may clone Apple company logo or trademark to its content to increase the chance of potential victim's response or increase the "believability" if you will. Thus, we hypothesize that

---

H11: Phishing emails that have authority principle will likely to include an image to its content

---

Apart from the target analysis, we also investigate the reason why potential victim responds to phisher's request. Phishing email that implies our account expiration would have scarcity principle because the account itself may very valuable for us and is in danger to be expired or terminated. Therefore, we hypothesize that

---

H12: There will be a significant association between account related reason and scarcity principle

---

Similar from the hypothesis H12, it is sensible if a phishing email which contains account related reason such as reset password or security update, may tend to have a URL for the potential victim to be redirected towards phisher's bogus website or malware. Regardless of the target, based on our initial coding of the dataset we found that account related reason in a phishing email needs an immediate action greater than other reasons. Therefore, phishers may likely to include a URL to have immediate response from the potential victim. This leads to our hypothesis that say

---

H13: Phishing emails which have account related reason will likely to have URL

When a phishing email has document related reason such as review some document reports or court notice, it may tend to impersonate government to make the email sensible enough to persuade potential victim more than other targets. We therefore hypothesize that

---

H14: Phishing emails which targeting government sector will likely to have document related reason

---

Analogous with the hypothesis *H14*, it is make sense if a phishing email which has document related reason such as reviewing contract agreement or reviewing resolution case, would tend to have a file to be attached. We therefore hypothesize that

---

H15: Phishing emails which have document related reason will likely to include attachment

---

We think it is make sense if a phishing email which use HTML to present their email design may tend to increase the attractiveness to the potential victim. Consequently, unsuspected victim may respond to the request just because of the email design is attractive. Therefore, we hypothesize that

---

H16: Phishing emails which use HTML will have a significant association with likeability principle

---

5 DATA AND ANALYSIS

This chapter explain our research methodology and results in detail. We begin by explain the framework of our method, that consists of steps taken in order to get our results. By the end of this chapter, we present the results of our analyses to answer the research questions that we explained in chapter 4.

5.1 RESEARCH METHODOLOGY

As we illustrate in Figure 12, we processed our data into several steps. Firstly, we collect the data from a security organization in the form of suspected phishing email reports. Next, we performed our data categorization and we divided it into three categorization tasks. The next step, we determined what variables are needed for our analysis. In the next step, we executed our data classification into these variables, so that we could reconstruct into SPSS readable dataset. Lastly, we conducted our SPSS analysis to answer our hypotheses.

5.1.1 *Data collection*

The data is obtained from a security organization based in Netherlands that handles reports on online crime and fraud including phishing in the form of phishing emails, which were reported between august 2013 and december 2013. Since this research is based on self report, the data will be processed based on the self conception. Furthermore, the data consists of 8444 suspected phishing emails in total that we will be categorized and classified in the following subsection.

5.1.2 *Data categorization*

We manually categorized 8444 suspected phishing emails by sorting all the emails by the subject so that we could tell which emails were being distributed with the exact same content. We then examined individual email which has no or empty subject and determined in which language it was delivered. Initial categorization resulted as follows:

- 7756 suspected phishing in Dutch language
- 688 suspected phishing in English language

Within 688 suspected phishing emails in English group we further categorized based on phishing, spam and others. We label this pro-

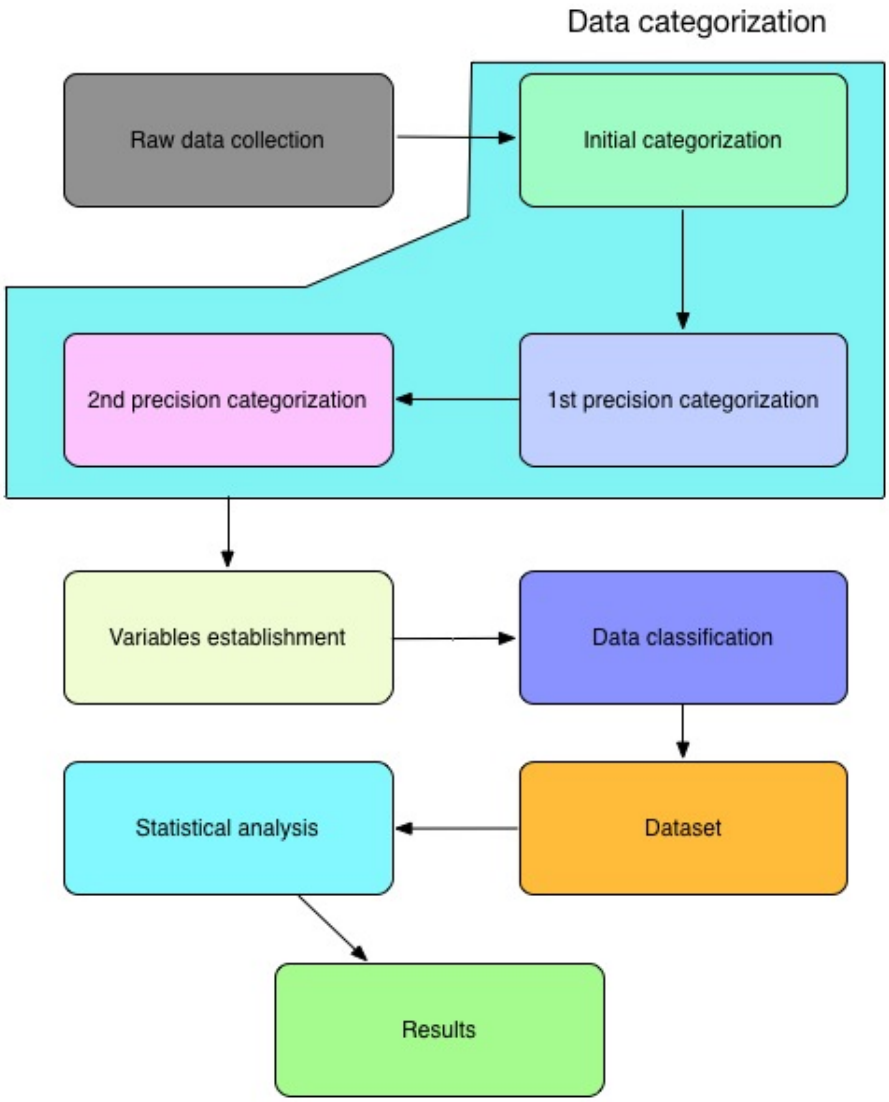


Figure 12: Research methodology diagram

cess as 1st level categorization. Phishing group is determined which emails were indeed phishing, spam group is specified which emails were commercial advertising and others group is established by the following guidelines:

- The email which has no content, if it was removed automatically by antivirus program. For instance, one email displayed only "INVOICE-983761039847.pdf" in the body. Therefore, we assumed the content of this email is removed by some sort of antivirus and we put this kind of email in "Others" group.
- The email which is presented in the language other than Dutch or English. We find some emails that use French language, thus we also put them in "Others" group.

The process is resulted as follows:

- 486 phishing
- 150 spam
- 52 others

To get a further precision from the 1st level categorization, we executed 2nd level categorization within 688 suspected phishing emails in English group, which basically repeated process of 1st level categorization with highly scrutinized. Basically, the 2nd level of categorization is to find if there is any mistake from the 1st categorization. This process is resulted as follows:

- 440 phishing
- 180 spam
- 50 others
- 18 legitimate

Interestingly, based on the outcome of 2nd categorization, we have 18 legitimate emails that were mistakenly reported as phishes or false positive. Although they are only 18 false positive, this suggest there are still misinterpretation of fraudulent email in our dataset. 30 phishes from the 1st categorization were moved to spam, 14 phishes were moved to legitimate and 2 "others" were moved to legitimate group. From this point we would only code 440 phishing emails to an excel sheet with necessary variables so that we could convert it into SPSS readable file. From 440 phishing emails, we performed query search which email that has a duplicated content. Query search was done by putting a query in mozilla thunderbird client, which has this feature. For example, a query string "Barclays Bank PLC" was taken from the body of the email and we found there are 11 emails which contain this string. Thus, we input "11" in "CounterSameContents", which we will explain in the next subsection. Query searches were resulted in 207 unique phishing emails. At the end, in total we analyze 207 unique phishing emails in our corpus.

5.1.3 *Variables and concepts*

As we study phishing email properties in subsubsection 2.6.1.3, variables are needed to be statistically analyzed in the SPSS. Based on our findings in the literature survey on phishing email properties, 25 Variables were created as part of the methodology processes prior data classification. Generic properties are depicted by the general structural phishing email properties except persuasion principles. These variables are explained as follows:

1. *Mail ID* : Unique ID [Scale measurement]
2. *Timestamps*: Implies the date and time when the email is being reported [Scale measurement]
3. *Attachments:* Indicates whether the phishing email has an attachment(s), if so, what kind of attachment
    a) PDF [0 = No, 1 = Yes]
    b) ZIP [0 = No, 1 = Yes]
    c) HTML [0 = No, 1 = Yes]
4. *Requests*: Implies the inquiry by the phishers in the contents
    a) ReqOpenAttachment; A request to respond by opening an attachment(s) [0 = No, 1 = Yes]
    b) ReqClickLink; A request to respond by clicking URL(s) [0 = No, 1 = Yes]
    c) ReqEmailReply; A request to respond by email reply [0 = No, 1 = Yes]
    d) ReqCallingByPhone; A request to respond by calling through phone [0 = No, 1 = Yes]
5. *Contents*: Indicates what elements are included in the body
    a) ContainHyperlink [0 = No, 1 = Yes]
    b) UseHTML [0 = No, 1 = Yes]
    c) IncludesImage [0 = No, 1 = Yes]
6. *ObfuscatedURL*: Specifies whether a phishing email has obfuscatedURL. We have explained URL obfuscation in subsubsection 2.6.1.3 [0 = No, 1 = Yes]
7. *CountMessageReporter*: A counter where the reporter includes an extra information with the minimum value 0. For instance, A reporter said "Geen spam, maar phishing!", we put a value 1 in this variable [Nominal measurement]
8. *Target*: Determined the target institutions
    a) TargetType [Values can be seen in Table 33]
9. *Reason*: Implies he reason why unsuspected victim must grant the phisher's request
    a) ReasonType [Values can be seen in Table 34]
10. *Cialdini's Principles*: Specifies what principle(s) the phishing email signifies. Coding consideration will be explained on subsection 5.1.5
    a) Reciprocation [0 = No, 1 = Yes]

b) Consistency [0 = No, 1 = Yes]
c) SocialProof [0 = No, 1 = Yes]
d) Likeability [0 = No, 1 = Yes]
e) Authority [0 = No, 1 = Yes]
f) Scarcity [0 = No, 1 = Yes]

11. *CounterSameContents*: A number that specifies how many emails are duplicated. The minimum value of this variable is 1, which indicates a unique email. Duplicated means the exact same text in the body. For example, value 2 indicates that there is (2-1) duplicated email with the same text in the body, value 3 means there are (3-1) duplicated emails.

We have established variables based on phishing email properties. Furthermore, we distinguished generic properties and persuasion properties. Generic properties of a phishing email is effected by these variables: attachments, requests, contents, obfuscatedURL, target and reason. On the other hand, persuasion properties effected by these variables: reciprocation, consistency, social proof, likeability, authority and scarcity.

5.1.4 *Data Classification*

We classified our data accordingly into our variables. As a result, usable dataset has been made to be analyzed. Data classification is conducted in a straightforward way. For example, a phishing email has a PDF attachment, we put "1" in our "PDFattachment" variable. Similarly, if phishing email has a hyperlink in the content, we put "1" in our "ContainHyperlink" variable. Lastly, we conducted integration on our data with Cialdini's six principles of persuasion will be discussed in subsection 5.1.5.

5.1.5 *Cialdini's principle and conception*

Part of our analysis, we tried to synthesize phishing emails dataset with Cialdini's principles titled "The science of persuasion". The decision making and the rationale in this process are achieved based our perspective of Cialdini's principles in the following details.

Reciprocation: The norm that obligates individuals to repay in kind what they have received. Return the favor. Adjustment to smaller request [8]. When a phisher sends an email containing a message that perceived as a request or obligation towards the recipient to "return the favor". It might be natural for an individual to feel "obligated" to return the favor for things or information that he/she is given and is deemed to be valuable. For example in the phishing email context, when PayPal has detected there are suspicious activities on our

account, we sometimes believe that PayPal has done a good job in detecting security risk on their system and we feel "obligated" to return the favor of that valuable information. Another example, if the sender gave the information that they have added "extra security" on their system so that we also feel obligated to grant their request.

Consistency: Public commitment. When people become psychologically become vested in a decision they have made [79]. When a phishing email contains a message that perceived to request recipient's "consistency" on a decision they have made. For example in the phishing email context, when a hotel agent asks us to review the payment details of our reservation that we have previously made, we might feel committed or agreed to review the payment details that has been given. Another example, if Facebook gave a link to change your password that you requested previously to change it. It might be not applicable to those who are not requesting password previously, but we believe it will impact to those who are committed to change the password previously.

Social proof: It occurs when people model the behavior of their peer group, role models, important others or because it is generally "fashionable" [79]. For example, when someone tells us that there are a hundreds of other people who use particular system, so we might want to agree to use it as well just because a lot of other people use it as well. Another example, when Facebook give information that someone wants to be our friend, and we knew who that someone is. We might tend to follow that request and click the link to accept the request.

Likeability: It occurs when people trust and comply with requests from others who they find attractive or are perceived as credible and having special expertise or abilities such as sports figures or actors they like [79].When a phishing email contains a message that attracts recipient to comply the sender's request based the reference on something or someone that likeable for the recipient. Cialdini [8] identified that people usually "trust those they like". For example, if someone is asking us to download and listen to a music that Michael Jackson made, we might be attracted to download and listen to it just because we happen to love Michael Jackson music. It is like someone is asking us to watch a concert and he/she said, "Coldplay will be there", if we are devoted fan of Coldplay, we might find it very interesting. Another example, when a sender gives compliments to us or committed to help us to safeguard our account from the hackers, we tend to think that the sender cares about our safety, which is good for us, and consequently it might attract us to comply with the sender's request.

Authority: It can be used to engender fear, where people obey commands to avoid negative consequences such as losing a privilege or something of value, punishment, humiliation or condemnation [79]. When a phishing email contains logo or image or signature or any-

thing that looks like legitimate institutions. It can be used to makes it look trustworthy so that the recipient might accept and obey the sender's request. For example, when an email is presenting somehow authentic looking signature like "Copyright 2013 PayPal, Inc. All rights reserved" or PayPal logo. Cialdini [8] suggests that authoritative persuasion could be achieved by only presenting "aura of legitimacy". Another example, when the content of the email stated that it is from "System Administrator" asking for password update. It would be not authoritative if only random people asking us to change our password.

Scarcity: Based on the principle of reactance, where people respond to perceived shortages by placing greater psychological value on perceived scarce items [79]When a phishing email contains a message that tells a recipient to react or respond to scarce/turns-into scarce items or things or privileges. For example in the phishing email context, if a sender tells us that he/she will suspend/deactivate/limit our account if not respond to his/her request, we might want to respond to their request because we are worried we will not able to access our account again or in other words our account become scarce or limited.

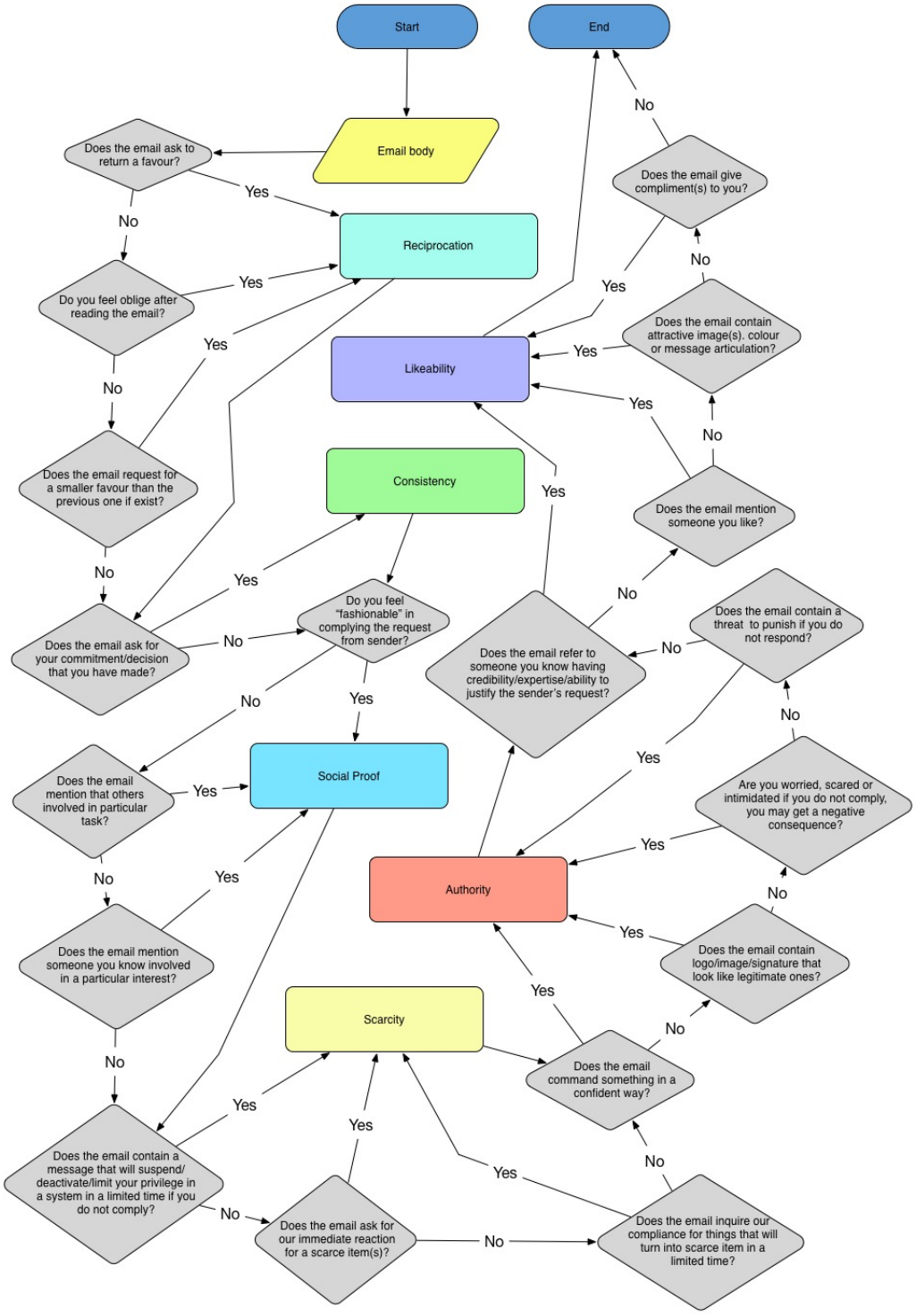


Figure 13: Integration pseudo-code of cialdini's principles

We have made a flowchart[1] in Figure 13 to illustrate our integration of cialdini's principles with the dataset.

1 Shapes and lines were created based on http://www.rff.com/how_to_draw_a_flowchart.htm

5.1.6 *Data entry and analyses*

In the previous section, we described the framework of our methodology in some detail. Until data classification, we have used Microsoft Excel to code our data. To perform the analyses, we transform the data into SPSS readable file. We initially record our data in 25 variables which could be expanded depending on our analyses, for example, selecting cases which have all instructions or contents.

The data has been analyzed by quantitative analysis from mainly three different viewpoints; general properties characteristics, persuasion principles characteristics and their relationships. We used frequency analysis to answer questions related to occurrences. For instance, we used frequency analysis to answer the most targeted institution in chapter 4. Furthermore, we used Pearson chi-square to test our hypotheses to discover if there is a significant relationship between two variables. If the resulted p-value is less than 0.05, 0.01, or 0.001, thus we are confident 95%, 99% and 99.9% respectively that the two chosen variables are having a significant relationship. By combining frequency analysis and chi-square test, we will see how they can answer our research questions in the next section. As our data is not continuous (i.e. interval or ratio) but nominal (i.e. categorical), therefore, we do not analyze our data by pearson correlation. However, to test the strength of association involves nominal variables, the appropriate measurements are using phi and Cramer's V. Phi is used for 2 by 2 table and Cramer's V can be used for more than 2 by 2 table. Since our data will be analyzed on 2 by 2 table, therefore, Phi measurements will be used. Values close to 0 indicate a very weak relationship, and values close to -1 or +1 indicate a very strong negative or positive relationship respectively.

5.2 RESULTS

In this section, we will discuss our findings in a considerable detail. We find that 36.2% of the total phishing emails have attachment(s) included within its content, On the other hand, 63.8% of them do not have attachment. In addition, we also look at what types of attachment does one has. Of the total emails having attachment, we find that 4% have PDF attachment, 78.7% have ZIP attachment, 12% have HTML attachment and 5.3% of them have unknown attachment. We are not sure what type of attachments they have, but we determined that the attachment element is still there if the request to open attachment within its content is presence. Table 8 illustrates our finding on attachment variables.

When we look at what are the instructions or requests used in the dataset. We find 202 emails or 97.6% of all total reported phishing

Type of attachment	Frequency	Percent
ZIP	59	78.7
HTML	9	12
Unknown	4	5.3
PDF	3	4
Total	75	100

Table 8: Attachment analysis

emails with clear instructions; whether it requests to click URL(s), request to open attachment, request to reply by email or request to respond by phone calling. we find that 37.2% of the total phishing emails request to open attachment, 52.7% of them request to click URL(s), 16.9% request for email reply and 4.3% request to call by phone. Moreover. One single email can request multiple requests. If we look deeper, we have 8 valid emails or 3.9% of all emails which have both request to open attachment and request to click URL. However, we do not find any email which request all instructions in the content. Table 9 illustrates our findings in respect of requests are used. In addition, of all phishing emails which have clear instructions, 54% request to click URL(s), 38.1% request to open attachment(s), 17.3% request for email reply and 4.5% request to respond by phone calling.

Request	Frequency	Percent
click URL	109	52.7
open Attachment(s)	77	37.2
Email Reply	35	16.9
call by phone	9	4.3

Table 9: Request analysis of all total emails (one email can contain more than one instructions so the total here does not sum up to 100%)

As we discussed before, we have also analyzed the content of phishing emails in our corpus. We look at whether it has URL(s), using HTML code or includes image(s) within its content. We find that 60.4% have URL(s) while 39.6% do not have URL. 66.2% of the emails use HTML code while 33.8% do not use HTML code within its content. We find 35.3% of them include image(s) while 64.7% do not include image. Table 10 highlights our findings in respect of content analysis. The percentage depicted of all total emails. If we look fur-

ther of all emails which utilized HTML, 120 emails or 87.6% of them have provided URL(s) and 73 or 53.3% of them include image(s). Furthermore, of all 73 emails that include image, 67 emails or 91.8% of them have provided URL(s). Based on this result, we know that one variable overlap with other variables. Therefore, the total percentage does not sum up to 100%.

Content	Frequency	Percent
utilizing HTML	137	66.2
URL presence	125	60.4
include Image	73	35.3

Table 10: Content analysis of all total emails (one email can contain more than one content variables so the total here does not sum up to 100%)

When we look at target classification table in Table 11, we find that financial sector is the most targeted sector and ISP is the least common target in our corpus. Furthermore, E-Commerce/retails, administrator and government allocated in the second, third and fourth respectively as the most targeted sectors. Figure 14 illustrates the pie chart of target analysis derived from Table 11. As one email does not have more than 1 targeted sector, therefore the total sums up to 100%. Note that, we initially had 92 targets in our corpus and we had to shrink it down into 10 target sectors in our data classification.

Target	Frequency	Percent
Financial	78	37.7
E-commerce/retails	40	19.3
Administrator	30	14.5
Government	14	6.8
Non-existence/individuals	13	6.3
Social media	11	5.3
Postal service	9	4.3
Travel agency	5	2.4
Industrial	5	2.4
ISP	2	1
Total	207	100

Table 11: Target analysis

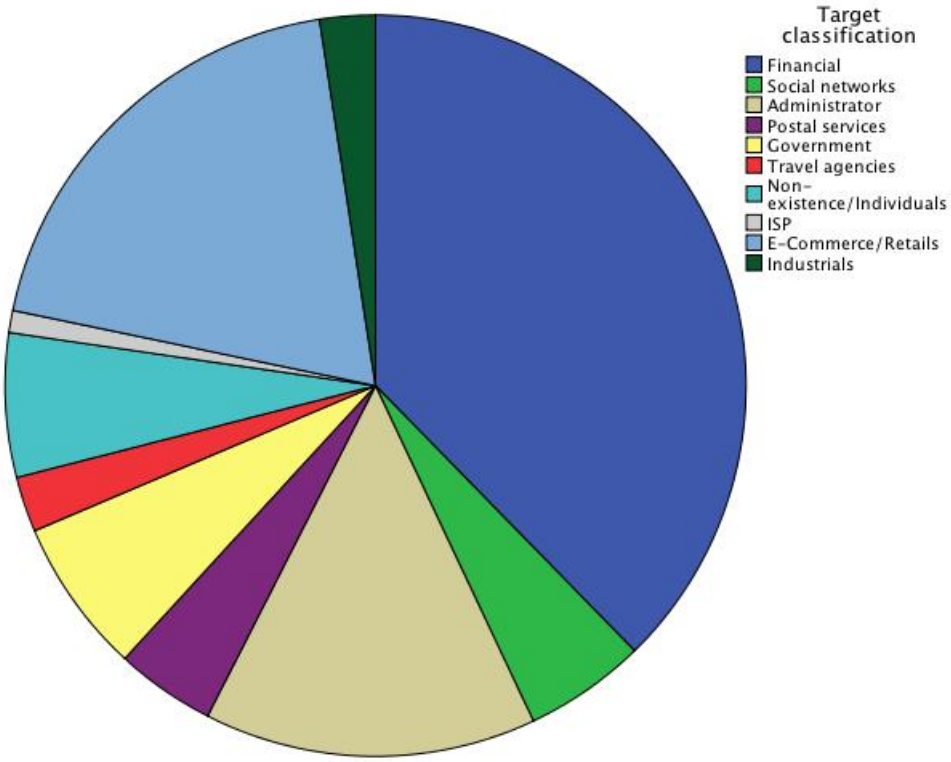


Figure 14: Target classification pie chart

When we look at what reasons are used in Table 12, we find that 48.8% of the total emails are account related, 25% are financial reason and 11.1% are document related reason. In addition, only 9.7% are product and services reason and only 4.8% are social reason. This suggests that account related is the most common pretext to manipulate recipients in our corpus while social is evidently the least common pretext. Figure 15 illustrates the bar chart of the reason classification derived from Table 12.

Reason	Frequency	Percent
Account related	101	48.8
Financial	53	25.6
Document related	23	11.1
Product and services	20	9.7
Social	10	4.8
Total	207	100

Table 12: Reason classification

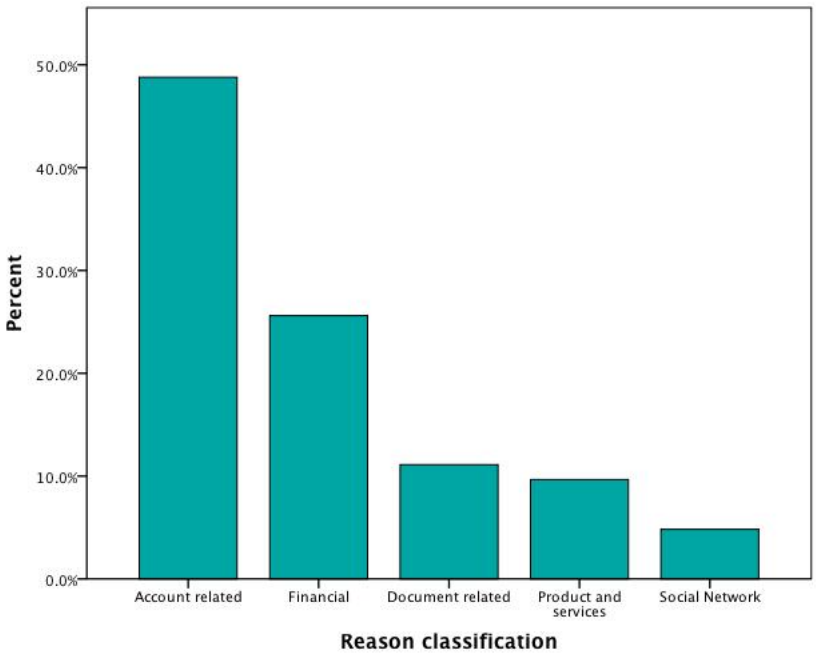


Figure 15: Reason classification bar chart

We look at the result of persuasion theory synthesis based on Cialdini's principles with our corpus. As we can see from Table 13, we find that 96.1% of the total phishing emails are having authority principle, which holds the most prevalent principle. Followed by scarcity

principle at 41.1%. While 21.7% of the total are having likeability principle, only 17.4% of them are having consistency principle. We find 9.7% are having reciprocation principle and 5.3% of them are having social proof principle. It is important to know that one email can have multiple principles. Therefore, the total percentage does not sum up to 100%.

Cialdini's principles	Frequency	Percent
Authority	199	96.1
Scarcity	85	41.1
Likeability	45	21.7
Consistency	36	17.4
Reciprocation	20	9.7
Social proof	11	5.3

Table 13: Persuasion principles analysis

Based on the result of persuasion principles analysis, we know that authority principle is the most used principle in our corpus. Now, we look at the relationship between government and authority principle to test hypothesis 1. We find that 95.9% of non-government targeted emails have authority principle and 4.1% of them do not impersonate government nor having authority principle. We find 100% of government targeted emails are having authority principle. On the other hand, we find 93% of all authority emails are non government targeted email and 7% of them are government targeted emails. Table 14 depicted the relationship between government targeted email and authority principle. Furthermore, of all phishing emails, 6.8% of them which have both authority and government targeted sector. A chi-square test was performed and we find that there is no significant association between government sector and authority principle, $X^2(1) = 0.604, p = 0.473$. since p is not less than 0.05, thus we reject hypothesis 1.

Table 14: Government sector and authority principle

Type of Target	Non-authority	Authority	N
Non-government	8	185	193
Government	0	14	14
N	8	199	207
Pearson chi-square		0.604	

When we look at the relationship between phishing emails which impersonate administrator and authority principle to test hypothesis 2. We find that 96.7% of administrator targeted emails have authority principle and 96% of non-administrator emails have authority principle. On the other hand, 85.4% of all authority emails are non administrator and 14.6% of them are administrator targeted emails. A chi-square test was performed and we find that there is no significant relationship between administrator target and authority principle, $X^2(1) = 0.027, p = 0.870$. Since p is not less than 0.05, therefore we reject hypothesis 2. Table 15 highlights the relationship between administrator sector and authority principle.

Table 15: Administrator sector and authority principle

Type of Target	Non-authority	Authority	N
Non-administrator	7	170	177
Administrator	1	29	30
N	8	199	207
Pearson chi-square		0.027	

Now we look at the association between financial sector and scarcity principle to test hypothesis 3. We find that 39.7% of all phishing emails which target financial sector have scarcity principle while 60.3% do not have scarcity principle. Furthermore, 41.9% of all non financial targeted emails have scarcity principle, while 58.1% of them do not have scarcity principle. On the other hand, 63.5% of scarcity emails is non financial targeted emails, inversely 36.5% of them is financial targeted emails. We performed a chi-square test and we find that there is no significant association between financial sector and scarcity principle, $X^2(1) = 0.090, p = 0.764$. Since p is not less than 0.05, thus, we reject hypothesis 3. Table 16 illustrates the relationship between financial targeted emails and scarcity principle.

Table 16: Financial sector and scarcity principle

Type of Target	Non-scarcity	Scarcity	N
Non-financial	75	54	129
Financial	47	31	78
N	122	85	207
Pearson chi-square	0.090		

We look at the association between phishing emails which are targeting e-commerce/retails and likeability principle to test hypothesis 4. We find that 20% of e-commerce/retails targeted emails have likeability principle. Furthermore, 22.2% of non e-commerce/retails targeted emails have likeability principle. On the other hand, only 17.8% of all likeability emails are e-commerce/retails targeted emails. A chi-square test was performed and we find that there is no significant association between phishing emails targeting e-commerce/retails and likeability principle, $X^2(1) = 0.088, p = 0.767$. Since p is not less than 0.05, therefore we reject hypothesis 4. Table 17 illustrates the relationship between e-commerce/retails targeted sector and likeability principle.

Table 17: E-commerce/retails sector and likeability principle

Type of Target	Non-likeability	Likeabillity	N
Non-ecomm/retails	130	37	167
Ecomm/retails	32	8	40
N	162	45	207
Pearson chi-square	0.088		

Now we look at the association between phishing emails targeting social networks and social proof principle to test hypothesis 5. We find that 18.2% of social media targeted emails have social proof principle. Furthermore, 4.6% of non social media targeted emails have social proof principle. On the other hand, 18.2% of all social proof email is social media targeted emails and 81.8% of them are not social media targeted emails. A chi-square test was performed and we find that there is no significant association between phishing emails targeting social networks and social proof principle, $X^2(1) = 3.823, p = 0.051$. Therefore since p is not less than 0.05, we reject hypothesis 5. Table 18

depicted the relationship between social media and social proof principle.

Table 18: Social media sector and social proof

Type of Target	Non-social proof	social proof	N
Non-social media	187	9	196
Social media	9	2	11
N	196	11	207
Pearson chi-square	3.823		

Furthermore, we look at the relationship between authority principle and scarcity principle to test hypothesis 6. Based on the result in Table 19, we find that 41.7% of authoritative emails have scarcity principle while 58.3% do not have scarcity principle. However, we find that 97.6% of all scarcity emails have authority principle and only 2.4% of them do not have authority principle. A chi-square test suggests that there is no significant relationship between authority principle and scarcity principle, $X^2(1) = 0.887, p = 0.346$. Thus, we reject hypothesis 6.

Table 19: Authority and scarcity

	Non-scarcity	Scarcity	N
Non-authority	6	2	8
Authority	116	83	199
N	122	85	207
Pearson chi-square	0.887		

We look at the relationship between likeability principle and consistency principle to test hypothesis 7. Based on our result in Table 20, we find that only 6.7% of likeability emails have consistency principle while 93.3% of them do not have consistency principle. In addition, we find that 20.4% of non likeability emails have consistency principle while 79.6% of them do not have likeability principle. On the other hand, 8.3% of all consistency emails are likeability emails while 24.6% of non consistency emails are likeability emails. A chi-square test suggests that there is a significant relationship between likeability principle and consistency principle $X^2(1) = 4.603, p = 0.032$. Phi measurement suggests a very weak negative (inverse) relationship at -0.149 that indicate as one variable increases, the other variable de-

crease. This suggests, that higher likeability principle, the less chance of consistency principle in a phishing email. Thus we accept hypothesis 7 that says the occurrence of likeability principle will impact the occurrence of consistency.

Table 20: Likeability and consistency

	Non-consistency	Consistency	N
Non-likeability	129	33	162
Likeability	42	3	45
N	171	36	207
Pearson chi-square	4.603*		

*p < 0.05 (significant).

Now we move on to find out the association between URL presence and obfuscated URL in our corpus to test hypothesis 8. Based on our result in Table 21, we find 76% of URL(s) are obfuscated while 24% are not. A chi-square test suggests that there is a highly significant association between URL presence and obfuscated URL, $X^2(1) = 115.191, p < 0.001$. Moreover, Phi measurement suggests a strong positive relationship at 0.746. This indicates a strong relationship between them. Therefore, we accept hypothesis 8.

Table 21: URL presence and obfuscated URL

URL	Not obfuscated	Obfuscated	N
Not exist	82	0	82
Exist	30	95	125
N	112	95	207
Pearson chi-square	115.191***		

*** p < 0.001 (significant).

We look at the relationship between URL presence and the emails which request to click URL in Table 22 to test hypothesis 9. We find 87.2% of phishing emails which have URL also requested to click it, while only 12.8% do not request to click. A chi-square test was performed and suggests that there is a highly significant relationship between URL presence and request to click URL, $X^2(1) = 151.034, p <$

0.001. Phi measurement suggests that they have a strong positive relationship at 0.854. Thus, this data supports hypothesis 9.

Table 22: URL presence and Request to click URL

URL	does not request to click URL	requests to click URL	N
Not exist	82	0	82
Exist	16	109	125
N	98	109	207
Pearson chi-square	151.034***		

*** p < 0.001 (significant).

Similarly, We look at the association between the email that includes attachment and the emails which request to open attachment to test hypothesis 10. Based on our result in Table 23, We find 96% of phishing emails which include attachment also requested to open it, while only 4% do not request to open the attachment. A chi-square test was performed and suggests that there is a significant relationship between URL presence and request to click URL, $X^2(1) = 174.079, p < 0.001$. Phi measurement suggests that they have a strong positive relationship at 0.917. Therefore, we accept hypothesis 10.

Table 23: Includes attachment and request to open attachment

Attachment	does not request	requests	N
Not exist	127	5	132
Exist	3	72	75
N	130	77	207
Pearson chi-square	174.079***		

*** p < 0.001 (significant).

We look at the relationship between authority principle and the emails which include an image(s) to test hypothesis 11. Based on our result in Table 24, we find 35.7% of authoritative emails include image(s), while 25% of non authority emails include image(s). On the other hand, 97.3% of emails which include image is authority emails and 95.5% of emails which do not include image(s) is authority emails. A chi-square test was performed and suggests that there is no significant relationship between authority principle and image

presence, $X^2(1) = 0.384, p = 0.535$. Thus, based on this result, we reject hypothesis 11.

Table 24: Authority and image presence

Cialdini's principle	does not include image	Includes image	N
Non-authority	6	2	8
Authority	128	71	199
N	134	73	207
Pearson chi-square	0.384		

Now we look at the association between account related reason and scarcity principle to test hypothesis 12. Based on the result in Table 25, we find 68.3% of account related phishing emails have scarcity principle, while 31.7% of them do not. In addition, we find 81.2% of scarcity emails have account related reason while 18.8% of them do not. A chi-square test was performed and suggests that there is a significant association between account related reason and scarcity principle, $X^2(1) = 60.535, p < 0.001$. Phi measurements suggests that they have a strong positive relationship at 0.541. Therefore, we accept hypothesis 12.

Table 25: Account related reason and scarcity

ReasonType	Non-scarcity	Scarcity	N
Not account related	90	16	106
Account related	32	69	101
N	122	85	207
Pearson chi-square	60.535***		

*** p < 0.001 (significant).

Furthermore, we look at the relationship between account related reason and URL presence to test hypothesis 13. Based on the result in Table 26, we find 78.2% of account related emails include URL and 63.2% of emails which include URL(s) are account related emails. Furthermore, 38.2% of total phishes are account related and include URL(s). A chi-square test was performed and suggests that there is a significant relationship between these two variables, $X^2(1) = 26.216, p < 0.001$. Phi measurement suggests that they have a weak positive relationship at 0.356. Therefore, we accept hypothesis 13.

Table 26: Account related reason and URL presence

ReasonType	URL does not exist	URL exists	N
Not account related	60	46	106
Account related	22	79	101
N	82	125	207
Pearson chi-square	26.216***		

*** p < 0.001 (significant).

Now we look at the relationship between document related reason and government sector to test hypothesis 14. Based on the result in Table 27, we find only 21.7% of document related reason phish emails were targeting government and inversely 78.3% of them are not targeting government. However, A chi-square test suggests that there is a highly significant relationship between these variables, $X^2(1) = 9.203, p = 0.002$. Phi measurement indicates that they have a weak positive relationship at 0.211. Therefore, we accept hypothesis 14.

Table 27: Document related reason and government sector

ReasonType	Non-government	Government	N
Not document related	175	9	184
Document related	18	5	23
N	193	14	207
Pearson chi-square	9.203**		

** p < 0.01 (significant).

Now we look at the relationship between document related reason and attachment variables to test hypothesis 15. Based on our result in Table 28, we find 78.3% of document related reason phish emails have attachment included, while 21.7% of them do not. A chi-square test suggests that there is a significant relationship between these variables, $X^2(1) = 19.783, p < 0.001$. Phi measurement indicates that they have a weak positive relationship at 0.309. However, the result still supports hypothesis 15.

Table 28: Document related reason and includes attachment

ReasonType	Does not include attachment	includes attachment	N
Not document related	127	57	184
Document related	5	18	23
N	132	75	207
Pearson chi-square	19.783***		

*** $p < 0.001$ (significant).

Lastly, we look at the association between HTML usage variable and likeability principle to test hypothesis 16. Based on the result in Table 29, we find 80% of likeability phish emails use HTML, while 20% of them do not use HTML code within its content. Furthermore, 37.7% of non likeability emails do not use HTML and 62.3% of them use HTML. On the other hand, 26.3% of emails which use HTML are likeability emails and 17.4% of total phishes use HTML and have likeability principle. A chi-square test suggests that there is a significant relationship between these variables, $X^2(1) = 4.904, p = 0.027$. Phi measurement suggests that have a very weak positive relationship at 0.154. However, we still accept hypothesis 16.

Table 29: use HTML and likeability

Content	non-likeability	likeability	N
Not use HTML	87.1	12.9	70
use HTML	73.7	26.3	137
N	162	45	207
Pearson chi-square	4.904*		

* $p < 0.05$ (significant).

6 DISCUSSION

In our attempt to discover the existence of bad neighborhood, we have conducted a basic exploration with the help of tools available freely on the Internet which capable of reverse IP and domain lookup. These tools can check an IP address or a domain for other IP addresses within the same web server or the same network. However, the shortcoming of this exploration is that we do not know how exactly these tools check the neighbors of an IP address. Therefore, our result is dependent on the information given by these tools.

Our result suggests that 10% of our dataset are hacked domains and 35% of them show the existence of phishing bad neighborhoods. However, 55% of them were not having bad neighborhoods. Therefore, we can conclude while there is a small chance of "bad neighborhood" in a phishing domain, phishers generally do not host their phish in one certain network of internet infrastructures. The consequence is that, it hinders the effort of security experts to detect phishing in a certain network of internet infrastructures.

When we look at phishing vs original experiment, it suggests 87% of the total HTML codes are maintaining the appearance to be the same as the target while 12.5% are completely different with its target. This suggests that most of the phishing web pages appear the same in terms of the layout as its target. With this in mind, considering to replicate a web page can be done with a little effort, however, fragments of suspicious codes can be spotted when the HTML source is closely examined. Our result does not imply that the phishers are lazy in terms of only replicating legitimate websites, but how much modification they did while maintaining the same "look" as their targets. This can help an automated phishing detection in HTML source level. Automated detection might be achieved firstly by creating a library that contains suspicious HTML codes that previously have been identified. Secondly, a toolbar that can extract HTML structures from a website and compare against the library. If a match found, then the web page likely to be a phish. However, this might be another field of research that needs to be explored.

6.1 RESEARCH QUESTIONS

At the beginning of our research, we have stated the following research questions that needed to be answered. In this section, we will shortly discuss our findings to answer our research questions.

What are the characteristics of reported phishing emails?

In chapter 4, we defined seven parameters to characterize the phishing emails in our dataset. Based on our findings, we can conclude in the following points:

- When attachment(s) is included in a phishing email, it is likely to attach ZIP or HTML file
- Requesting to click URL(s) is the most prevalent instruction in phishing emails
- Most of the phishing emails are using HTML code and provide URL(s)
- The financial sector is the most common target
- Most of the phishing emails use account related as a pretext
- Authority principle is the most used persuasion technique in phishing emails
- Phishing email which has account related as a pretext, is likely to include URL(s)
- Clear instructions to request an action from recipients; phishing emails which include attachment(s) are likely to request to open it and phishing emails which provide URL(s) are likely to request to click it.
- URL(s) in a phishing email is most likely different than the actual destination
- A government targeted phish is likely to have a document related reason
- Phishing email which has document related reason as a pretext, is likely to include attachment(s)

To what extent the persuasive principles are used in phishing emails?

To answer the second research question, we look at the relationships between the persuasive principles and the generic properties. With this in mind, we have established 10 hypotheses related to these relationships and we look wether the findings are consistent with these hypotheses. Table 30 summarizes the overview of verified hypotheses. Because almost all phishing emails use authority principle, therefore, this implies all phishing email properties related to authority principle are resulted in no significant relationship.

When we look at financial targeted sector and scarcity principle, we find that both financial and non financial targeted emails are less

chance to have scarcity principle. Apart from our hypothesis related to financial sector and scarcity principle, if we look deeper, administrator targeted emails are likely to have scarcity principle. In contrary, non administrator targeted emails are less likely to have scarcity principle. However, our finding suggests that the strength of association between administrator targeted emails and scarcity principle is weak.

The next finding on the relationship between e-commerce/retails targeted emails, it indicates that this sector contributes less number of likeability principle as both e-commerce/retails and non-ecommerce/retails targeted emails have high number of non-likeability principle. Similarly, our data suggests that there are no significant association between social media targeted emails and social proof principle.

Another observation whether likeability and consistency have a relationship, our data suggests that there is a significant association between them. Our result signifies that the higher likeability, the lower chance to have consistency principle. This support our hypothesis which says the occurrence of likeability will impact the occurrence of consistency. However, we find the strength of association between the variables is very weak.

When we look at account related phishing and scarcity, we find that there is a highly significant relationship between them. This means that if a phishing email uses an account related as a reason, it will likely to use scarcity principle as a persuasion technique. Moreover, the result suggests that account related reason and scarcity principle have a strong relationship.

Lastly, we find that there is a significant association between the use of HTML and likeability principle. This suggests that likeability phishing emails tend to use HTML code to persuade unsuspecting victim. However, their strength of association is very weak.

Overall, seven hypotheses in respect of persuasive principles are rejected and three of them are accepted. Based on this, we can answer our research question by three underlying perspectives. First, the extensive use of authority as persuasion technique in phishing emails as oppose to social proof technique. Secondly, scarcity principle will likely to be used when phishing emails are coming from administrator and using account related reason. Third, likeability principle affects the usage of HTML-based email and consistency principle.

Table 30: Overview of verified hypotheses

Hypotheses	Category	Accept	Reject
H1	A[a]		X
H2	A		X
H3	A		X
H4	A		X
H5	A		X
H6	A		X
H7	A	X	
H8	B[b]	X	
H9	B	X	
H10	B	X	
H11	A		X
H12	A	X	
H13	B	X	
H14	B	X	
H15	B	X	
H16	A	X	

a Related to persuasion principles
b Related to generic structural properties

6.2 CONCLUSION

Our research was aimed at understanding the characteristics of phishing emails considering persuasion techniques in the real world analysis which has not been done yet. The analysis consists of finding relationships between persuasion techniques and generic properties of phishing emails.

An important aspect of our research was that persuasion techniques have variety of strengths in respect of influence depending on the individual perspective. Although, we had made a flowchart in subsection 5.1.5 to model our decisions in terms of data coding, we believe

persuasion techniques are personal and difficult to find a consensual decision.

Nevertheless, by using parameters and hypotheses in chapter 4, we have been able to find the characteristics of phishing email considering persuasion techniques. Our approach has shown to be useful in identifying critical characteristics and relationships between generic properties phishing emails and persuasion techniques. This will help security experts to identify the underlying issue in phishing emails in terms of psychological aspect. As a security expert, the reflection from this research is that we can position ourselves as the phishers, which and how persuasion techniques are used to generate phishing emails. To regular user readers, we recommend to scrutinize when an email requires to quick respond to scarce or turn-into scarce items or things or privileges. We also do not recommend to be negligent when an email is coming from others who give attractive impression or perceived as credible and having special expertise. Lastly, an email which gives negative consequences, fear and authentic impression, does not always truthful. Therefore, should be doubted and questioned its legitimacy.

Continued research on persuasion techniques in phishing emails is required to stay ahead of the phishers, our method being a solid starting point in a real world analysis to identify the underlying issue in phishing emails.

6.3 LIMITATION

Although the research produces a conclusive results, our findings need to be assimilated in the backdrop of some limitations which arise due to the complex nature of our methodology and research environment. It is important for us to explain these limitations so that the readers can understand the findings of our research in the proper context. The first limitation is that we only get the data from one organization. Our study is totally dependent on the information documented by this organization, in which we do not know whether the sample data represents national-wide or represents a certain area or criterion. The second limitation is that, sometimes the email does not show the complete structures because the reporter forwarded a suspected phishing email as an attachment which removes essential element of it such as original attachment(s) included in the original email. This causes our study to be dependent on the reporter that reports to this organization as well. The third limitation is language barrier. A few number of suspected English-based phishing emails forwarded by the reporter along an information in Dutch language. It might be useful to know to understand the information provided by the reporter. The fourth limitation is that, our data classification is done by one person. It means the coding of the data into associated

variables could be inaccurate. While the data coding to the generic structural properties of phishing email could be justified, however, the data coding into the persuasive principles could be an issue in terms of accuracy. For example, one person claimed an email is attractive while another person claimed it is not. This introduces the greatest limitation to our research because it significantly impact our results. The fifth limitation is being unique dataset of the reported phishing emails. This resulted in smaller sample of data and therefore, gives a challenging task to find associations using pearson chi-square method.

6.4 FUTURE WORK

Despite our considerable limitations, a recommendation as a follow-up study with a larger sample encompassing more security organization as a data source and extended time period of reported phishing emails would be extremely desirable in order to test our findings. We feel that any future research along this line will find our work to be a useful starting point. Furthermore, we also recommend to add resilient validation in data classification in terms of persuasive principles by involving several people to have an objective decision. It is also interesting to identify authority principle in regular mails to make an objective perspective compare with phishing emails. Lastly, as we understand that persuasion principles in phishing email have some influence in user's decisions, therefore, it is interesting if the future research can build a simple game in terms of persuasion awareness to grab user's attention in making the right decision. For instance, the flowchart in Figure 13, can be adapted to "snakes and ladders" game to alert users of the presence of persuasive principles in an email they received.

A

PRELIMINARY ANALYSES APPENDIX

A.1 PHISHING URLS FROM PHISHTANK

Table 31: Phistank URL list

URL ID	Phishing URL
1	http://update-mypaypal.woa.wa.directtosignin-cgi-sys-defaultwebpage.cgi.defaultwebpage.cgi.munduslc.com/7d119be5b314a9a159244f884bc87ad0/36fd4df0d83094c6d466fdfdc5ad4aec/
2	http://daff-inc.com/PayPal/cgi-bin/webscr%3fcmd=_login-submit&dispatch=5885d80a13c0db1f8e263663d3faee8d8cdcf517b037b45005cf5d4eda3b985b/f4c476e425cd92c31d6d6452b0ac80b3/
3	http://cntsiam.com/logs/
4	http://www.hockeyfollonica.com/app2/media/bearleague/events/windhoek_tours/7b4b770d7284f852d17dbd7fe3b3154f/validate.php?cmd=53026&dispatch=68c92e699bc27f49aef2aaa5f3293d38
5	http://douban.co.uk/ss/e9023fd16f4f0d79785e91f4b06f6c46/?cmd=_home&dispatch=5885d80a13c0db1f8e&ee=9220c39c535c3317b3743e915aef3adc
6	http://appsgeo.org/paypal/mmp/web.php?#/_flow&SESSION=54cc48e97ad73aaa2519dbb379719301FpG7mo2DssMkja2121545487KJJ9ecd33e80218623592ca5d52281eb16eHHG5548782121548LLOp
7	http://paypal.com.cgi.bin.webscr.cmd.login.submit.15.cgi.bin.webscr.cmd.login.submit.15.cgi.bin.webscr.cmd.login.submit.15154.fatihdabak.com/sc/y/rev/488d0f6f819beb02b29791e111576dec/
8	http://edirnewebtasarimi.com/help/support/help/PP-1124-075-998/00b056620c4cc49fd79b6cb5aa773b0b/
9	http://www.clientel-pl.com/pl/b1999504af89588d08f4679d126f7720/scr.php?cmd=53026&dispatch=61ea2516a84f51c22d86a8fb0151b008
10	http://totalwhiteboard.com.au/.pp/0053d4ae3e2c78154d29d413c1236341/webscr.php?cmd=_login-run&dispatch=5885d80a13c0db1f998ca054efbdf2c29878a435fe324eec2511727fbf3e9efcd8
11	http://classiclogin.altervista.org/-/dados/time/AtualizandoDiaenoite2014/

12	http://ssl.paypal.secure.your.billing.information.mytrickworld.com/update-your-billing-information/8db3caa65cd255d3ae984b35c683952d/Security/Update/Account/Login/?cmd=_login-run&dispatch=e04a132adbe8a628371887da515b33e9e04a132adbe8a628371887da515b33e9
13	http://paypal.com.update.account.toughbook.cl/8a30e847925afc5975161aeabe8930f1/?cmd=_login-run&dispatch=70d1e179bda95563c92cdbb41bd380f670d1e179bda95563c92cdbb41bd380f6
14	http://paypal.com-cgi-bin.webscr.cmd.flowsession.home.locale.en-update.doctorsantis.cl/?cmd=_home&dispatch=5885d80a13c0db1f8e&ee=c2fdde4e9dbbb604104ee20987ae47db
15	http://paypal.ankarabayanmodel.com/PP/?cmd=_home&dispatch=5885d80a13c0db1f8e&ee=37c89453f89e8330ec833a3eb8ca0a77
16	http://www.theripe.tv/wp-content/plugins/justified-image-grid/languages/EN/c40ab55b0b2d4614ef0980c72fbe6007/?cmd=_home&dispatch=b47f8b3f68e40295c379148eb5c7a257b47f8b3f68e40295c379148eb5c7a257
17	http://www.yoursyours.com/templates/ilu/a/e5ad280236ce655dc34f3dedab589e97/scr.php?cmd=53026&dispatch=5b9bcc80064eac725312483dad5f6d46
18	http://www.re-update-your-information-1qs5dc1qs5941q5sf1sqf1qs5.vidavallarta.com.mx/reactivation/e66aea1c3732c4e6f5528af492d34ca0/?cmd=_home&dispatch=adcc48290abfe8c419eebf1df1ac7373adcc48290abfe8c419eebf1df1ac7373
19	http://alvaroestrella.com/secure/webapps/mpp/home/?cmd=_home&dispatch=5885d80a13c0db1f8e&ee=805eb67f9400bc03e8daa639613a16f7
20	http://178.210.162.252/~zeedee/9a70c0acdb2584924f5d0/web.php?#/confirm.php
21	http://310bxgg.com/aa/index.php?cmd=_home&dispatch=6e2b830562361bda74cc627c58f7e9306e2b830562361bda74cc627c58f7e930
22	http://smak.affordablebestwebsitehosting.com/~wxacad99/modules/fr/PayPal.fr/c.html?webscr?cmd=_login-done&login_access=2265929062
23	http://alwaysplotting.com/mokhtarhome/989e5e37528c1cbab8a0e6410ea45ee8/?cmd=_home&dispatch=5885d80a13c0db1f8e&ee=016b39232ee2655d5281db69fdf27aca
24	http://kgmu.kcn.ru/gigitru/images/banners/pp/webapps/mpp/?cmd=_home&dispatch=5885d80a13c0db1f8e&ee=eb70143576d5f26d479f02e2637fe227
25	http://www.cleanwheel.net/images/pics/informationen3/initsec.html
26	http://www.avatur.net/admin/cx/9772515528495d59759dc10765940daa/?cmd=_login-run&dispatch=f9ec5d6a8a348bce8818eb32e1b1d3eff9ec5d6a8a348bce8818eb32e1b1d3ef
27	http://www.vicfer.mx/~screamvi/PayPal/1bfc5540549a80993495b28f7a7b4d07/AccountLogin.php

28	http://www.latitud-x.com.mx/~screamvi/PayPal/1bfc5540549a80993495b28f7a7b4d07/AccountLogin.php?Userid=fxwspckm5
29	http://www.carycar.com.mx/~screamvi/PayPal/1bfc5540549a80993495b28f7a7b4d07/AccountLogin.php
30	http://gentilee.com.ar/~screamvi/PayPal/1bfc5540549a80993495b28f7a7b4d07/AccountLogin.php
31	http://www.uniredmx.com/~screamvi/PayPal/1bfc5540549a80993495b28f7a7b4d07/AccountLogin.php

A.2 HTML CODE ANALYSIS OF PHISHING VS ORIGINAL IN PHISHLOAD

Table 32: Differences phishing webpage vs legitimate website; target: PayPal

No.	What's changed/added?	Details
1	Irregular scripts	- Creation of function zzz() contains variables that holds input value from a form <form onsubmit="return zzzz()" class=" edit" action="getinfo.php" method="post" name="fox">. - There are 17 script tag in the original paypal, whereas there are only 5 script tags - If... else.. logic is added to manage how a user inputs his information
	Irregular link tag	- We are not sure whether the link tag is added by the phisher or it refers to the original paypal with the different source (PayPal has changed its web page from time to time, it means that the source also change periodically) - Most of the links are redirected to pass.php
	Suspicious form	A form used to ask user to submit his creditcard information, upon clicking, it will parse the information to variables hold by zzz() function
2	Html language	Lang=en whereas the original has lang=de
	Country difference	It based on Algeria PayPal country code DZ whereas the original has DE country code

	Irregular scripts	- Script to measure anticlickjack is not present. - Script addition clearly stating that it is a hack script at line 91 - s.prop1="p/gen/login-processing"; s.pageName="p/gen/login-processing::_login-processing"; whereas the original has s.prop1="xpt/Marketing_CommandDriven/homepage/MainHome"; and no s.pageName variable - Function scOnload() is present in the original whereas it is not present
3	Title differences	It has different title than the original
		Paypal object redirected based on 20101108 whereas in the original phishload database is based on 20120210
	Irregular scripts	- Anticlickjack script is not present - New script addition is present which hide from JavaScript-challenged browsers. - More new script addition is present to manage PayPal login flow <script type="text/javascript">PAYPAL.common.loginflow = 'p/gen/login';.... - s.prop1="p/gen/login";
	Css object difference	Paypal object redirected based on 20101108
	Suspicious form	It has <form action="websrc.php" name="login_form" method="post"> as user input for PayPal username and password
6	Different title	Italian vs German language
	Suspicious Form	It has <form action="error_login.php" name="login_form" method="post"> which ask for paypal credentials and redirect them to the wrong action
	Irregular scripts	- It has script to hide from javascript challenge browser. - It has script <script type="text/javascript">PAYPAL.common.loginflow = 'p/gen/login'; contains malicious operation - s.prop1="p/gen/login"; - It does not have the function scOnload()
	Flash object	It seems that there is an additional div tag at the bottom and I think it contains flash object.

7	URL encoding	It has weird URL encoding inside script tag document.write(unescape(. . . . I think the URLencoding inside the document.write is similar with the non URLencoding. So the rest of the result referred to the URLencoding will be similar as well.
	Suspicious form	If I decode the URL encoding, this form is present <form method="post" name="login_form" action="error_logins.php"> The form above exists again in the non encoding html
	Irregular scripts	- Script to hide from javascript challenged browser - The script that contains YAHOO.util.Event.addListener also quite different from the original, yet it exists again at the end - The tag <noscript></noscript> is removed - There is no function scOnload()
8	Different language	Lang=de vs. lang=en
	Suspicious Form	<form action="processing.php" name="login_form" method="post"> which ask for login email and password
	Irregular scripts	- Script to hide from javascript challenged browser - Suspicious script YAHOO.util.Event.addListener
	Flash object	flash object is added at the bottom
9	Title difference	“Log in” title page
	Irregular scripts	Function validateFormOnSubmit(theForm)
	Suspicious Form	form onsubmit="return validateFormOnSubmit(this)" method="post" action="cnd_pay.php">
	Suspicious link tag	All the links are redirect to itself / no absolute URL path
	Based on individual examination I would say that the web page is completely different than the original	

10	Suspicious Form	<form action="error_login.php?cmd=_login-run&dispatch=5885d80a13codb1f998ca054efbdf2c29878a435fe324eec2511727fbf3e9efcd8" name="login_form" method="post"> asking for login email and login password
	Irregular input tag	- <input type="submit" class="button primary" value="Log In" name="submit.x" /> - <input type="hidden" name="operating_system" value="Windows" /><input type="hidden" id="flow_name" name="flow_name" value="p/gen/login" /> - <input type="hidden" id="bp_ks2" name="bp_ks2" /><input type="hidden" id="bp_ks3" name="bp_ks3" /><input type="hidden" name="flow_name" value="p/gen/login" />
	Irregular scripts	- Script to hide from javascript challenged browser - s.prop1="p/gen/login"; - no function scOnload()
	Flash object	flash object is added at the bottom
11	Irregular scripts	- src="/js/lib/yui/animation.js - Script anticlickjack removed - Script PAYPAL.util.lazyLoadRoot removed - Great amount of scripts are removed - Suspicious script <script type="text/javascript">if(typeof PAYPAL != 'undefined'){ PAYPAL.core.Navigation.init(); }</script> - Suspicious script s.prop1="p/gen/login-processing"; - Function scOnload removed - Suspicious script setTimeout("location.href =... at the bottom
12	Irregular input	- <input type="hidden" name="flow_name" value="p/gen/login" /> - <input type="hidden" value="ok" name="login_cmd" /> - <input type="hidden" value="" name="login_params" />

	Irregular scripts	- <script src="files/globaloo.js" type="text/javascript"> - script that hides from javacscript challenged browser - anticlickjack script is removed - <script src="files/pp_jscod.js" type="text/javascript"></script> - s.prop1="p/gen/login"; • function scOnload removed
	Suspicious Form	<form action="" name="login_form" method="post"> which ask user input for login credential
	Suspicious link	some of the links are redirected to itself / not absolute URL path
	Flash object	flash object is added at the bottom
15	Irregular input	<input type="hidden" name="flow_name" value="p/gen/login" /> • <input type="hidden" value="" name="login_cmd" /><input type="hidden" value="" name="login_params" /><fieldset>
	Suspicious form	<form action="error_login.php?cmd=_login-run&dispatch=5885d80a13codb1f998ca054efbdf2c29878a435fe324eec2511727fbf3e9efcd8" name="login_form" method="post"> that asks login credential
	Irregular scripts	- script that hides from javascript challenged browser - anticlickjack script is removed - suspicious script <script type="text/javascript">if(typeof PAYPAL != 'undefined'){ PAYPAL.core.Navigation.init(); }</script> - suspicious script <script type="text/javascript">PAYPAL.common.loginflow = 'p/gen/login'; - s.prop1="p/gen/login"; - function scOnload is removed
	Flash object	flash added at the bottom (perhaps it is from the latest legitimate paypal website)
16	Suspicious form	<form action="Submit.php" name="login_form" method="post">

	Irregular scripts	- script that hides from javascript challenge browser - YAHOO.util.Event.addListener script - s.prop1="xpt/Marketing_CommandDriven/ homepage/IndividualsHome"; - function scOnload() is removed - YUE.addListener script at the bottom
	Irregular links	href="#content
17	Suspicious Form	<form action="processing.php" name="login_form" method="post">
	Irregular scripts	script that hides from javascript challenge browsers function scOnload is removed
	Irregular links	href="#"
18	Irregular input	- <input type="hidden" name="flow_name" value="p/gen/login" /> - <input type="hidden" value="" name="login_cmd" /> - <input type="hidden" value="" name="login_params" />
	Irregular from	<form action="error_login.php" name="login_form" method="post">
	Irregular script	- script that hides from javascript challenge browsers - anticlickjack script removed - PAYPAL.tns.loginflow script - PAYPAL.common.loginflow = 'p/gen/login' added - s.prop1="p/gen/login"; - function scOnload is removed
	Flash object	Flash added at the bottom
19	Irregular script	- <script type="text/javascript"> if (parent.frames.length > 0) {top.location.replace(document.location); }</script> - Script that hides from javascript challenge browsers - <script type="text/javascript">PAYPAL.common.loginflow = 'p/gen/login'; - s.prop1="p/gen/login"; - function scOnload is removed
	Irregular links	<li class="login">Einloggen
	Suspicious form	<form action="websrc.php" name="login_form" method="post">

	Irregular input	<input type="hidden" name="flow_name" value="p/gen/login" />
	Flash object	flash added at the bottom
20	Suspicious form	<form action="asu.php" name="login" method="POST">
	Irregular link	all the links are redirected to itself
	Irregular script	<script type="text/javascript" language="JavaScript">
	Based on individual examination we would say that the web page is completely different than the original	

A.3 TARGET TYPES

Value	Label
1	Financial
2	Social networks
3	Administrator
4	Postal Services
5	Government
6	Travel agencies
11	Non-existence/individuals
23	ISP
24	E-Commerce/Retails
26	Industrials

Table 33: Target classification

A.4 REASON TYPES

Value	Label
1	Account related
2	Social network
3	Financial
4	Product and services
5	Document related

Table 34: Reason classification

BIBLIOGRAPHY

[1] PA Barraclough, MA Hossain, MA Tahir, Graham Sexton, and Nauman Aslam. Intelligent phishing detection and protection scheme for online transactions. *Expert Systems with Applications*, 40(11):4697–4706, 2013. (Cited on page 20.)

[2] Mark Blythe, Helen Petrie, and John A Clark. F for fake: four studies on how we fall for phish. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3469–3478. ACM, 2011. (Cited on page 1.)

[3] Ashley Carman. Phishing scam targets michigan public schools. URL http://www.scmagazine.com/phishing-scam-targets-michigan-public-schools/article/343177/. [Online; accessed 13-May-2014]. (Cited on page 15.)

[4] Madhusudhanan Chandrasekaran, Krishnan Narayanan, and Shambhu Upadhyaya. Phishing email detection based on structural properties. In *NYS Cyber Security Conference*, pages 1–7, 2006. (Cited on pages 2, 5, and 26.)

[5] Zesheng Chen and Chuanyi Ji. Optimal worm-scanning method using vulnerable-host distributions. *International Journal of Security and Networks*, 2(1):71–80, 2007. (Cited on page 18.)

[6] Sidharth Chhabra, Anupama Aggarwal, Fabricio Benevenuto, and Ponnurangam Kumaraguru. Phi. sh/$ ocial: the phishing landscape through short urls. In *Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference*, pages 92–101. ACM, 2011. (Cited on page 21.)

[7] Pern Hui Chia and Svein Johan Knapskog. *Re-evaluating the wisdom of crowds in assessing web security*, pages 299–314. Springer, 2012. (Cited on page 20.)

[8] Robert Cialdini. The science of persuasion. *Scientific American Mind*, 2001. ISSN 1555-2284. (Cited on pages 12, 19, 30, 31, 45, 51, 52, and 53.)

[9] Richard Clayton, Drew Dean, Markus Jakobsson, Steven Myers, Stuart Stubblebine, and Michael Szydlo. A chat at the old phishin'hole. *Lecture Notes in Computer Science*, 3570:88, 2005. (Cited on page 7.)

[10] M Patrick Collins, Timothy J Shimeall, Sidney Faber, Jeff Janies, Rhiannon Weaver, Markus De Shon, and Joseph Kadane. Using

uncleanliness to predict future botnet addresses. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 93–104. ACM, 2007. (Cited on page 18.)

[11] Organización Internacional de Normalización. *ISO/IEC 27002: Information Technology, Security Techniques, Code of Practice for Information Security Management*. ISO/IEC, 2005. (Cited on page 27.)

[12] Belgium Police Department. Overzicht per criminele figuur, . URL http://www.polfed-fedpol.be/crim/crim_statistieken/app_crimestat/app_crimestat_dashboard_crimfig_misdrijven_nl.php. [Online; accessed 13-May-2014]. (Cited on page 17.)

[13] Belgium Police Department. Portaal van de belgische politie, . URL http://www.polfed-fedpol.be/. [Online; accessed 13-May-2014]. (Cited on page 17.)

[14] Rachna Dhamija, J Doug Tygar, and Marti Hearst. Why phishing works. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 581–590. ACM, 2006. (Cited on pages 1, 7, and 27.)

[15] Oxford Dictionaries. Phishing. URL http://www.oxforddictionaries.com/definition/english/phishing. (Cited on pages 5 and 7.)

[16] Collins English Dictionary. Phishing. URL http://www.collinsdictionary.com/dictionary/american/phishing. (Cited on page 7.)

[17] Ronald C Dodge Jr, Curtis Carver, and Aaron J Ferguson. Phishing for user security awareness. *Computers & Security*, 26(1):73–80, 2007. ISSN 0167-4048. (Cited on page 27.)

[18] Douglas P Dotterweich and Kimberly S Collins. The practicality of super bowl advertising for new products and companies. *Journal of Promotion Management*, 11(4):19–31, 2006. (Cited on page 43.)

[19] Christine E Drake, Jonathan J Oliver, and Eugene J Koontz. Anatomy of a phishing email. In *CEAS*, 2004. (Cited on page 26.)

[20] Shaun Egan and Barry Irwin. An evaluation of lightweight classification methods for identifying malicious urls. In *Information Security South Africa (ISSA), 2011*, pages 1–6. IEEE. ISBN 1457714817. (Cited on pages 21, 23, and 24.)

[21] Aaron Emigh. Online identity theft: Phishing technology, chokepoints and countermeasures. *ITTC Report on Online Identity Theft Technology and Countermeasures*, 3, 2005. (Cited on pages xi, 10, 14, 27, and 28.)

[22] Philip V Fellman and Robert Rodriguez. The dark side of the internet. In *International Federation for Information Processing, International Meeting, "IT Innovation for Adaptability and Competitiveness"*. (Cited on page 6.)

[23] RSA FraudAction. Rsa monthly fraud report. URL http://www.emc.com/security/rsa-identity-protection-and-verification/rsa-fraudaction/rsa-fraudaction-antiphishing-service.htm. [Online; accessed 6-August-2014]. (Cited on pages xi and 9.)

[24] Edwin Donald Frauenstein and Rossouw von Solms. *An Enterprise Anti-phishing Framework*, pages 196–203. Springer, 2013. (Cited on pages xi, 10, 11, 12, 27, and 28.)

[25] Adam Greenberg. Medical staffers fall for phishing emails, data on 8,300 compromised. URL http://www.scmagazine.com/medical-staffers-fall-for-phishing-emails-data-on-8300-compromised/article/340590/. [Online; accessed 13-May-2014]. (Cited on page 15.)

[26] Gaurav Gupta and Josef Pieprzyk. Socio-technological phishing prevention. *Information Security Technical Report*, 16(2):67–73, 2011. ISSN 1363-4127. (Cited on page 21.)

[27] Cormac Herley and Dinei Florêncio. A profitless endeavor: phishing as tragedy of the commons. In *Proceedings of the 2008 workshop on New security paradigms*, pages 59–70. ACM, 2009. (Cited on pages 8 and 9.)

[28] Jason Hong. The state of phishing attacks. *Communications of the ACM*, 55(1):74–81, 2012. ISSN 0001-0782. (Cited on pages 8, 9, 10, and 11.)

[29] Huajun Huang, Liang Qian, and Yaojun Wang. A svm-based technique to detect phishing urls. *Information Technology Journal*, 11(7), 2012. ISSN 1812-5638. (Cited on page 20.)

[30] Tom N Jagatic, Nathaniel A Johnson, Markus Jakobsson, and Filippo Menczer. Social phishing. *Communications of the ACM*, 50(10):94–100, 2007. (Cited on pages 1 and 5.)

[31] Markus Jakobsson. Modeling and preventing phishing attacks. In *Financial Cryptography*, volume 5. Citeseer. (Cited on page 7.)

[32] Markus Jakobsson and Steven Myers. *Phishing and countermeasures: understanding the increasing problem of electronic identity theft*. John Wiley & Sons, 2006. ISBN 0470086092. (Cited on pages 1, 5, 6, 7, 8, 14, 16, and 17.)

[33] Lance James. *Phishing exposed*. Syngress, 2005. ISBN 0080489532. (Cited on pages 1, 6, 7, and 27.)

[34] K Jansson and R Von Solms. Phishing for phishing awareness. *Behaviour & Information Technology*, 32(6):584–593, 2013. ISSN 0144-929X. (Cited on page 27.)

[35] Engin Kirda and Christopher Kruegel. Protecting users against phishing attacks with antiphish. In *Computer Software and Applications Conference, 2005. COMPSAC 2005. 29th Annual International*, volume 1, pages 517–524. IEEE, 2005. (Cited on page 38.)

[36] Iacovos Kirlappos and Martina Angela Sasse. Security education against phishing: A modest proposal for a major rethink. *IEEE Security and Privacy Magazine*, 10(2):24–32, 2012. (Cited on page 27.)

[37] Ponnurangam Kumaraguru, Steve Sheng, Alessandro Acquisti, Lorrie Faith Cranor, and Jason Hong. Lessons from a real world evaluation of anti-phishing training. In *eCrime Researchers Summit, 2008*, pages 1–12. IEEE. ISBN 1424429692. (Cited on page 27.)

[38] Ponnurangam Kumaraguru, Yong Rhee, Alessandro Acquisti, Lorrie Faith Cranor, Jason Hong, and Elizabeth Nunge. Protecting people from phishing: the design and evaluation of an embedded training email system. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 905–914. ACM, 2007. (Cited on page 2.)

[39] Ponnurangam Kumaraguru, Justin Cranshaw, Alessandro Acquisti, Lorrie Cranor, Jason Hong, Mary Ann Blair, and Theodore Pham. School of phish: a real-world evaluation of anti-phishing training. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, page 3. ACM, 2009. (Cited on pages xi, 28, 29, and 30.)

[40] Willy Lai. Fitting power law distributions to data. (Cited on page 21.)

[41] Elmer Lastdrager. Achieving a consensual definition of phishing based on a systematic review of the literature. 2014. (Cited on pages 7 and 8.)

[42] Anh Le, Athina Markopoulou, and Michalis Faloutsos. Phishdef: Url names say it all. In *INFOCOM, 2011 Proceedings IEEE*, pages 191–195. IEEE. ISBN 1424499194. (Cited on pages xi and 23.)

[43] Avivah Litan. Phishing victims likely will suffer identity theft fraud. *Gartner Research Note (May 14, 2004)*, 2004. (Cited on page 8.)

[44] Gang Liu, Bite Qiu, and Liu Wenyin. Automatic detection of phishing target from phishing webpage. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 4153–4156. IEEE, . ISBN 1424475422. (Cited on page 20.)

[45] Haotian Liu, Xiang Pan, and Zhengyang Qu. Learning based malicious web sites detection using suspicious urls. . (Cited on pages xi, 23, 24, 25, and 26.)

[46] Christian Ludl, Sean McAllister, Engin Kirda, and Christopher Kruegel. On the effectiveness of techniques to detect phishing sites. In *Detection of Intrusions and Malware, and Vulnerability Assessment*, pages 20–39. Springer, 2007. (Cited on page 38.)

[47] Justin Ma, Lawrence K Saul, Stefan Savage, and Geoffrey M Voelker. Identifying suspicious urls: an application of large-scale online learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 681–688. ACM, . ISBN 1605585165. (Cited on pages xi and 25.)

[48] Justin Ma, Lawrence K Saul, Stefan Savage, and Geoffrey M Voelker. Beyond blacklists: learning to detect malicious web sites from suspicious urls. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1245–1254. ACM, . ISBN 1605584959. (Cited on pages xi, 23, 24, and 25.)

[49] Liping Ma, Bahadorrezda Ofoghi, Paul Watters, and Simon Brown. Detecting phishing emails using hybrid features. In *Ubiquitous, Autonomic and Trusted Computing, 2009. UIC-ATC'09. Symposia and Workshops on*, pages 493–497. IEEE, 2009. (Cited on page 27.)

[50] Steve Mansfield-Devine. Interview: Joe ferrara–fighting phishing. *Computer Fraud & Security*, 2013(7):17–20, 2013. (Cited on page 28.)

[51] Max Emanuel Maurer. Phishload. URL `http://www.medien.ifi.lmu.de/team/max.maurer/files/phishload/index.html`. [Online; accessed 3-April-2014]. (Cited on page 37.)

[52] Tom McCall. Gartner survey shows phishing attacks escalated in 2007; more than $3 billion lost to these attacks. *Stephane GALLAND*, 2007. (Cited on pages 8 and 9.)

[53] Tom McCall. Gartner survey shows phishing attacks escalated in 2007; more than $3 billion lost to these attacks. *Stephane GALLAND*, 2007. (Cited on page 8.)

[54] Kevin D Mitnick and William L Simon. *The art of deception: Controlling the human element of security*. Wiley, 2001. ISBN 0471432288. (Cited on pages 2 and 30.)

[55] Tyler Moore and Richard Clayton. An empirical analysis of the current state of phishing attack and defence. In *WEIS*. Citeseer. (Cited on page 17.)

[56] Tyler Moore and Richard Clayton. *Evaluating the wisdom of crowds in assessing phishing websites*, pages 16–30. Springer, 2008. ISBN 3540852298. (Cited on pages xi, 20, 21, and 22.)

[57] Giovane César Moura. *Internet bad neighborhoods*. Giovane Cesar Moreira Moura, 2013. ISBN 9036534607. (Cited on pages 17 and 18.)

[58] Giovane CM Moura and Aiko Pras. Scalable detection and isolation of phishing. In *Scalability of Networks and Services*, pages 195–198. Springer, 2009. (Cited on pages 8 and 9.)

[59] Philip J Nero, Brad Wardman, Heith Copes, and Gary Warner. Phishing: Crime that pays. In *eCrime Researchers Summit (eCrime), 2011*, pages 1–10. IEEE. ISBN 1457713403. (Cited on pages 11 and 14.)

[60] National Plant Diagnostic Network. Types of social engineering. URL http://www.npdn.org/social_engineering_types. [Online; accessed 16-July-2014]. (Cited on page 42.)

[61] OpenDNS. Phishtank: Out of the net, into the tank. URL http://www.phishtank.com/faq.phpk. [Online; accessed 13-May-2014]. (Cited on page 19.)

[62] Parth Parmar and Kalpesh Patel. Comparison of phishing detection techniques. In *International Journal of Engineering Research and Technology*, volume 3. ESRSA Publications. ISBN 2278-0181. (Cited on page 19.)

[63] Bryan Parno, Cynthia Kuo, and Adrian Perrig. *Phoolproof phishing prevention*. Springer, 2006. ISBN 3540462554. (Cited on page 7.)

[64] James W Pennebaker and Deborah Yates Sanders. American graffiti: Effects of authority and reactance arousal. *Personality and Social Psychology Bulletin*, 2(3):264–267, 1976. (Cited on page 42.)

[65] Phishing.org. History of phishing. URL http://www.phishing.org/history-of-phishing/. (Cited on page 6.)

[66] Swapan Purkait. Phishing counter measures and their effectiveness - literature review. *Information Management & Computer Security*, 20(5):382–420, 2012. ISSN 0968-5227. (Cited on page 18.)

[67] Anirudh Ramachandran and Nick Feamster. Understanding the network-level behavior of spammers. In *ACM SIGCOMM Computer Communication Review*, volume 36, pages 291–302. ACM. ISBN 1595933085. (Cited on page 18.)

[68] Teri Robinson. Phishing scam aimed at google docs, drive users. URL http://www.scmagazine.com/phishing-scam-aimed-at-google-docs-drive-users/article/338369/. [Online; accessed 13-May-2014]. (Cited on page 16.)

[69] Stuart E Schechter, Rachna Dhamija, Andy Ozment, and Ian Fischer. The emperor's new security indicators. In *Security and Privacy, 2007. SP'07. IEEE Symposium on*, pages 51–65. IEEE. ISBN 0769528481. (Cited on page 27.)

[70] Wombat security technology. Phishguru: Assess and motivate your employees using simulated phishing attacks. URL http://www.wombatsecurity.com/phishguru. [Online; accessed 23-May-2014]. (Cited on page 28.)

[71] Frank Stajano and Paul Wilson. Understanding scam victims: seven principles for systems security. *Communications of the ACM*, 54(3):70–75, 2011. (Cited on page 31.)

[72] Henri Tajfel and John C Turner. The social identity theory of intergroup behavior. 2004. (Cited on page 43.)

[73] Gregg Tally, Roshan Thomas, and Tom Van Vleck. Anti-phishing: Best practices for institutions and consumers. *McAfee Research, Mar*, 2004. (Cited on pages 7, 10, and 14.)

[74] Inspired Telemarketing. 5 tips for getting past receptionists!, 2013. URL http://inspiredtelemarketing.wordpress.com/2013/09/13/5-tips-for-getting-past-receptionists/. [Online; accessed 16-July-2014]. (Cited on page 42.)

[75] Merriam Webster. Phishing. URL http://www.merriam-webster.com/dictionary/phishing. (Cited on page 7.)

[76] Liu Wenyin, Ning Fang, Xiaojun Quan, Bite Qiu, and Gang Liu. Discovering phishing target based on semantic link network. *Future Generation Computer Systems*, 26(3):381–388, 2010. ISSN 0167-739X. (Cited on page 21.)

[77] Rebecca Wetzel. Tackling phishing. *Business Communications Review*, 35(2):46–49, 2005. (Cited on pages xi, 10, 12, and 13.)

[78] Joshua S White, Jeanna N Matthews, and John L Stacy. A method for the automated detection phishing websites through both site characteristics and image analysis. In *SPIE Defense, Security, and Sensing*, pages 84080B–84080B–11. International Society for Optics and Photonics. (Cited on page 20.)

[79] Michael Workman. Wisecrackers: A theory - grounded investigation of phishing and pretext social engineering threats to information security. *Journal of the American Society for Information Science and Technology*, 59(4):662–674, 2008. ISSN 1532-2890. (Cited on pages 2, 5, 31, 43, 52, and 53.)

[80] Guang Xiang, Jason Hong, Carolyn P Rose, and Lorrie Cranor. Cantina+: a feature-rich machine learning framework for detecting phishing web sites. *ACM Transactions on Information and System Security (TISSEC)*, 14(2):21, 2011. ISSN 1094-9224. (Cited on pages xi, 23, 24, 25, and 26.)

[81] Huiping Yao and Dongwan Shin. Towards preventing qr code based attacks on android phone using security warnings. In *Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security*, pages 341–346. ACM. ISBN 1450317677. (Cited on page 20.)

[82] Yue Zhang, Serge Egelman, Lorrie Cranor, and Jason Hong. Phinding phish: Evaluating anti-phishing tools. ISOC, . (Cited on page 2.)

[83] Yue Zhang, Jason I Hong, and Lorrie F Cranor. Cantina: a content-based approach to detecting phishing web sites. In *Proceedings of the 16th international conference on World Wide Web*, pages 639–648. ACM, . ISBN 1595936548. (Cited on pages 2 and 25.)

DECLARATION

Put your declaration here.

Enschede, August 2014

Nurul Akbar

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography *"The Elements of Typographic Style"*. `classicthesis` is available for both LaTeX and LyX:

http://code.google.com/p/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured at:

http://postcards.miede.de/

Final Version as of September 11, 2014 (`Nurul Akbar` version 1).