ANALYSING PERSUASION PRINCIPLES IN PHISHING EMAILS

NURUL AKBAR

UNIVERSITY OF TWENTE.

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

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

Nurul Akbar: *Analysing persuasion principles in phishing emails,* Master thesis, © August 2014

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

LOCATION:
Enschede

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

— Quran 3:185

ABSTRACT

As the barrier to abuse system vulnerabilities has been raised significantly with time, attacking users' psyche has rapidly become a more efficient and effective alternative. The usage of email as an electronic means 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 from the unsuspecting victims. Although persuasion can be used as a tool for a good reason, it can also be used for a malicious reason by phishers to get a positive response from an intended victim in phishing emails.

To protect users from phishing attacks on 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 analysis of persuasion techniques in phishing emails. Our research is aimed at understanding the characteristics of phishing emails, by considering persuasion techniques in the real world analysis.

We have conducted a quantitative analysis on our dataset that consists of reported phishing emails between August 2013 and December 2013. The findings are mainly observed from three different viewpoints: general structural properties; persuasion principles characteristics; and their relationships. We have found that financial institutions are the most common target with high number of occurrences in our dataset. Three important findings of our research are that: (1) authority is the most popular persuasion technique regardless of the target and the reason used; (2) depending on the target types and the reason types, the next most popular persuasion principles are scarcity, consistency, and likeability; and (3) scarcity principle has a high involvement with administrator target type and account-related concerns.

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, which have made this research possible. They have gone beyond the expected duties as supervisors and helped 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 a phishing emails corpus with Cialdini's principles as the core of my research. He has assisted me in obtaining the data as it is confidential and sensitive.

I would like to thank PhD candidates in the SCS group for letting me pick their brains when I did a brief presentation at the beginning of my research. Thanks to Geert Jan for letting me work in the lab. and to Suse and Bertine for lending me the key when no one else was in the lab. I would also like to express my appreciation and gratitude to Drs. Jan Schut for providing me invaluable advice 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, and Gaurav and Vignesh who have assisted me by giving feedback on my writing. I appreciate all my friends, Aldi, Saud and all the others who supported me either directly and indirectly during my master studies. Without all their support, accomplishing my studies would not have been possible. I would like to thank my family members and relatives who have supported me 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.

– Nurul Akbar (Nolie)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AOL American Online

URL Uniform Resource Locator

1 INTRODUCTION

With the advancement of information technology in the modern generation, the evolution of the digital era has become more mature in the sense of effectiveness and ease for societies. People can sell and buy goods, conduct banking activities and even participate in political activities such as elections 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, society has been utilizing technological means such as emails, websites, online payment systems and social networks to achieve their tasks efficiently, affordably and in a more focus way. However, the advancement in information and communication technology has been a double-edged sword. As the internet increasingly becomes 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, birthdates, relationships or workplaces can 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. There are many possible incentives that drive phishing attacks including illicit corporate espionage, political power, and, most common of all, financial benefits. The attacker generally masquerades as a legitimate institution to trick users into disclosing personal, financial or computer account information [28]. 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 [28][10]. 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 [28][3][10][29][26]. Social engineering involves the techniques used to deceive people in order to get compliance and response by specific actions that will dis-

close their sensitive information, such as replying to a fake email or clicking a link within an email [50]. Moreover, phishers often use persuasion techniques to manipulate potential victims to engage in certain emotions such as excitement and fear as well as interpersonal relationships, such as trust and commitments, to divert users' attention [76]. 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 via email can be technical or non-technical. One of technical approaches to detect phishing email is achieved by distinguishing between phishing and original emails based on their structural properties, using machine learning [5]. One of the non-technical approaches to defending against phishing attacks is to make people aware of the threats. Security awareness concerning phishing attacks might be achieved by embedded training methods that teach people about phishing during their normal use of email [37].

The common characteristics of a phishing email include structural properties such as misleading hyperlinks and misleading header information [81, 80]. To make a phishing email effective, 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 compliance from the recipient in a phishing email, persuasion is used as an underlying technique to get a positive response from an intended victim [76].

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. Phishers often misuse persuasion techniques to get positive responses from the recipients. Unfortunately, not many studies have included persuasion techniques as important aspects in phishing emails. Consequently, not many people are aware of the existence of persuasion in the emails they are getting. Current phishing email countermeasures are greatly rely on the technical aspect alone rather than integrating the technical aspect with psychological aspects such as persuasion. Based on Cialdini's persuasion principles, there are six basic tendencies of human behavior that can generate a positive response; reciprocation, consistency, social proof, likeability, authority and scarcity [8]. As we mention earlier, persuasion techniques can also be exploited by the phishers to get positive responses from potential victims. Based

Keywords	Scopus	Web of science
TITLE-ABS-KEY (phishing persuasion)	[31, 77, 33, 3, 65, 19]	[77, 33, 76]
(TITLE-ABS-KEY (phishing) AND REF (cialdini))	[77, 35, 23, 71, 32]	

Table 1: Query searches in Scopus and Web of Science

on this reasoning, we will conduct a quick scan of the existing studies regarding persuasion techniques in phishing emails.

We performed a combination of query strings such as "phishing" and "persuasion" to search based on title, abstract, and keyword fields in the Scopus database and we also searched based on topic in the Web of Science database. As we adopt the concept of persuasion based on Cialdini's six principles, we also search "phishing" in TAK(Title, Abstract, Keywords) fields and "Cialdini" in the references field. From Table 1, we can see that two papers that occur in both databases [77, 31] and one paper appearing in both queries [77]. However, Sharma's paper is retracted for having non-original materials, so we will not conduct a review of it. We will describe what the remaining papers do and how they are different from our research in the following points:

- As observed by Kavianto, the overall security risk in an organization depends on its individuals' behavioral decisions to respond to security threats such as phishing attacks. Kavianto also found that an individual's behavioral decisions can be separated into two levels: tradeoffs between uncertainties, losses and benefits; and their susceptibility to persuasion techniques [8] and emotion [31]. Kavianto develops a model of an individual's behavioral decisions in aggregate levels [31]. The outcome of the study is the possibility of incorporating individual-level behavioral biases into the analysis of system level risk (i.e. network security risk). Although Kavianto identifies that successful deception can be linked with how persuasion techniques are used by the perpetrator, no actual data of phishing emails is analyzed.

- Blythe et al. used four methods to show how people are often susceptible to phishing attacks [3]. These four methods are: content analysis; online surveys; interview with blind email users; and literary analysis [3]. Content analysis suggests that phish-

ing email is advancing with better spelling and grammar and supported by the present of visual graphics such as logos. Online surveys shows that while their participants are computer literate, they are not always successful in detecting phishing, even more so with the presence of logo. Blythe et al. found that blind email users are more attentive to the context of the email that is presented [3]. Thus, the detection of phishing by blind email users is higher than non-disabled email users. As it became clear that careful reading is the core process of identifying phishing emails, Blythe et al. then consider a phishing email as literature [3]. Literary analysis shows that the phishers who imitate personalized email from banking and security services allow phishing emails to remain successful, as they exploit people's anxieties in terms of the content of the email itself. Although Blythe et al. conducted content analysis and found that literary analysis was "very" persuasive, they did not conduct their content analysis based on persuasion techniques. Instead, they based the content analysis on structural properties such as sender, grammatical error, logos and style.

- Ghorbani et al. argue that some general approaches designed by attackers to obtain sensitive information to exploit a computer system using social engineering [19]. Social engineering includes aggressive persuasion and interpersonal skills to obtain unauthorized access to a system. Moreover, Ghorbani et al. discussed network attack taxonomy, probes, privilege escalation attacks, Denial of Service (DOS) attacks, Distributed Denial of Services (DDoS) attacks, and worm and routing attacks [19]. However, their study only discussed these network attacks in considerable detail without addressing persuasion theory and little explanation in terms of phishing emails.

- The paper of Krombholz et al. provides a taxonomy of well-known social engineering attacks and studies the overview of advanced social engineering attacks on knowledge workers [35]. What they meant by knowledge worker here is the worker that characterized knowledge as their capital. The paper used Cialdini's persuasion principles as the background study of social engineering. The taxonomy was classified based on attack channel (e.g. emails, instant messenger, social network, etc.), the operator (e.g. human, software), different types of social engineering (e.g. phishing, shoulder surfing, dumpster diving, etc.), and specific attack scenarios. Krombholz et al. provide an overview of social engineering attacks by creating a taxonomy to support further development of social engineering attack countermeasures [35].

- Herzberg et al. tested the effectiveness of different defense mechanisms that use forcing and negative training functions [23]. Their methods involved using an online exercise submission system called "Submit" to simulate phishing attacks. It involved a population of ~400 students and two years of observation. Their outcomes claimed that forcing and negative training functions are very effective in both prevention and detection. However, their defense mechanisms do not consider on the persuasion techniques at all, or analyze data from real phishing emails.

- Vishwanath et al. tested the individual differences in phishing vulnerability within the integrated information processing model [71]. The model focuses on four contextual factors: the individual level of involvement; domain specific knowledge; technological efficacy; and email load. The method involves 161 samples. The conclusion of their study is to show the model can be used as an insight into how individuals get phished. However, they do not consider persuasion principles in terms of their contextual factors, or analyze real phishing attacks. This suggests that persuasion techniques do not play an important role in determining the success of phishing attacks.

- Kawakami et al. developed an e-learning system that uses animation for information security education [32]. The paper only mentions the commitment and consistency based on Cialdini's principles to be used to influence people to take their e-learning system as a security education.

- An interesting paper from Wright et al. which analyzes how people are influenced by Cialdini's six persuasion principles in phishing messages [77][8]. The study involved creating phishing messages that represented persuasion principles and testing them on 2,600 participants to see how many would respond [77]. The outcome of the study is that liking, social proof, scarcity and reciprocation do increase the likelihood of recipients responding to phishing emails [77]. Despite the fact that Wright et al. used the same persuasion principles as our study [8], they tried to find the implication of persuasion principles from the users' perspectives. Although the direction of their paper is different, it can be used as a complementary to our study.

- Workman proposed a peripheral route of persuasion model and conducted a behavioral study based on this model [76]. The model relates basic human factors that respond to persuasion techniques based on Cialdini's six principles [8]. These factors are; normative commitment, continuance commitment, affective commitment, trust, fear, and reactance. Workman created six hypotheses to investigate whether the participants who are prone

to these factors exhibit a higher risk to phishing attacks [76]. Based on Workman's measurement, all six hypotheses are accepted. The data was obtained by a questionnaire and objective observation involving 850 participants. In the end, a total of 612 participants responded. The conclusion is that the participants who have the tendency for these factors are more vulnerable to phishing attacks [76]. For instance, one of Workman's hypotheses stated that: "people who are more reactance, will succumb to social engineering more frequently than those who are more resistance." This suggests that Workman tried to measure the implication of persuasion principles from the users' perspectives as well.

- We found only one paper that was similar to our study [33]. Apart from a different dataset, Kim et al. conducted a content analysis of phishing emails based on message argument quality rather than Cialdini's six persuasion principles [33][8]. They did not relate structural phishing properties such as URLs, attachments, usage of HTML, targeted sector and reason used with their persuasion theory. The message argument quality includes: rational appeals, emotional appeals, motivational appeals and time pressure [33]. The reasoning of rational appeals is determined by direct evidence of causality between events. For example, "a few days ago our online banking security team observed invalid logins to customer accounts. Thus, you are required to re-confirm your online access for account verification." This can be mapped as reciprocation based on Cialdini's persuasion principles [8]. An emotional appeal is defined by fear, sadness, guilt, anger, happiness, affection and humor. In our study this can be mapped into authority or likeability. Time pressure is identified by the limited amount of time the recipient has to respond to a phishing email. This also can be mapped into the scarcity principle. One of the factors representing motivational appeals is "the need of belongingness" which also can be portrayed as the need of being part of a group to forms bond with others. Based on our understanding, we can map motivational appeals such as the social proof principle. However, there is no conception of message argument quality that can be mapped to the consistency principle. Table 2 indicates our mapping from message argument quality into Cialdini's persuasion principles [8]. One interesting result from the study is that the number of time pressure emails (42%, n=285) is not as high as they expected. The study concludes that phishers indeed incorporate rational, emotional and motivational appeals in their dataset. However, the conception of persuasion theory adopted by Kim et al. is different from our study. We argue that diversity of per-

suasion theory needs to be incorporated to achieve an objective conclusion.

Message argument quality	Cialdini's persuasion principles
Rational appeals	Reciprocation
Emotional appeals	Authority and Likeability
Motivational appeals	Social proof
Time pressure	Scarcity

Table 2: A map of message argument quality [33] to Cialdini's persuasion principles [8]

Based on the review, we can say that the studies conducted by Workman [76] and Wright et al. [77] are similar. Both of them have tried to find the implication of persuasion principles from the users' perspectives, which can be complementary to our study.

We have conducted individual investigations on all papers found in both databases. The finding shows not only the different measurement instruments and methods but also the foundation of persuasion theory between the current studies and our study. The low number of results from the query searches also indicate that there is little academic research into study persuasion techniques and phishing area. Therefore, a real world analysis of phishing emails characterization based on persuasion techniques is needed to bridge this gap. This characterization can show to what extent persuasion techniques are used in phishing emails. Our research fills the void as a milestone towards countermeasures against phishing attacks with an insight of psychological aspects.

1.2 RESEARCH GOAL

The main goal of this research is to characterize phishing email properties by considering persuasion principles, finding the association between generic properties and persuasion 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 persuasion principles are introduced as variables in our methodology. We 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 analysis of persuasion principles in phishing emails can also be used to generate a new automated method of detecting phishing emails as one of the primary delivery techniques of phishing attacks.

1.3 RESEARCH QUESTIONS

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

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

Several aspects of phishing email characteristics and hypotheses related to the research questions are addressed in detail in detail in Chapter 3.

1.4 STRUCTURES

This research project is structured as follows:

Chapter 2 describes background and literature reviews about phishing in general. The subsections are: 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; general phishing countermeasures; and lastly the human factor in phishing.

In Chapter 3, we present the rationale of our main research questions and hypotheses. It includes what aspects to be considered to answer the characteristics of phishing emails based on persuasion principles in the dataset and the motivation of our hypotheses to support our research questions.

In Chapter 4, we 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 5 we present our discussion and conclusion of the research project, how the research questions are answered along with the recommendations, the limitations, and how these limitations could become the basis of future research.

2

BACKGROUND & LITERATURE REVIEW

In order to meet our research goal, some necessary knowledge on phishing in general is required. This chapter introduces a general understanding of phishing, an exploration of its damage in financial terms, the overview of its modus operandi, a brief explanation of types of 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 for malicious behavior such as online fraud on a massive scale with little cost to the attackers. The attackers can manipulate the Internet users instead of the computer systems (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 gain trust from online users by mimicking trustworthy and legitimate institutions [28]. Phishing has a similar basic principle to '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 responding to a phishing email, either replying to or clicking on a hidden URL within its content. There are diverse definitions of phishing in our literature reviews. Therefore, we would like to discuss about its universal definition in a later section. However, one of the definitions of phishing, according to the Oxford Dictionary, is as follows:

> "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" [11].

Several studies suggest that phishing is a form of online criminal activity by using social engineering techniques [28][76][26][5]. An individual or a group who uses this technique is called *Phisher(s)*. After successfully obtaining sensitive information from the victim, phishers use this information to access victim's financial accounts or to commit credit card fraud. 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 a later section.

The techniques or modus operandi of phishing may vary, but the most common is using fraudulent emails and websites [29]. A fraudulent website is designed in such a way that it may be identical to its legitimate target. However, a phishing website could also be completely different to its target as there is no level of identicalness. In the following subsections, we introduce how phishing was originally came about and how current literatures formally define phishing.

2.1.1 *The History*

The term "phishing" was first published by the American Online (AOL) UseNet Newsgroup on January 2, 1996 and its use started to expand in 2004 [61]. Since 1996, phishing has flourished. Jakobsson et al. [28] mentioned that in the early years of the '90s (according to [61] 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 thought that these accounts were legitimate and thus activated 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 and therefore deactivated the accounts.

While creating false AOL user accounts with fake credit card information was not phishing, AOL's effort to counter these attacks led to the development of phishing. AOL's countermeasure including directly verifying the legitimacy of credit card information and the associated billing identity, forced hackers to pursue alternative ways [28]. Hackers masqueraded as AOL employees, asking other users for credit card information through AOL's instant messenger and email system [61]. Jakobsson et al. suggest that phishing attacks originated from this incident [28]. Since such attack had not been done before, many users fell victim to then. Eventually, AOL enforced warning systems to most of its customers to be vigilant when it comes to sensitive information [61]. In 2004, phishing was recognized as fully industrialized in terms of economy of crime: In the underground market, "off-the-shelf" components for ready-made phishing attacks were available for sale [1]. To the present day, phishing attacks might not only be motivated by financial gain but also political reasons, and they emerged not only with 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 briefly explore the common definition of phishing. Currently, there is no consensus definition, since almost every research paper, academic textbook or journal has its own definition of phishing [28, 29, 68, 9, 59, 27, 10]. Phishing is also constantly evolving, so it might be very challenging to define its universal terminology. There are not so many studies that specifically address the standard of phishing definition. An exception is a piece of research conducted by Lastdrager [40], which addressed a consensual definition of phishing. Before we decide upon 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" [29]

> "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" [68]

> "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" [59]

It is noteworthy that the definitions described by James et al., Tally et al., and Clayton et al. [29, 68, 9] specify that the phishers only use email as a communication channel to trick potential victims. While it might be true because using email is greatly cost effective, 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, such as private messages on online social networks. This definition is also similar to dictionary libraries [11, 12, 72] that mention email as a medium communication between phishers and users.

We believe that the standard definition of phishing should be applicable to most phishing concepts that are presently defined. Consequently, a high level of abstraction is required to build a common definition of phishing. We are convinced that the formal definition of phishing should not focus on the technology that is being used but

rather on the techniques of how the deception is being conducted, the method of an "act" if you will. Therefore, we follow the definition of phishing by Lastdrager [40] which states that:

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

Lastdrager [40] states that to achieve this universal definition, a systematic review of literature up to August 2013 was conducted along with a manual peer review, which resulted in 113 distinct definitions to be analyzed. We thereby agree with Lastdrager [40] that this definition addresses all the essential elements of phishing and we will adopt it as the universally accepted definition throughout our research.

2.2 THE COSTS OF PHISHING ATTACKS

It is a challenging task to find the real costs from phishing attacks in terms of money or direct costs. This is because the financial damage on banks is only known by the banks themselves and most institutions do not share this information with the public. Jakobsson et al. argue that the phishing economy is consistent with black market economy and does not advertise its successes [28]. In this section, a brief explanation of direct and indirect costs of phishing attacks will be illustrated based on a literature review.

According to Jakobsson et al., direct costs are depicted by the value of money or goods that are directly stolen through phishing attacks [28]. Indirect costs are the costs that do not represent the money or goods that are actually stolen, but the costs which have to be paid by the people who handle these attacks [28], i.e. time, money and resources spent to reset people's passwords.

As we mentioned earlier, it is difficult to assess the damage caused by phishing attacks on banks and institutions as they keep this information to themselves. Furthermore, many users are unwilling to acknowledge that they have fallen prey to phishing attacks. This happens because of fear of humiliation, financial loses, or legal liability [28]. Studies estimate the damage of direct losses to victims in the US only [24][53] to range from $61 million [22] to $3 billion per year [49]. The Gartner Group estimated $1.2 billion direct losses from phishing attacks on US banks and credit card companies for the year 2004 [42]. By 2007, losses had escalated to more than $3 billion [49]. The estimation performed by TRUSTe and Ponemon Institute stated that the cost of phishing attacks was up to $500 million losses in the US for the same year [1].

The lack of information such as a detailed documentation survey on how these numbers were found by Gartner group or Ponemon

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

makes estimations more biased than is generally realized. It is interesting to investigate why their estimates are quoted without really analyzing their bias. One thing that comes to mind is that they might have a hidden agenda to make societies think that the cost of phishing is high. Consequently, people would be obliged to implement anti-phishing systems or engage in phishing awareness in their company, which requires money. With this in mind, we would like to emphasize that our findings on the costs of phishing attacks are only estimates without scrutiny from academic researchers, and they might be an exaggeration. Having said that, even if the cost of phishing attacks is zero, we believe that phishing is still a major problem in terms of trust among users and the misuse of email as a means of communication.

We now consider how phishing attacks are carried out, and if there are distinct stages involved in the attacks? In the next section we review phishing's modus operandi in term of phishing stages or phases.

2.3 MODUS OPERANDI

As we mentioned earlier, a phishing attack is a subset of identity theft. One modus operandi is first to create a fake website that spoofs legitimate website such as financial websites. These websites can be either identical or not identical to real websites; the aim is just to get a response from unsuspecting victims. After that, the phishers will try to trick the potential victim into submitting important information such as usernames, passwords and PINs 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 the target institution is a bank.

Phishers employ a variety of techniques to trick potential victims to access their fraudulent website. One of the typical ways is by sending illicit emails on a large scale claiming to be from a 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. While there are various techniques of phishing attack, we address the common phases of phishing that we analyzed in a literature survey of several studies and we also address our own phishing phases. These phases are compiled in Table 3.

Table 3: Compilation of phishing phases

J. Hong [24]
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. [18]
1. Planning 2. Email Design 3. Fabricated story 4. Threatening tone/Consequences 5. Spoofed website
Wetzel [74]
1. Planning 2. Setup 3. Attack 4. Collection 5. Fraud 6. Post-attack
Tally, et al. [68]
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 [17]
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. [54]
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 [24] argued that there are three major phases, while Frauenstein et al. [18] suggested that five main processes are used to perform phishing attacks based on the perspective of the attacker.

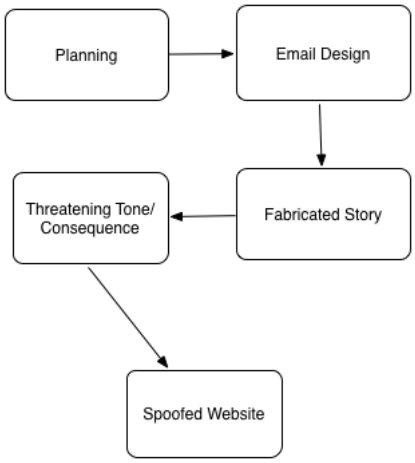


Figure 1: Phishing processes based on Frauenstein[18]

As we illustrated in Figure 1, the first process is called *Planning*. In this process a phisher would usually do some reconnaissance on how would the attack is executed and what information would be obtained from the victim.

In the second process, the phisher would think about the design of the email. This email is desired by the phisher to look as legitimate as possible to potential victim. For this purpose, target institutions' logos, trademarks, symbols, etc. are used to make the content look official to the victim. The author called this process as *Email Design*. Figure 2 illustrates the example design of a fake email that impersonates ING bank and can trick unsuspecting victims[2]. From the figure, we can spot a fake email by investigating the real sender email address by looking at the email header or the URL provided in the body, seeing whether it redirects to the official ING website or not.

In the third process, the phisher *fabricates* a story to make potential victims think that email is important. To achieve user attention, a phisher might build up a story about system upgrade, account hijacking or security enhancement so that the victim would feel obliged to be informed. Evidently, this technique corresponds with Cialdini [8], who suggests there are six principles to persuade people to comply with a request.

In the fourth process, a phisher usually includes a *threatening tone* or explain the urgency and consequences if the potential victim chooses not to take action desired by the phisher (e.g. account removal, account blocked). Consequently, users may fear for their account being deleted.

The last process involves a fraudulent website that has been created by the phisher. Users may falsely believe the message given in the email and may click on a Uniform Resource Locator (URL) that

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

Figure 2: Example of a phishing email impersonating ING bank

is embedded in the email. Subsequently, the URL redirects users to a *spoofed website* which may request 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 about 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 [74] 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 3. We speculate that Wetzel's taxonomy is not analogous with Frauenstein's main phishing processes [18]. The difference is that Frauenstein et al. only focused on the design of the attack while Wetzel has added several phases like *Collection*, *Fraud* and *Post-attack*. Therefore, Wetzel's taxonomy is more holistic in term of phishing.

As we listed Wetzel's taxonomy in Table 3, we explain more of the taxonomy as follows:

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

2. *Setup*: After the target, victim and the method are known, the phisher crafts a platform where the victim's information could be transmitted and stored, for example a fraudulent web- site/e-mail.

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.

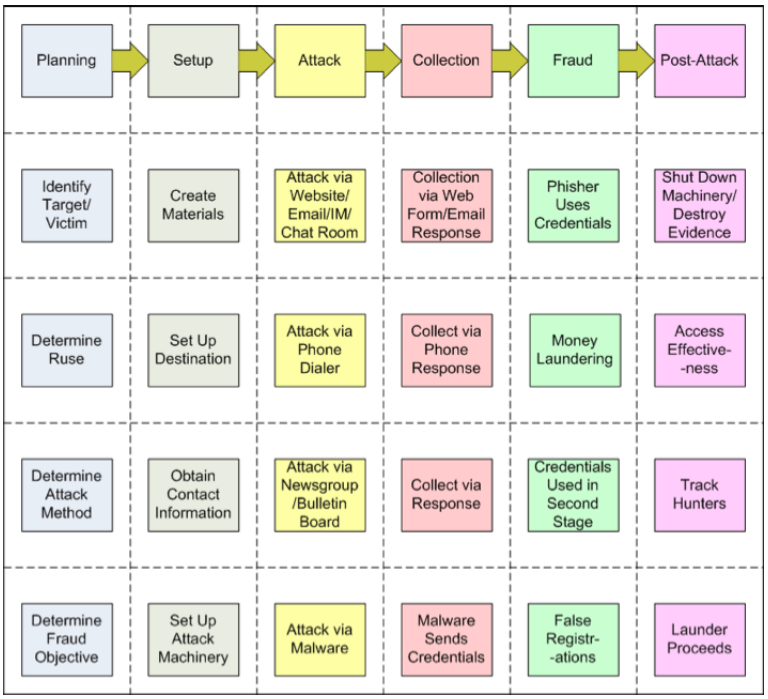


Figure 3: Phishing attack taxonomy and lifecycle[74]

5. *Fraud*: Phisher abuses victim's information by impersonating 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, they would not want to be noticed or detected by authority. Thus, the phisher might need to destroy evidence of the activities that he/she had undertaken.

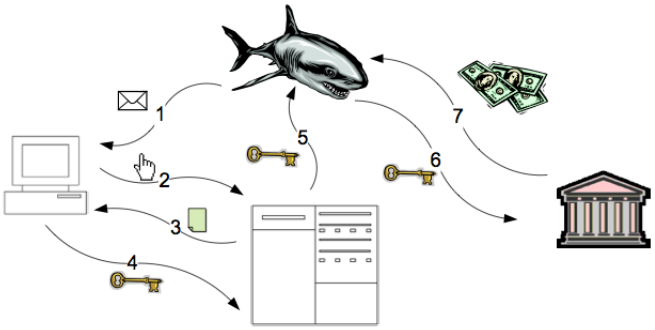


Figure 4: Flow of information in phishing attack [17]

As shown in Table 3, Tally et al. suggest that there are several phases involved in a phishing attack based on the attacker's point of view [68]. The first phase represents the planning, where 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, can be viewed as design phase. We consider the third phase as the delivery and attack phases as it involves the attacker sending a fake email to the unintended victims and hiding the true source. The fourth phase represents attack phase as it involves with the recipient complying with the attacker's request(s). Lastly, the fifth phase represents the fraud phase, as it related to the attacker harvesting and exploiting the victim's resources. Additionally, the phases described by Tally et al. [68] are comparable with the information flow explained by Emigh[17] that illustrated in Figure 4 and explained in Table 3. Phishing attack steps executed by the phisher are also addressed by Nero et al [54]. In their study, a successful phishing attack involves several phases, which can be seen and compared in Table 3.

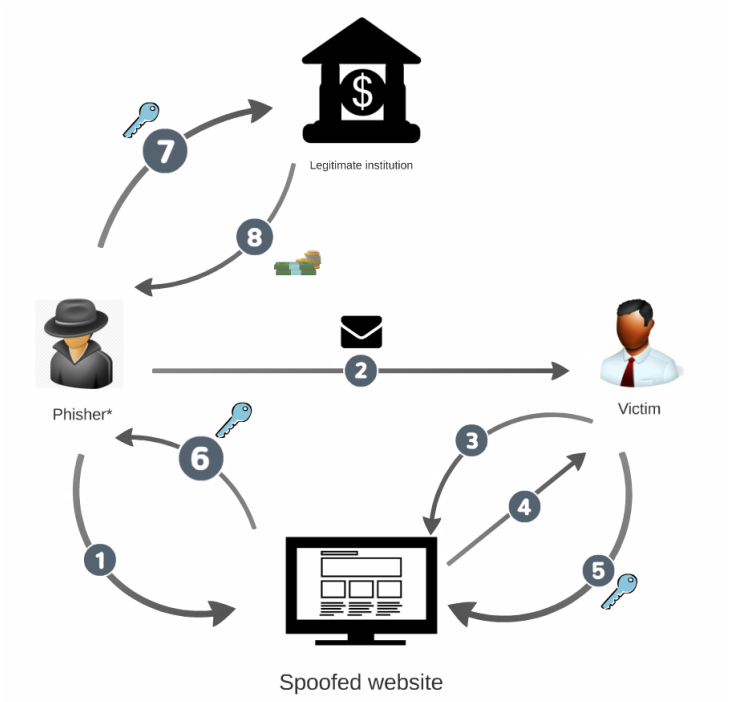


Figure 5: Information flow phishing attack

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. [28]: *the lure*, *the hook* and *the catch*. Figure 5 synthesizes these three components with our phases based on the attacker's 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
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 5 we have only illustrated the phases if the bait was using a spoofed website as a method.

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

2.4 TYPES OF PHISHING

In January 2014, the data of 8300 patients was compromised by US medical companies [20]. The data includes names, addresses, dates of birth and phone numbers. Other than demographic information, clinical information associated with this data was also stolen, including social security numbers. In April 2014, phishers successfully stole US$163,000 from a US public school based in Michigan [4]. It is said that the email prompting the transfer of money came from the finance director of the school. In March 2014, Symantec discovered a phishing attack aimed at Google Drive users [63]. The attack involved firstly an incoming email asking for the opening of document hosted at Google Docs. Users who clicked on the link were taken to a fraudulent Google login page that prompted Google users credentials. Interestingly, the URL seemed very convincing because it was hosted on Google's secure servers. We hypothesize that even more phishing incidents take place in the financial sector, but sometimes the news is kept hidden to maintain creditability. With this in mind, we believe fake websites might be hosted in the network that has more phishing domains than other networks. In the next section, we will discuss the general phishing countermeasures.

One may ask: what types of phishing are there? What are the general types of phishing relevant to our research? Based on the cost of phishing attacks in Section 2.2, the threat of such attacks is alarming and might evolve in the future with more sophisticated technique of attacks. For this reason, it might be useful to provide a brief insight into popular variants of phishing that currently exist. We will briefly explain the types of phishing that are the most relevant to our

research, based on the work of Jakobsson et al. [28]. These types of phishing are strongly related to the phishing definition that we use, 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. The typical scenario of phishing based on visual similarities is to send a large amount of illicit emails containing a call to action asking recipients to click embedded links [28]. These variations include cousin domain attacks. For example, there is a legitimate PayPal website addressed as wow.paypal.com. Cousin domain attacks confuse potential victims to believe that wow.paypal-security.com is a subdivision of the legitimate website due to similar-looking address. Similarly, homograph attacks create a confusion using similar characters to its addresses. For example, with wow.paypal.com and wow.paypa1.com, both addresses look the same but on the second link there is a "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 hostnames to redirect potential victim to phishing websites and JavaScript is used to take over the address bar of a browser to make potential victims believe that they are communicating with the legitimate institution.

Another type of deceptive phishing scheme is rock-phish attacks. They were held responsible for half of the reported phishing incidents worldwide in 2005 [51]. These attacks evade email filters by utilizing random text and GIF images that contain the actual message. Rock phish attacks also utilize a toolkit that is able to manage several fraudulent websites in a single domain. Sometimes, deceptive phishing schemes lead to installation of malware when users visit fraudulent website. We describe malware-based phishing schemes in the next section.

2.4.2 *Malware-based phishing*

Generally, malware-based phishing refers to any type of phishing that involves installing malicious software onto users' personal computers [28]. 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.

In the next section, we study general phishing countermeasures in term of phishing detection and prevention.

2.5 CURRENT COUNTERMEASURES

There are various types of phishing countermeasures that are implemented at different levels. Purkait conducted an extensive research reviewing countermeasures available up until 2012 and analyzing their effectiveness [62]. He suggests that there is a classification of phishing countermeasures into separate groups, as follows:

- Stop phishing at the email level
- Security and password management toolbars
- Restriction list
- Visually differentiate the phishing site
- Two factor and multi channel authentication
- Takedown, transaction anomaly detection, log files
- Anti-phishing training
- Legal solutions

In addition, Parmar et al. suggest that phishing detection can be classified into two types: user training approach and software classification approach [58]. They illustrate a diagram and a table that summarizes phishing detection as countermeasures in a broad view[58]. They also argue the advantages and disadvantages of each category [58]. However, as our research mainly focuses on an analysis of phishing emails based on Cialdini's six principles of persuasion [8], we briefly discuss the most relevant phishing countermeasures: restriction list group (e.g. Phishtank); machine learning approach (web-based phishing); properties or features in a phishing email; and anti-phishing training group (e.g. PhishGuru). In the last section of this chapter, we explore the human factor in phishing attacks, to see how phishing emails are engineered to gain the recipient's trust in order to get a response.

2.5.1 *Phishing detection*

In this subsection, we conduct a literature review related to Phishtank as restriction list, and machine learning approach to detect spoofed websites as phishing detection.

2.5.1.1 *Phishtank*

One of the most common approaches to detect phishing attacks is the implementation of a restriction list. As the name suggest, it prevents users from visiting fraudulent websites. One of the efforts to achieve

a 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 [57]. Phishtank stores phishing URLs in its database and is widely available for use by other companies for creating restriction lists. Some of the big companies that use Phishtank's data are: Yahoo Mail, McAfee, APWG, Web Of Trust, Kaspersky, Opera and Avira.

In this section, we discuss what current literatures deal with phishing data provided by Phishtank. The first step to getting a list of relevant literatures is by a keyword search in the Scopus online library. Putting "Phishtank" as a keyword search results in 12 pieces of literature. The next step is to read all the abstracts and conclusions of the resulting keyword search. We decided that 11 pieces of literatures were relevant to our research. Table 4 summarizes the papers selected and their relevance to Phishtank:

Table 4: Summary phishtank studies

Paper title	**First author**	**Relevance to Phishtank**
Evaluating the wisdom of crowds in assessing phishing website [52]	Tyler Moore	Examines 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	Examines 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 [43]	Gang Liu	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 [75]	Joshua S. White	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 [2]	P.A. Barraclough	Phishtank features are used as one of the inputs of neuro fuzzy technique to detect phishing websites. The study suggested 72 features from Phishtank by exploring journal papers and 200 phishing websites.
Towards preventing QR code based attacks on android phone using security warning [79]	Huiping Yao	Phishtank API is used to lookup whether the given QR containing phishing URL was in the Phishtank database.
A SVM based technique to detect phishing URLs [25]	Huajun Huang	Phishtank database is used as a validation resulting 99% accuracy by SVM method. Furthermore, top ten brand names in the Phishtank archive are used as features in SVM method.
Socio technological phishing prevention [21]	Gaurav Gupta	Analyzes the Phishtank verifiers (individual/organization) to be used as an anti-phishing model.
An evaluation of lightweight classification methods for identifying malicious URLs [16]	Shaun Egan	Indicates that lightweight classification methods achieve an accuracy of 93% to 96% with trained data from Phishtank.
Phi.sh/$oCiaL: The phishing landscape through short URLs [6]	Sidharth Chhabra	Phishtank database is used to analyze suspected phishing that is done through short URLs.
Discovering phishing target based on semantic link network [73]	Liu Wenyin	Phishtank database is used as a test dataset to verify their proposed method (Semantic Link Network)

From our literature survey, we know that Phishtank is a crowd-sourced platform to manage phishing URLs. For that reason Moore et al. aim to evaluate the wisdom of crowd platforms accommodated by Phishtank [52]. Moore et al. suggest that the user participation is distributed according to power law. It uses model data where the frequency of an event varies as a power of some attribute of that event [39]. 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 all wealth is controlled by 20% of the population. It makes sense that in Phishtank's verification system, a single

3 http://kottke.org/03/02/weblogs-and-power-laws

highly active user's action can greatly impact the system's overall accuracy. Table 5 summarizes the comparison performed by Moore et al. [52] between Phishtank and closed proprietary anti-phishing feeds.[4] Moreover, there are some ways to disrupt the Phishtank verification system, such as: submitting invalid reports accusing legitimate websites; voting legitimate websites 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 are a more subtle method, intended to undermine Phishtank's credibility.

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 5: Comparison summary [52]

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 the overall system [52]. Also, there is a high probability that bad users could extensively participate in submitting or verifying URLs in Phishtank.

2.5.1.2 *Machine learning approach in detecting spoofed website*

The fundamental task of a phishing detection system is to distinguish between phishing websites and legitimate ones. As we previously discussed, the aim of phishing attacks is to gather confidential information from potential victims. To do this, phishers often prompt for this

4 The author conceals the identity of the closed proprietary company

information through fraudulent websites and masquerade as legitimate institutions. It does not make sense if phishers created them in a way very distinctive from its target. It may raise suspicions with the result of unsuccessful attack. To put it another way, we speculate that most phishing websites are mostly identical to legitimate websites to reduce suspicion of potential victims.

In contrast to one of the 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 websites and 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, which are called features or characteristics.

Furthermore, there are several learning algorithms to classify the data, such as logistic regression, random forest, neural networks and support vector machines. As this particular topic is outside the scope of our research, we do not discuss the learning algorithm that is currently implemented. We limit ourselves to introducing three features that are used in machine learning based detection.

There are vast amount of features that can be used in machine learning to detect phishing attacks. Literature was selected by keyword search such as "phishing + detection + machine learning". We analyzed three features: lexical feature, host-based feature and site popularity feature. Each of these features are introduced briefly below:

- 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 [45]. These features include how many dots exist, the length, how deep the path traversal the URL has, or if there any sensitive words present in a URL. For example, with the URLs https://wow.paypal.com and http://wow.paypal.com.example.com/ or http://login.example.com/wow.paypal.com/, we can see that the domain paypal.com is positioned differently, with the first one being the benign URL. Le et al. suggest we can extract the features related to the full URL, domain name, directory, file name and argument [41]. For example, if we want to extract features related to the full URL, we can define the length of the URL, the number of dots in the URL, and whether the blacklisted word presents in the URL. The blacklisted words consist of sensitive words such as confirm, account, login or webscr.

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

Haotian Liu, et al. [44]	**Guang Xiang, et al. [78]**
- 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 6: Existing lexical features [44, 78]

utilizing lexical features combined with external features such as WHOIS data [41]. Egan et al. conducted an evaluation of lightweight classification that includes lexical features and host-based features in its model [16]. The study found that the classification based on these features resulted in extremely high accuracy and low overheads. Table 6 lists the existing lexical features that are currently implemented by two different studies [78, 44]. However, Xiang et al.[78] pointed out that URL structure can be manipulated with little cost, causing the features to fail. For example, attackers could simply remove the 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 [78], for example, http://wow.example.net/pathto/wow.paypal.com. Suspicious URL features 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 [78], 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 to use for first phase analysis of large amounts of data [16].

- Host based features

Since phishers often host phishing websites with less reputable hosting services and registrars, host-based features are needed to observe 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 [45]. Table 7 shows the host-based features from three studies

Justin Ma, et al.[45, 46]	**Haotian Liu, et al. [46][44]**	**Guang Xiang, et al. [78]**
- WHOIS data - IP address information - Connection speed - Domain name properties	- Autonomous system number - IP country - Number of registration information - Number of resolved IPs - Domain contains valid PTR record - Redirect to new site - All IPs are consistent	- Age of Domain

Table 7: Host-based features [45, 46, 44, 78]

that are currently used in machine learning phishing detection. These studies are selected only for example comparison.

Each of these features matters for phishing detection. However, as our main objective is an analysis of phishing emails based on Cialdini's persuasion principles, we do not describe each of these features in detail. It is noteworthy that some of the features are subset of other features. For instance, autonomous system number (ASN), IP country and number of registration information are derived from WHOIS information. Nevertheless, we only explain the few of them that we consider the most crucial:

1. WHOIS information: Since phishing websites and hacked domains are often created at a 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. 2. IP address information: Justin Ma et al. used this information to identify whether or not an IP address is blacklisted [46, 45]. 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

Guang Xiang, et al. [78]	Haotian Liu, et al. [44]
- 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 8: Site popularity features [78, 44]

popularity than a legitimate website. According to [78], some of the features indicated in Table 8 are well performed when incorporated with machine learning system.

1. Page in top search results: this feature was originally used by [80] to find whether or not a website shows up in 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 [78]. We believe this feature is similar to the Number of external links feature since both of them are implying the same technique.

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 [78], the intuitive rationale of this feature is that phishing websites often have very low PageRank due to their ephemeral nature and very low incoming links that are redirected to them. This feature is similar to Real traffic rank feature employed by [44] 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 [78] and the latter feature can be obtained from amazon.com [44].

2.5.1.3 *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. [5, 15] specify the structure of phishing emails properties as follows:

1. Spoofing of online banks and retailers. Impersonation of legitimate institutions may occur at the email level. Phishers may design a fake email to resemble the reputable company in order 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 hidden URL and this method is 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 is often not personalized, unlike a legitimate email that address its recipient with personalized information such as the last four digits of their account.

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.

Ma et al. experimented with seven properties to consider in phishing emails: 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 [47]. Based on this survey, we established phishing email properties as variables in order to classify our data in Section 4.1.3.1.

2.5.2 *Phishing prevention*

Phishing attacks aim to bypass technological countermeasures by manipulating users' trust and can lead to monetary losses. Therefore, human factors play a big part in the phishing taxonomy, especially in the organizational environment. Human factor in phishing taxonomy is comprised of education, training and awareness [18]. Figure 6 illustrates where human factors play a part in phishing threats [18]. User's awareness of phishing has been explored by several studies [29, 18, 17, 36, 30, 13] as preventive measures against phishing attacks. According to ISO/IEC 27002 [18][56], information security awareness is important and it has been a critical success factor in mitigating security vulnerabilities that attack user's trust. One approach to hopefully prevent phishing attacks is to implement an anti-phishing warning/indicator. Dhamija et al. suggest that users often ignore security

indicators, thus making them ineffective [10]. Even if users notice the security indicators, they often do not understand what they represent.

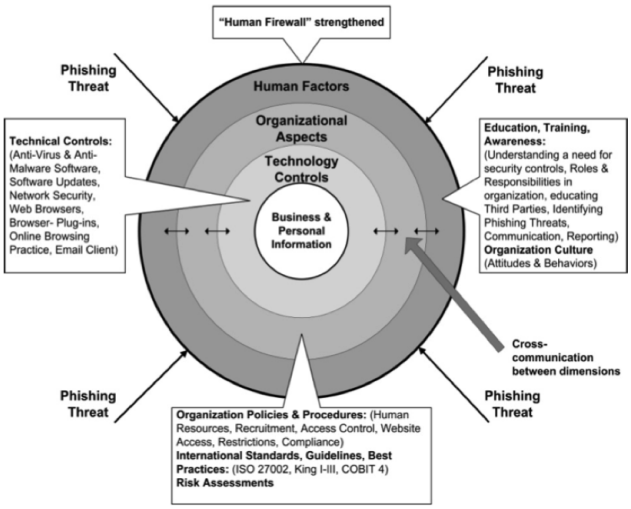
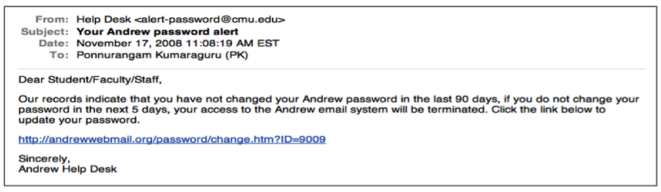


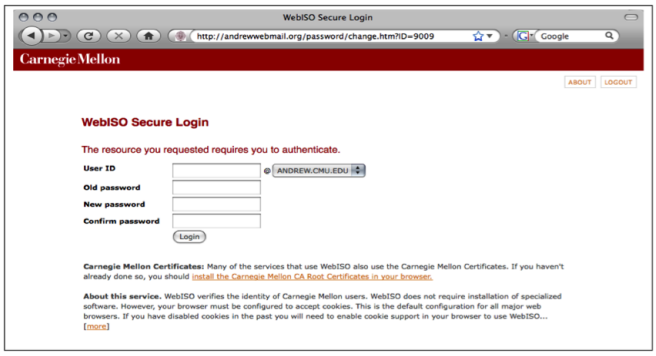
Figure 6: Holistic anti-phishing framework [18]

Moreover, the inconsistency of positioning on different browsers makes it difficult to identify phishing [34]. Schechter et al. pointed out that 53% of their study participants were still attempting to provide confidential information even after their task was interrupted by strong security warning [64]. This suggests that an effective phishing education must be added as a complementary strategy to complete technical anti-phishing measures.

Phishing education for online users often includes instructing people not to click links in an email, ensure that the SSL is present and to verify that the domain name is correct before giving information. This traditional practice evidently has not always been effective [17]. One may ask, therefore, what makes phishing education effective. A study suggests that in order for online users to be aware of phishing threats, they need to be really engaged so that they understand how vulnerable they are [48]. To do this, simulated phishing attacks are often performed internally in an organization. Figure 7 shows a simulated phishing email and website carried out by Kumaraguru et al. from PhishGuru [38]. This scenario delivers the ultimate teachable moment if they fall for these attacks.



(a) Simulated phishing email [38]



(b) Simulated phishing website [38]

(c) Simulated phishing message [38]

Figure 7: Simulated phishing attack [38]

Phishguru is a security training system operated by Wombat security technology that teaches users not to be deceived by phishing attempts through simulation of phishing attacks [69]. They claimed Phishguru provides more effective training than traditional training as it is designed to be more engaging. Figure 8 illustrates how embedded phishing training was presented by PhishGuru.

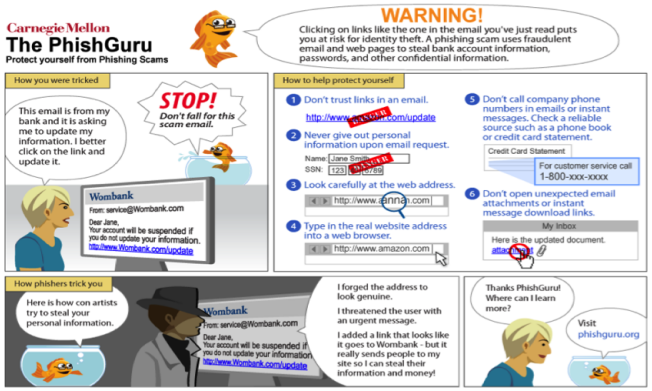


Figure 8: Embedded phishing training [38]

Kumaraguru et al. investigated the effectiveness of embedded training methodology in a real world situation [38]. They indicated that 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 websites 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 is less likely a trained user gave a false positive when he or she was requested to give information by true legitimate emails [38]. This suggests that a user training strategy or effective phishing education is necessary to improve phishing awareness, especially in organizational settings.

2.6 HUMAN FACTOR AND PERSUASION

Phishing attacks generally aim to manipulate end users to comply with the 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? Why are online users gullible to these attacks?

Kevin Mitnick, who obtained millions of dollars by performing social engineering techniques, is probably the best known person who had used the social engineering technique to carry out attacks. His book entitled "The art of deception: Controlling the Human Element of Security" [50] 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 Mitnick's definition we can learn that people are the main target of the attack. He 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 are: *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 a smaller request [8]. Consistency is a public commitment where people commit to the decision they have made [76][8]. Social proof is when people follow the behavior of their peer group, role models or important others because it is generally "fashionable" [76]. Stajano et al. suggest people will let their guard down when

everybody around them appears to share the same risk [66]. Likeability is when people trust those they find attractive or credible [76, 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; losing something valuable; and fear of punishment, humiliation or condemnation [8, 76]. Stajano et al. suggest that scarcity is related to the time principle: that is, when we are under time pressure to make an important choice, we tend to have less reasoning in making a decision [66].

The human being as the "weakest link" in computer security has been exploited for a long time. Security designers blame users and complain that: "the system I designed would be secure, if only users were less gullible" [66]. Stajano et al. stated that: "a wise security designer would seek a robust solution which acknowledges the existence of these vulnerabilities as unavoidable consequence of human nature and actively build countermeasures that prevent this exploitation" [66]. With this in mind, the exploration of persuasion principles is congruent with our research goal. Cialdini's six persuasion principles are the foundation of our research.

3

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 with analysis of data collected by 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 phishing emails?

The characteristics of phishing emails in our dataset are determined by the following parameters:

- How often phishing emails 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 subsubsection 4.1.3.1.

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

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

We established 16 hypotheses to indicate the relationship between generic properties and relevance of persuasion principle to these properties. H8, H9, H10, H13, H14, H15 will answer RQ1 in respect to the relationship between generic properties and the rest will answer RQ2. We conducted an analysis of phishing emails based on Cialdini's principles. In order to conduct the analysis, we established our decision making to classify which persuasive elements that exist in a phishing email. This process will be explained in subsubsection 4.1.3.2.

In our coding of Cialdini's principles and phishing email dataset, we identified phishing emails with fake logos and signatures that may mistakenly be regarded as legitimate by average internet users. For example, in the context of phishing email, signatures such as "Copyright 2013 PayPal, Inc. All rights reserved" or "Administrator Team" and the Amazon logo were used to show the "aura of legitimacy". In the real world, telemarketers and sellers use authoritative element to increase the chance of potential consumers' compliance [70]. This means that they have to provide information in a confident way. Consumers will have their doubts if sellers are unsure and nervous when they offer their products and services. This principle has been one of the strategies in a social engineering attack to acquire action and response from a target [55].

It makes sense if a government has the authority to compose laws and regulations and to control its citizens. Government sectors including court and police departments are also authorized to execute penalties if any wrongdoing happens within their jurisdiction. However, a government may not have to be likeable to enforce their rules and regulation. An administrator who controls his/her network environment may behave in a similar fashion to 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 an administrator will likely have authority principle

---

Similar to the authority principle that may trigger compliance, scarce items and shortage may produce immediate compliance from people. In essence, people will react when their freedom is restricted about a valuable matter when they think they are capable of making a choice among different options [60]. For example, in the phishing email context, we may get an email from Royal Bank informing us that we have not been logged into our online banking account for quite some time, and as a security measure they must suspend our online account. If we would like to continue to use the online banking facility, we must click the URL provided. The potential victim may perceive their online banking account as their valuable matter to access facility and information about their savings. Consequently, they may react to the request because their account could be scarce and restricted. In a real world example, a hard-working bank customer who perceives money is a scarce item may immediately react when their bank informs them that they are in danger of losing their savings due to a "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, as they are relying on emails, websites and phone calling [14]. To exhibit charm or charisma to the potential victims, perpetrators may gain their trust by establishing friendly emails, affectionate websites and soothing voices 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, the e-commerce/retail sector may apply likeability principles to gain potential customers. We therefore hypothesize that:

H4: Phishing emails that target e-commerce/retail companies will likely have a significant relationship with likeability principle

---

Tajfel et al. argue that people often form their own perception based on their relationship with others in certain social circles [67]. This leads to an 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 [76]. 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 visited the website as well. Thus, we hypothesize that:

H5: Phishing emails that target social networks will likely have significant 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 deem valuable. For example, if a perpetrator masquerading as an authority and dressed as police officer stops us on the road, the perpetrator may tell us that we did something wrong and that they will take our driving license if we do not pay them the fine. In the phishing email context, an email masquerading as "System Administrator" may tell us that we have exceeded our mailbox quota, so the administrator must freeze our email account, and that we can reactivate it by clicking the URL provided in the email. This scenario uses both the authority principle and scarcity principle. Therefore, we hypothesize that:

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

We often stumble upon a group of people requesting us to donate some of our money to more unfortunate people. Of course, they use physical attractiveness and kind words to get our commitment to support those people. Once they have got our commitment, they start asking for a donation, and we tend to grant their request and give some of our money to show that we are committed. Phishing emails can work in a similar way. For example, an email may say that Paypal appreciates our membership and kindly notifies us that in the membership term of the agreement they must perform an annual membership confirmation of its customers. Based on this scenario, we know that the email has the likeability principle and 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 their fake product as genuine as possible and hide the fabricated element of their product. There are also fraudsters that do not make their product identical to the legitimate product. In the phishing email context, we perceive fake products as URLs in an email. Phishers do not necessarily hide 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:

H8: Phishing emails that include URLs will likely different than the actual destination

We know from experience that 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 would not make sense if they try to sell their product but 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 for people's attention to look at our product. In the phishing email context, phishers may include a URL or attachment in the body of the email and they may also request the unsuspecting victim to click the URL or to open the attachment. This leads us to the following two hypotheses:

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

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

We sometimes find it suspicious if a person dressed as a police officer does not have a badge. Consequently, a fake police officer may use a fake badge to build up even more "aura of legitimacy". Cialdini suggests the increment of passers-by who stop and stare at the sky by 350 percent with a person in 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 an image in the phishing mail context. Similarly, an email that masquerades as Apple may clone the Apple company logo or trademark to its content to increase the chance of a potential victim's response – to increase the "believability". Thus, we hypothesize that:

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

Apart from the target analysis, we also investigate the reason why potential victims respond to phishers' requests. Phishing emails that imply account expiration incorporate the scarcity principle because the account itself may be very valuable for us and we fear it expiring or being terminated. Therefore, we hypothesize that:

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

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

H13: Phishing emails that have account-related reasons will likely include URL(s)

When a phishing email has document-related reasons such as reviewing some document reports or court notice, they tend to impersonate a government to make the email realistic enough to persuade the potential victim more than other targets. We therefore hypothesize that:

H14: Phishing emails which targeting government sector will likely have document-related reasons

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

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

We think it makes sense if a phishing email that uses HTML to present their email design to be more attractive to the potential victim. Consequently, an unsuspecting victim may respond to the request just because the email design is attractive. Therefore, we hypothesize that:

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

4 DATA AND ANALYSIS

This chapter explains our research methodology and results in details. We begin by explaining the framework of our methods, which consists of several steps in order to get our results. By the end of this chapter, we present the results of our analysis to answer the research questions that we explained in chapter 3.

4.1 RESEARCH METHODOLOGY

As we illustrate in Figure 9, we processed our data in several steps. We collected the data from a security organization in the form of suspected phishing email reports, and then performed a data selection that consists of selecting 8444 raw emails into 207 unique phishing emails. The selection process can be found in subsection 4.1.2. In the next step, we executed data classification into the variables that we established in subsubsection 4.1.3.1 so that we could reconstruct into an SPSS readable dataset. Finally, we conducted statistical analyses to answer our hypotheses.

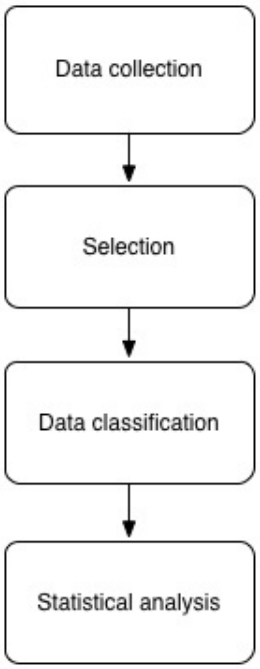


Figure 9: Research methodology diagram

4.1.1 *Data collection*

The data was obtained from an organization based in the 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. The data consists of 8444 suspected phishing emails in total that are selected and classified in the following sections.

4.1.2 *Selection*

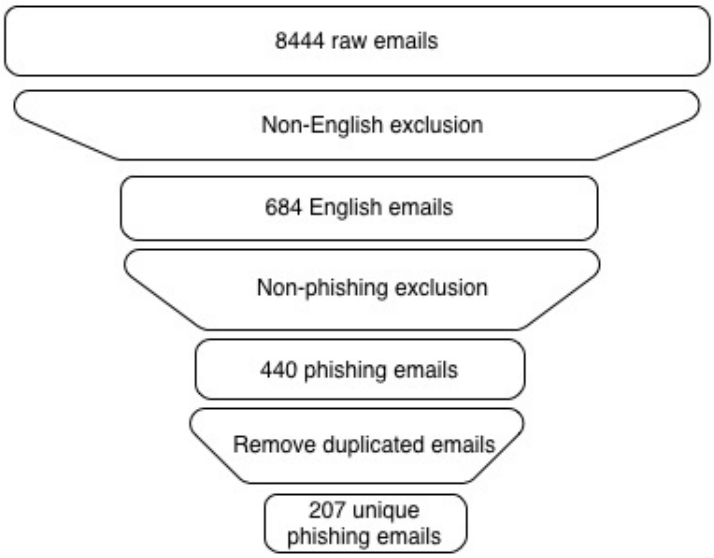


Figure 10: Selection diagram

The selection process consists of non-English exclusion, non-phishing exclusion and removing duplicated emails. Figure 10 illustrates the selection process.

4.1.2.1 *Non-English exclusion*

By manually inspecting each of suspected phishing emails, we can separate the emails based on languages. These languages consist of English, Dutch and other languages. The raw data was sorted by the subject to help ease the separation process. This process gave the following results:

- 7756 suspected phishing in Dutch language
- 684 suspected phishing in English language
- 4 suspected phishing in other languages

We excluded the suspected phishing emails in non-English languages because our proficiency of non-English languages is not sufficient. More detail on why we excluded the emails with non-English languages can be found in section 5.3 and section 5.4.

4.1.2.2 *Non phishing exclusion*

From 684 suspected phishing emails in the English group, we excluded the non-phishing emails by categorizing them into Phishing, Legitimate and Others groups. The phishing group consists of the emails that were indeed phishing. The legitimate group consists of legitimate emails. The others group contains spam emails that represent commercial advertisements and the emails that have no content – for instance, when the content has been removed before it was forwarded.

This process gave the following results:

- 440 Phishing
- 18 Legitimate
- 226 Others

Interestingly, based on the results of the categorization process, we found 18 legitimate emails that were mistakenly reported as phishes (i.e. false positives). This suggests that although there are only 18 false positives, misinterpretation of a fraudulent email among the reporters is still occurred.

4.1.2.3 *Removing duplicated emails*

We coded the 440 phishing emails to an Excel sheet with necessary variables so that we could convert the Excel sheet into an SPSS readable file. Our aim was to analyze only the unique phishing emails, so that the dataset would not be redundant. Duplicated emails in our dataset were defined as those having exactly the same text in the entire body of the emails. To find duplicated phishing emails, we conducted the following steps:

1. Sorting the 440 emails by the subject, to show which emails had the same subject.
2. Manually investigating each email with other emails with the same subject to make sure all the text in the entire body is exactly the same. If it did we excluded it.
3. Sometimes duplicated emails have slightly different subjects. To find more duplicated emails, we searched based on random phrase from the body.

4. If other emails found, we manually investigate each email to make sure they had exactly the same text in the entire body.
5. We indicate the number of duplicated emails in "CounterSameContents" variable that we explain in subsubsection 4.1.3.1.

These steps gave 207 unique phishing emails.

4.1.3 *Data Classification*

We classified our data into our variables to give a usable dataset that could be analyzed. We put either 0 or 1 in our variables except Mail ID, Timestamps, CountMessageReporter, Target and Reason variables. For example, when a phishing email had a PDF attachment, we put value "1" in our "PDFattachment" variable. Similarly, if the phishing email had a hyperlink in the content, we put value "1" in our "ContainHyperlink" variable. As we want to analyze our dataset based on Cialdini's persuasion principles, it is important for us to explain our rationale and conception based on Cialdini's principles. We explain our variables and persuasion conception in the following sections.

4.1.3.1 *Variables and concepts*

As we studied phishing email properties in subsubsection 2.5.1.3, variables are needed to code our dataset into, so that we can conduct the statistical analysis in the SPSS application. Based on our findings in the literature survey on phishing email properties, 23 Variables were created as part of the methodology processes prior to data classification. Generic properties are depicted by the structural properties in phishing emails except persuasion principles. The variables are explained in the following list:

1. *Mail ID* : Unique ID [Scale measurement]
2. *Timestamps*: Implies the date and time when the email is reported [Scale measurement]
3. *Attachments:* Indicates whether the phishing email has an attachment(s), and 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. *Instructions*: 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 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. *HiddenURL*: Specifies whether a phishing email has a hidden URL(s) [0 = No, 1 = Yes]

7. *CountMessageReporter*: A counter where the reporter includes extra information with the minimum value 0. For instance, when a reporter said "Geen spam, maar phishing!" ("Not spam, but 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 37]

9. *Reason*: Implies the reason why the unsuspected victim must grant the phisher's request

   a) ReasonType [Values can be seen in Table 38]

10. *Cialdini's Principles*: Specifies what principle(s) the phishing email signifies:

    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. 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. The reason for this variable was to make sure we can track back from 207 unique phishing emails to 440 phishing emails.

We established the variables based on the phishing email properties. We also distinguished generic properties and persuasion properties. The generic properties of a phishing email is affected by these variables: attachments, requests, contents, hiddenURL, target and reason. On the other hand, the persuasion properties are affected by these variables: reciprocation, consistency, social proof, likeability, authority and scarcity.

4.1.3.2 *Cialdini's principles and conception*

As part of our analysis, we analyzed the phishing emails dataset based on Cialdini's principles entitled "The science of persuasion". The decision-making and the rationale in this process are achieved based on our perspective of Cialdini's principles in the following details.

Reciprocation: The norm that obligates individuals to repay in kind what they have received. In other words, to return the favor, or adjust to smaller request [8]. This occurs when a phisher sends an email containing a message that is perceived as a request or obligation towards the recipient to "return the favor". It might be natural for an individual to feel "obliged" to return the favor for things or information that he/she is given and deems 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 "obliged" to return the favor of that valuable information. Another example is that if the sender gave the information that they have added "extra security" on their system we also feel obliged to grant their request.

Consistency: Public commitment occurs when people become psychologically become vested in a decision they have made [76]. This happens when a phishing email contains a message that is perceived to request the 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 have been given. Another example is if "Facebook" gives a link to change your password when you previously requested to change it. It might be not applicable to those who are not requesting password previously, but we believe it will impact on those who previously committed to change their password.

Social proof: This occurs when people model the behavior of their peer group, role models, important others or because it is generally "fashionable" [76]. For example, when someone tells us that there are hundreds of other people who use a particular system, we might want to agree to use it as well just because a lot of other people use it. Another example is when Facebook gives information that someone

wants to be our friend, and we know 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 [76]. In the context of a phishing email, the email contains a message that attracts the recipient to comply with the sender's request, based on reference on something or someone 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 music that Michael Jackson made, we might be attracted to download and listen to it just because we happen to love Michael Jackson's music. It is like someone asks us to watch a concert and they said, "Coldplay will be there". If we are a devoted fan of Coldplay, we might find it very interesting. Another example is when a sender gives compliments to us or commits to help us safeguard our account from 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 [76]. This happens when a phishing email contains a logo, image, signature or anything that looks like a legitimate institution. It can be used to makes it look trustworthy so that the recipient might accept and obey the sender's request. For example, an email may present a somewhat authentic-looking signature like "Copyright 2013 PayPal, Inc. All rights reserved" or with the PayPal logo. Cialdini [8] suggests that authoritative persuasion can be achieved by presenting an "aura of legitimacy". Another example is when the content of the email states that it is from the "System Administrator" asking for password update. It would be not authoritative if random people asked us to change our password.

Scarcity: This is based on the principle of reactance, where people respond to perceived shortages by placing greater psychological value on perceived scarce items [76]. A phishing email containing such a message tells a recipient to react or respond to scarce/becoming-scarce items, things or privileges. In the phishing email context, if a sender tells us that he/she will suspend/deactivate/limit our account if we do 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 – in other words, our account becomes scarce or limited.

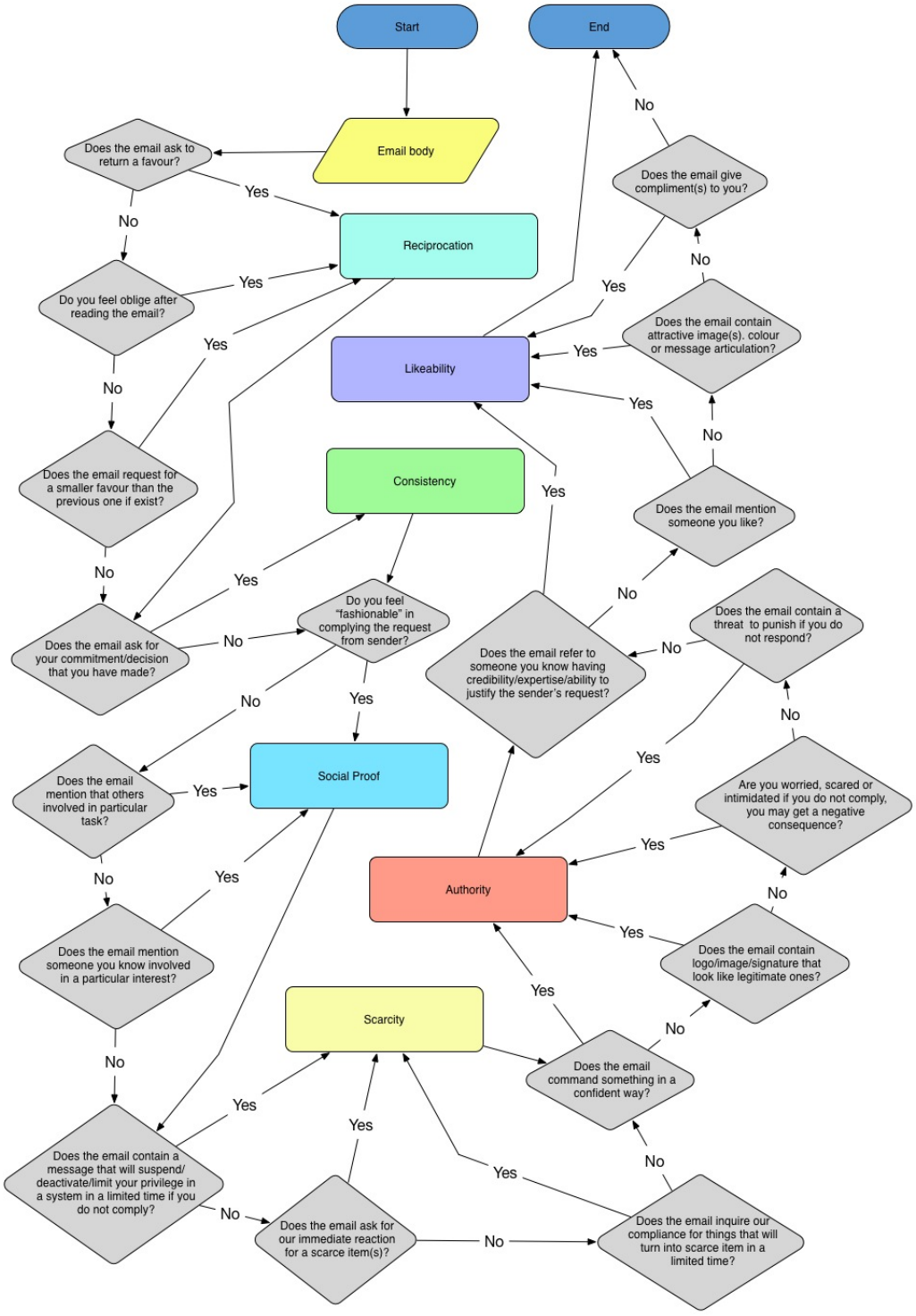


Figure 11: Integration pseudo-code of Cialdini's principles

We have made a flowchart[1] in Figure 11 that illustrates our analysis of the dataset based on Cialdini's principles.

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

4.1.4 *Statistical analysis*

In the previous section, we described the framework of our methodology in considerable detail. Until data classification, we used Microsoft Excel to code our data. To perform the analyses, we transformed the data into an SPSS readable file. We initially recorded our data in 23 variables that could be expanded depending on our analyses, such as selecting cases that have all instructions or selecting a specific target sector.

The data was analyzed by quantitative analysis from 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 3. Furthermore, we used Pearson chi-square to test our hypotheses to discover if there was a significant relationship between two variables. If the resulted p-value was less than 0.05, 0.01, or 0.001, we are 95%, 99% and 99.9% confident, respectively, that the two chosen variables have a significant relationship. By combining frequency analysis and chi-square test, we 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), we do not analyze our data by Pearson correlation. However, to test the strength of association involving nominal variables, the appropriate measurements are using phi and Cramer's V. Phi is used for 2 by 2 tables and Cramer's V can be used for more than 2 by 2 tables. Since our data is analyzed on a 2 by 2 table, Phi measurements are 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.

4.2 RESULTS

In this section, we elaborate on the results we obtained through our analyses. We begin this section by describing the frequency analyses of the general structural properties and persuasion principles. We then describe the relationship analysis between the general structural properties and the persuasion principles. We conclude the section by mentioning the results related to persuasion principles used in phishing emails.

We find that 36.2% of the total phishing emails have attachment(s) included within their content. 63.8% of them therefore do not have attachments. Of the emails with attachments, 4% have PDF attachments, 78.7% have ZIP attachments, 12% have HTML attachments and 5.3% of them have had the attachments removed before the emails were forwarded.

We are not sure what type of attachments they had, but we determined that an attachment element was there if there was a request to open an attachment within the email content. For example, in an email dated December 20th 2013 11:29am, it was said in the email's body: "...we have sent the attached as a secure electronic file". As nothing was actually attached, we suspect it was removed by antivirus software. Table 9 illustrates our findings on attachment variables.

When we look deeper, we find that there is a significant relationship between ZIP file and Attachment variable. A chi-square test resulted in $X^2(1) = 145.236, p < 0.001$. Similarly, there is a significant association between HTML file and Attachment variable with a chi-square test resulting in $X^2(1) = 16.560, p < 0.001$. However when we test the relationship between PDF and attachment, the significance level is not as strong as ZIP and HTML, with $X^2(1) = 5.358, p = 0.021$.

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

Table 9: Attachment analysis

When we look at the instructions or requests used in the dataset, we find 202 emails or (97.6% of total) with clear instructions: requests to click URL(s); requests to open attachments; requests to reply by email; or requests to respond by phone. 2.4% of the total emails do not contain a clear instructions to the recipients. For instance, on 24 November 2013 at 19:59, an email was sent that only include an attachment but no instruction to open it. However, with the subject of "Payrolls reports" we have the impression that this is a targeted phishing email, which means it only aims at a small audience as the recipients, usually a certain institution. Similarly, an email sent 15 August 2013 at 17:38 contains HTML suggesting a recipient to check the interesting pages on Facebook. However, there are no instructions to click on the URL, nor any other instructions.

Apart from the instruction to click URL(s), we find that 37.2% of the total phishing emails request to open attachments, 52.7% of them request to click URLs, 16.9% request for email replies and 4.3% request a phone call. Moreover, one single email can contain multiple requests. If we look deeper, we have 8 valid emails (3.9% of all emails)

that have requests to both open attachments and click URLs. However, we do not find any email which has all requests in the content.

Table 10 illustrates our findings in respect of requests used. Of all phishing emails that have clear instructions, 54% request to click URL(s), 38.1% request to open attachment(s), 17.3% request an email reply and 4.5% request a response by calling on the phone.

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 10: 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 analyzed the content of phishing emails in our corpus. We looked at whether they had URLs, use HTML code or included images within its content. We find that 60.4% have URLs, while 39.6% do not. 66.2% of the emails use HTML code while 33.8% do not use. Finally, 35.3% of them include images while 64.7% do not. Table 11 highlights our findings in respect of content analysis. The percentage depicted is of all total emails. If we look further at all emails that utilized HTML, 120 emails (87.6%) provided URLs, and 73 (53.3%) include images. Furthermore, of all 73 emails that include images, 67 emails (91.8%) provided URLs. Based on these 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 11: 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 the target classification table in Table 12, we find that financial sector is the most targeted sector and ISP is the least common target in our corpus. Furthermore, e-Commerce/retail sec-

tor, administrator and government are the second, third and fourth most targeted sectors, respectively. When we look deeper at the detailed list of targeted brands, we find PayPal has the highest frequency (37.2%) of the total financial targeted emails. Bank of America contributes 6.4%, American Express 5.1%, Visa contributes 5.1% and Western Union contributes 3.8%. Other financial institutions contribute to less than 3%. Figure 20 illustrates the detailed target brands of the financial sector.

Target	Frequency	Percent
Financial	78	37.7
E-commerce/retails	40	19.3
Administrator	30	14.5
Government	14	6.8
Non-existence/individual	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 12: Target analysis

As one email does not have more than 1 targeted sector, the total sums up to 100%. Note that we initially had 92 targets in our corpus and we had to classify them into 10 target types in our data classification.

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

Table 13: Reason classification

When we look at what reasons are used in Table 13, we find that 48.8% of the total emails are account related, 25% have a financial reason and 11.1% a document-related reason. Only 9.7% have a product/services reason and only 4.8% a social reason. When we break these down, account-related reasons consists of a security risk that contributes 28.7%, and system upgrade, new system requirement and account expiration that contribute below 11%. This suggests that an account-related reason is the most common pretext to manipulate recipients in our corpus while social is evidently the least common pretext. Figure 12 illustrates the detailed list of account-related reasons.

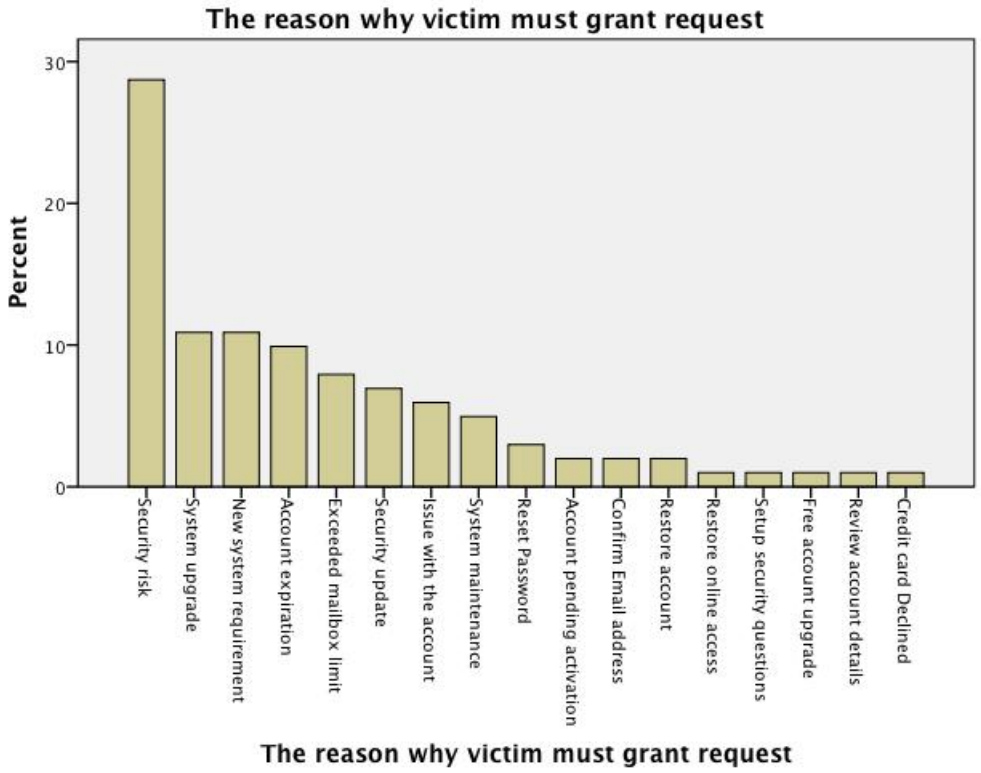


Figure 12: Detailed account related reason graph

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 14: Persuasion principles analysis

We look at the result of persuasion techniques analysis based on Cialdini's principles with our corpus. As we can see from Table 14, we find that 96.1% of the total phishing emails are using authority principle, which is the most used technique in our dataset, followed by the scarcity principle at 41.1%. 21.7% of the total use the likeability principle, while 17.4% use the consistency principle. 9.7% use the reciprocation principle and 5.3% of them use the social proof principle. Since the authority principle is the highest persuasion technique in our corpus, it is interesting to know why phishers often use author-

ity as the main technique. Perhaps, most people do not want negative consequences as a result of disobedience to authoritative figures. Consequently, those people who respond more obediently to authority are more likely to comply with the emails' requests than people who are more skeptical about authoritative figures. Note that one email can use multiple principles. Therefore, the total percentage does not sum up to 100%.

Table 15: 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	

Based on the results of persuasion principles analysis, we know that the 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 use the authority principle, while 4.1% of them do not impersonate government nor use the authority principle. We find 100% of government-targeted emails have the authority principle. On the other hand, we find 93% of all authority emails are non-government targeted emails and 7% of them are government-targeted emails. Table 15 depicts the relationship between government-targeted emails and the authority principle. Furthermore, of all phishing emails, 6.8% of them both use the authority principle and target the government sector. A chi-square test was performed and we find that there is no significant association between the government sector and authority principle, as $X^2(1) = 0.604, p = 0.473$. Since p is not less than 0.05, we reject hypothesis 1.

Table 16: 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	

When we look at the relationship between phishing emails that impersonate administrators and the authority principle to test hypothesis 2, we find that 96.7% of administrator-targeted emails use the authority principle and 96% of non-administrator emails use the 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, we reject hypothesis 2. Table 16 highlights the relationship between administrator sector and the authority principle.

Table 17: 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		

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 that target financial sector use the scarcity principle, while 60.3% do not. Furthermore, 41.9% of all non-financial targeted emails use the scarcity principle, while 58.1% do not. 63.5% of scarcity emails are non-financial targeted emails, while 36.5% of them are financial targeted emails. We performed a chi-square test and we found that there is no significant association between the financial sector and scarcity principle, with $X^2(1) = 0.090, p = 0.764$. Since p is not less than 0.05, we reject hypothesis 3. Table 17 illustrates the relationship between financial-targeted emails and the scarcity principle.

Table 18: E-commerce/retail 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		

Turning our attention to the association between phishing emails that target the e-commerce/retail sector and the likeability principle, we can test hypothesis 4. We find that 20% of e-commerce/retail sector targeted emails use the likeability principle. Furthermore, 22.2% of non e-commerce/retail sector targeted emails use the likeability principle. On the other hand, only 17.8% of all likeability emails are e-commerce/retail sector targeted emails. A chi-square test was performed and we found that there is no significant association between phishing emails targeting the e-commerce/retail sector and the likeability principle, with $X^2(1) = 0.088, p = 0.767$. Since p is not less than 0.05, we reject hypothesis 4. Table 18 illustrates the relationship between the e-commerce/retail targeted sector and the likeability principle.

Table 19: 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		

Now we look at the association between phishing emails targeting social media and the social proof principle to test hypothesis 5. We find that 18.2% of social media targeted emails employ the social proof principle. Furthermore, 4.6% of non-social media targeted emails employ the social proof principle. 18.2% of all social proof emails are social media targeted emails and 81.8% of them are not. A chi-square test was performed and we found that there is no significant association between phishing emails targeting social networks and social proof principle, with $X^2(1) = 3.823, p = 0.051$. Therefore since p is not less than 0.05, we reject hypothesis 5. Table 19 depicts the relationship between social media and the social proof principle.

Table 20: 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	

Next, we look at the relationship between authority principle and scarcity principle to test hypothesis 6. Based on the result in Table 20, we find that 41.7% of authoritative emails use the scarcity principle while 58.3% do not. However, we find that 97.6% of all scarcity emails use the authority principle and only 2.4% of them do not. A chi-square test suggests that there is no significant relationship between the authority principle and the scarcity principle, with $X^2(1) = 0.887$, $p = 0.346$. Thus, we reject hypothesis 6.

Table 21: 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).

We now consider the relationship between the likeability principle and the consistency principle to test hypothesis 7. Based on our results in Table 21, we find that only 6.7% of likeability emails have the consistency principle while 93.3% of them do not. In addition, we find that 20.4% of non-likeability emails have the consistency principle while 79.6% of them do not. 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 the likeability principle and the consistency principle, with $X^2(1) = 4.603, p = 0.032$. Phi measurement suggests a very weak negative (inverse) relationship at -0.149, indicating that as one variable increases, the other variable decreases. This suggests that the higher the use of the likeability principle in

a phishing email, the less chance of the consistency principle being used. Thus we accept hypothesis 7 that says the occurrence of the likeability principle will impact the occurrence of consistency.

Table 22: URL presence and hidden URL

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

*** p < 0.001 (significant).

Now we move on to find out the association between URL presence and hidden URLs in our corpus to test hypothesis 8. Based on our results in Table 22, we find that 76% of URLs are hidden while 24% are not. A chi-square test suggests that there is a highly significant association between URL presence and hidden URLs, with $X^2(1) = 115.191, p < 0.001$. Moreover, Phi measurement suggests a strong positive relationship at 0.746. This indicates a strong relationship between them, so we accept hypothesis 8.

Table 23: 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).

We now look at the relationship between URL presence and the emails which request to click on URLs in Table 23 to test hypothesis 9. We find 87.2% of phishing emails that have URLs also requested receivers to click on them, while 12.8% did. A chi-square test was performed and suggests that there is a highly significant relationship between URL presence and request to click on URLs, $X^2(1) = 151.034, p < 0.001$. Phi measurement suggests that they have

a strong positive relationship at 0.854. Thus, this data supports hypothesis 9.

Table 24: 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).

Similarly, we look at the association between the emails that include attachments and the emails which request to open attachments to test hypothesis 10. Based on our results in Table 24, 96% of phishing emails that include attachments also have a request for the attachments to be opened, while only 4% do not. A chi-square test was performed and suggests that there is a significant relationship between URL presence and request to click URL, with $X^2(1) = 174.079, p < 0.001$. Phi measurement suggests that they have a strong positive relationship at 0.917. Therefore, we accept hypothesis 10.

Table 25: 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		

To test hypothesis 11, we look at the relationship between the authority principle and emails that include images. Based on our results in Table 25, 35.7% of authoritative emails include images, while 25% of non-authority emails include images. 97.3% of emails that include images are authority emails and 95.5% of emails that do not include image(s) are authority emails. A chi-square test was performed and suggests that there is no significant relationship between authority principle and image presence, with $X^2(1) = 0.384, p = 0.535$. Thus, based on this result, we reject hypothesis 11.

Table 26: 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).

Now we look at the association between account-related reasons and the scarcity principle to test hypothesis 12. Based on the results in Table 26, 68.3% of account related phishing emails feature the scarcity principle, while 31.7% of them do not. 81.2% of scarcity emails have account-related reasons 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, with $X^2(1) = 60.535, p < 0.001$. Phi measurement suggests that they have a strong positive relationship at 0.541. Therefore, we accept hypothesis 12.

Table 27: 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).

Furthermore, we look at the relationship between account-related reasons and URL presence to test hypothesis 13. Based on the results in Table 27, we find 78.2% of account-related emails include URLs and 63.2% of emails that include URLs are account-related emails. Furthermore, 38.2% of total phishes are account-related and include URLs. A chi-square test was performed and suggests that there is a significant relationship between these two variables, with $X^2(1) = 26.216, p < 0.001$. Phi measurement suggests that they have a weak positive relationship at 0.356. Therefore, we accept hypothesis 13.

Table 28: 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).

To test hypothesis 14, we now look at the relationship between document-related reasons and the government sector. From Table 28, we find only 21.7% of document-related phish emails targeted government while 78.3% of them did not. However, a chi-square test suggests that there is a highly significant relationship between these variables, with $X^2(1) = 9.203, p = 0.002$. Phi measurement indicates that they have a weak positive relationship at 0.211. Therefore, we accept hypothesis 14.

Table 29: 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).

Now we look at the relationship between document-related reasons and attachment variables to test hypothesis 15. Based on our results in Table 29, 78.3% of document-related phish emails have attachments included, while 21.7% of them do not. A chi-square test suggests that there is a significant relationship between these variables, with $X^2(1) = 19.783, p < 0.001$. Phi measurement indicates that they have a weak positive relationship at 0.309. However, the result still supports hypothesis 15.

Table 30: The use of HTML and likeability

Content	Non-likeability	Likeability	N
Not use HTML	61	9	70
Use HTML	101	36	137
N	162	45	207
Pearson chi-square		4.904*	

* $p < 0.05$ (significant).

Lastly, we look at the association between HTML usage variables and the likeability principle to test hypothesis 16. Based on the result in Table 30, we find 80% of likeability phish emails use HTML, while 20% of them do not. 37.7% of non-likeability emails do not use HTML and 62.3% do. 26.3% of emails that use HTML are likeability emails. 17.4% of total phishes use HTML and the likeability principle. A chi-square test suggests that there is a significant relationship between these variables, with $X^2(1) = 4.904$, $p = 0.027$. Phi measurement suggests that have a weak positive relationship at 0.154. Although they have a weak relationship, HTML variable and the likeability principle still have a significant relationship. Thus, we accept hypothesis 16.

4.2.1 *Relationship between persuasion principles and target types*

We have seen the results according to the research questions and hypotheses in chapter 3. As we mentioned earlier, we find a significant relationship between administrator and scarcity. It is important for us to know whether the other target types and persuasion principles share any kind of relationship so that we can compare our findings and strengthen our conclusion.

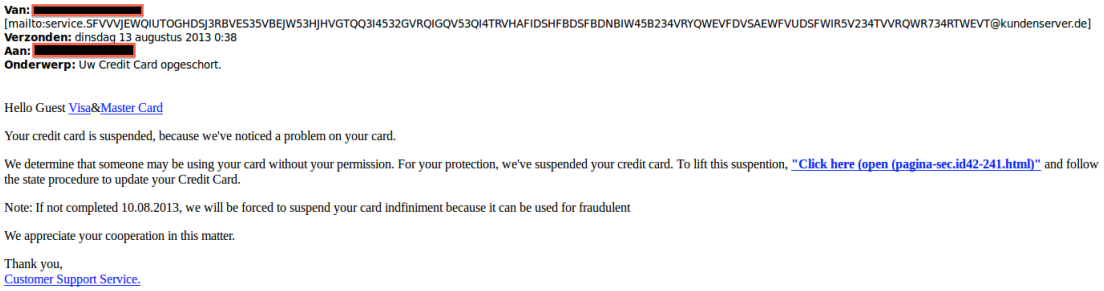
Table 31: Persuasion principles vs Target types in percentage

	Authority	Scarcity	Likeability	Consistency	Reciprocation	Social Proof	N
Financial	98.7	39.7	29.5	24.4	16.7	2.6	78
E-Commerce / Retails	100.0	60.0	20.0	5.0	5.0	5.0	40
Administrators	96.7	66.7	16.7	3.3	0.0	0.0	30
Government	100.0	7.1	0.0	35.7	7.1	21.4	14
Non-existence / Individual	61.5	23.1	23.1	0.0	15.4	15.4	13
Social media	100.0	0.0	36.4	18.2	0.0	18.2	11
Postal services	100.0	44.4	11.1	11.1	0.0	0.0	9
Travel agencies	80.0	20.0	0.0	60.0	20.0	0.0	5
Industrials	100.0	0.0	20.0	40.0	20.0	0.0	5
ISP	100.0	50.0	0.0	50.0	0.0	0.0	2

Note: N = total number

It is clear from Table 31 that the authority principle contributes high percentages among all target types, whereas social proof principle contributes the least in all target types. The highest percentage of social proof principle is used in government target type (21.4%), but this is still low compared to the use of consistency and authority principles. The social proof principle is not used in administrators, social media, postal services, travel agencies, industrials and ISP target types.

Depending on the target types, we can observe the next most popular principle for financial (39,7%), e-commerce/retail sector (60.0%) and administrator (66.7%) targets is the scarcity principle. When we look into our dataset and investigate why scarcity is the second most used principle, we can notice from Figure 13, Figure 14 and Figure 15 that these three target types use something that might be valuable that belong to the recipients:. accounts. By this reasoning, it makes sense if phishers that impersonate financial, e-commerce/retails and administrator typically use the scarcity principle.



Van: [redacted]
[mailto:service.SFVVVJEWQIUTOGHDSJ3RBVES35VBEJW53HJHVGTQQ3I4532GVRQIGQV53QI4TRVHAFIDSHFBDSFBDNBIW45B234VRYQWEVFDVSAEWFVUDSFWIR5V234TVVRQWR734RTWEVT@kundenserver.de]
Verzonden: dinsdag 13 augustus 2013 0:38
Aan: [redacted]
Onderwerp: Uw Credit Card opgeschort.

Hello Guest Visa&Master Card

Your credit card is suspended, because we've noticed a problem on your card.

We determine that someone may be using your card without your permission. For your protection, we've suspended your credit card. To lift this suspention, **"Click here (open (pagina-sec.id42-241.html)"** and follow the state procedure to update your Credit Card.

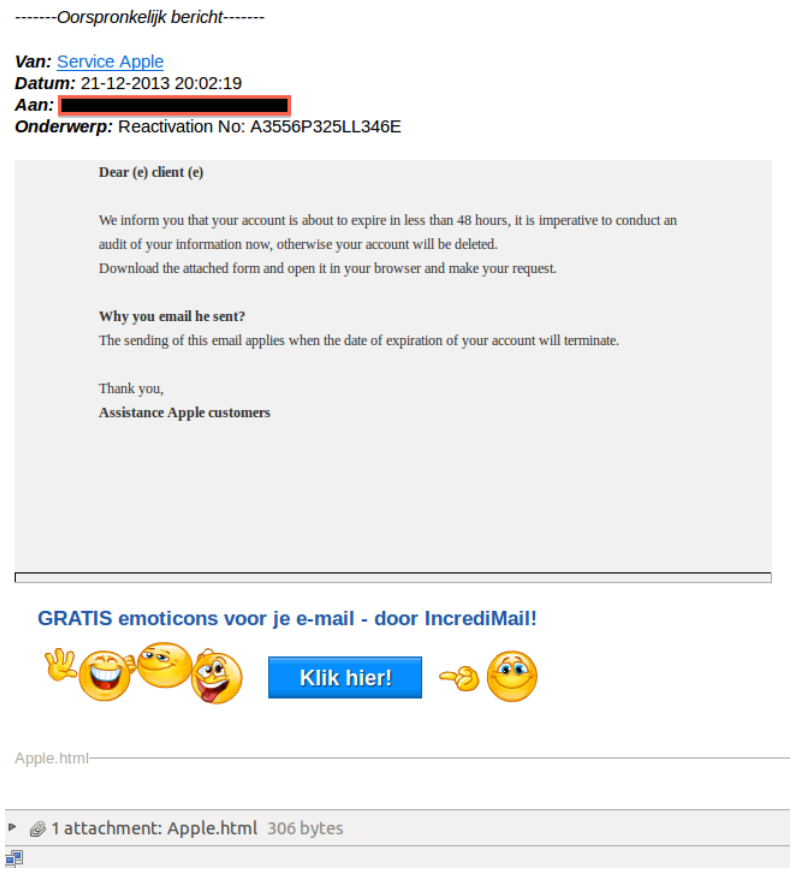
Note: If not completed 10.08.2013, we will be forced to suspend your card indfiniment because it can be used for fraudulent

We appreciate your cooperation in this matter.

Thank you,
Customer Support Service.

Figure 13: Financial target and scarcity

Figure 13 shows a financial targeted email (Visa and MasterCard) stating: "Your credit card is suspended,". In other words, the recipient's belongings will be scarce indefinitely if the recipient does not respond to the email within a limited time. A similar scenario is illustrated in Figure 14 and Figure 15 – e-commerce/retails and administrator targeted emails respectively – which mention account issues with a limited period of time to respond.



-------*Oorspronkelijk bericht*-------

Van: Service Apple
Datum: 21-12-2013 20:02:19
Aan: [redacted]
Onderwerp: Reactivation No: A3556P325LL346E

Dear (e) client (e)

We inform you that your account is about to expire in less than 48 hours, it is imperative to conduct an audit of your information now, otherwise your account will be deleted.
Download the attached form and open it in your browser and make your request.

Why you email he sent?
The sending of this email applies when the date of expiration of your account will terminate.

Thank you,
Assistance Apple customers

Apple.html

1 attachment: Apple.html 306 bytes

Figure 14: E-Commerce/Retails and scarcity

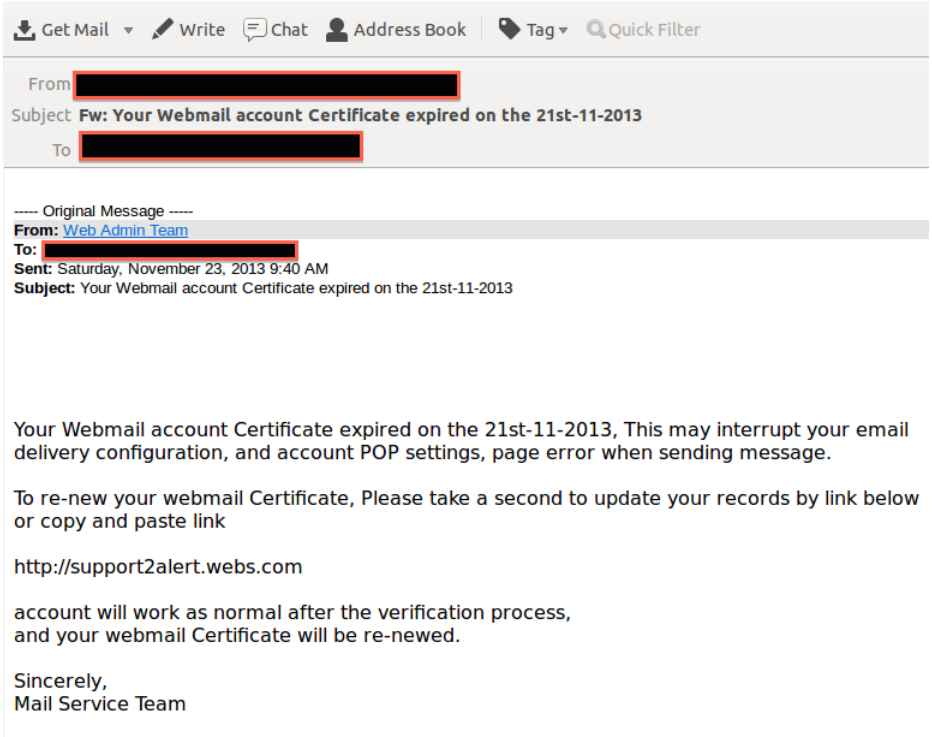


Figure 15: Administrator and scarcity

Consistency is the next most popular principle (35.7%) for government target type. When we look into our dataset and observe two government targeted emails, we find that both required consistency from the recipient. From Figure 16, the email stated "This message has been generated in response to the company complaint submitted to Companies House WebFilling service", implying that the email has been sent due to a complaint submitted previously. If the recipient has in reality submitted a complaint they might feel committed to respond to this email. We can also observe similar scenario from Figure 17, with the email stating that "...you have been scheduled to appear for your hearing...," implying that the sender required a public commitment from the recipient. However, this scenario may not impact those who do not have involvement in the targets chosen by the phishers.

Van: "Companies House" <webfiling@companieshouse.gov.uk>
Onderwerp: FW: Case IRU6YCFJE864IR6
Datum: 15 oktober 2013 (w42) 15:17:13 CEST
Aan: [redacted]

This message has been generated in response to the company complaint submitted to Companies House WebFiling service.

(CC01) Company Complaint for the above company was accepted on 15/10/2013.

The submission number is IRU6YCFJE864IR6

Please quote this number in any communications with Companies House.
All WebFiled documents are available to view / download for 10 days after their original submission. However it is not possible to view cop

Not yet filing your accounts online? See how easy it is...

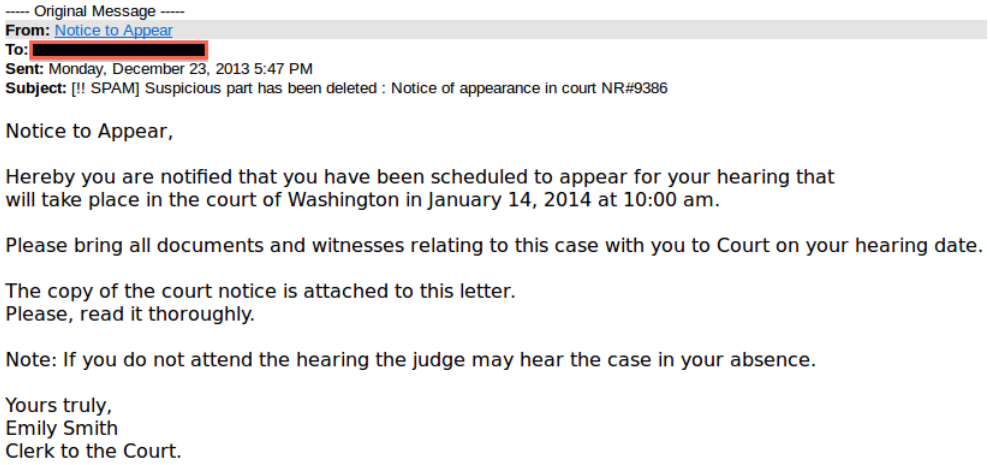
Note: reference to company may also include Limited Liability Partnership(s).

Thank you for using the Companies House WebFiling service.

Service Desk tel +44 (0)303 3989 520 or email enquiries@companieshouse.gov.uk

Note: This email was sent from a notification-only email address which cannot accept incoming email. Please do not reply directly to this

Figure 16: Government and consistency (a)



----- Original Message -----
From: Notice to Appear
To: [redacted]
Sent: Monday, December 23, 2013 5:47 PM
Subject: [!! SPAM] Suspicious part has been deleted : Notice of appearance in court NR#9386

Notice to Appear,

Hereby you are notified that you have been scheduled to appear for your hearing that
will take place in the court of Washington in January 14, 2014 at 10:00 am.

Please bring all documents and witnesses relating to this case with you to Court on your hearing date.

The copy of the court notice is attached to this letter.
Please, read it thoroughly.

Note: If you do not attend the hearing the judge may hear the case in your absence.

Yours truly,
Emily Smith
Clerk to the Court.

Figure 17: Government and consistency (b)

It is interesting that we do not find the scarcity principle in social media target type. If we put ourselves as potential victims getting an email from social media, our account in social media may be less important than our account in the financial sector (e.g. bank). On the other hand, our desire to respond to attractiveness in social media targeted emails might be higher. This explains why we find likeability as the next most popular principle instead of scarcity principle in social media target type.

Table 32: Chi-square tests of Persuasion principles vs Target types

Relationship	Authority	Scarcity	Likeability	Consistency	Reciprocation	Social Proof
Financial	2.247	0.90	4.416*	4.230	7.036**	1.881
E-Commerce/Retails	1.993	7.347**	0.088	5.299*	1.235	0.010
Administrator	0.027	9.504**	0.531	4.826*	-	-
Government	0.604	7.139**	-	3.509	0.109	7.749**
Non-existence/Individual	44.687***	1.854	0.015	-	0.520	2.796
Social media	0.467	-	1.460	0.005	-	3.823
Postal services	0.378	0.044	0.625	0.258	-	-
Travel agencies	3.590	0.939	-	6.475*	0.627	-
Industrials	0.206	-	0.009	1.823	0.627	-
ISP	0.081	0.067	-	1.495	-	-

Note: df = 1, *p < .05, **p < .01, ***p < .001.

When we look at the relationship between target types (sectors) and persuasion principles in Table 32, we do not find a significant relationship between the financial sector and the scarcity principle. However, we find that the financial sector has a significant relationship with the reciprocation principle. This explains that even if the number of financial sector targeted emails is low in terms of reciprocation (16.7%), the reciprocation principle is more likely to be used in the financial sector than the other sectors. Moreover, we find that the e-commerce/retail sector- and administrator-targeted emails have a significant relationship with the scarcity principle. This supports our previous finding that scarcity is the next popular principle in both target types. However, the p-value indicates that the administrator target type has a more statistically significant relationship with the scarcity principle than the other two sectors. Although we also find a significant relationship between government and scarcity, phi measurement indicates they have an inverse relationship (phi = -0.186). This explains that the scarcity principle is unlikely to be used in government-targeted emails. This also may supports our finding that consistency is the next most used principle in the government target. Despite the fact that we find a significant relationship between non-existence/individual sector and authority, when we look deeper, we find that they have an inverse relationship (phi = -0.465). This suggests that the authority principle is likely not to be used in non-existence/individual targets. This explains why the occurrence of au-

thority is lower in non-existence/individual targets (61.5%) than the other target types.

4.2.1.1 *Findings*

Based on our analysis of the relationship between persuasion principles and target types, we learn that depending on the target types, three persuasion principles – scarcity, consistency and likeability – are the next most popular persuasion principles in our dataset.

4.2.2 *Relationship between persuasion principles and reason types*

Another important aspect of a phishing email is the reason that is used by the phishers as a pretext to trick the recipients. Apart from the result in Table 26 which implies a strong relationship between account-related reasons and the scarcity principle, it is important for us to compare and strengthen our findings by seeing whether the other reason types and persuasion principles have any kind of relationship.

Table 33: Persuasion principles vs Reason types in percentage

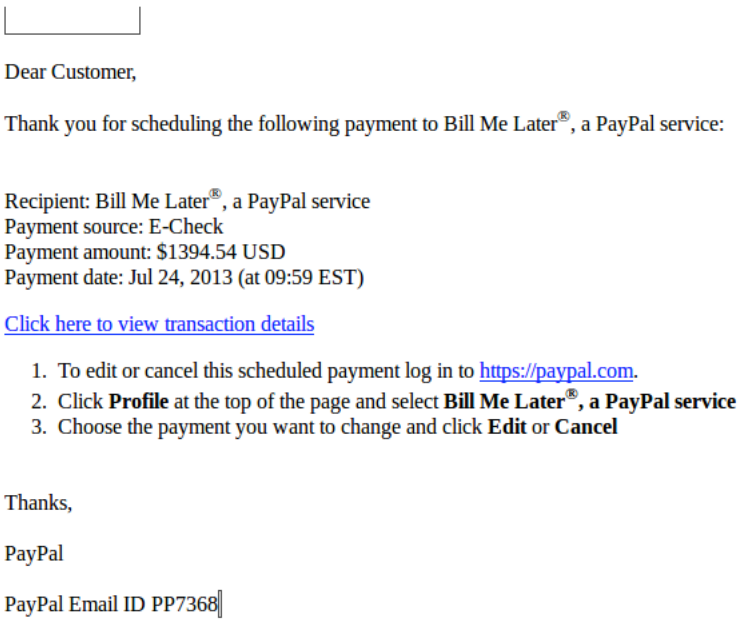
	Authority	Scarcity	Likeability	Consistency	Reciprocation	Social Proof	N
Account related	100.0	68.3	25.7	10.9	11.9	4.0	101
Financial incentive	92.5	20.8	20.8	26.4	15.1	5.7	53
Document related	95.7	4.3	4.3	34.8	0.0	8.7	23
Product/services	100.0	20.0	15.0	15.0	0.0	0.0	20
Social	80.0	0.0	40.0	0.0	0.0	20.0	10

Note: N = total number

As is shown by Table 33, apart from the authority principle, we find less than 50% contributions of likeability, consistency, reciprocation and social proof principles to all of the reason types. Only the scarcity principle is often used in account-related reasons (68.3%). We find reciprocation is the least principle used in terms of the reason types. We do not find reciprocation principle in document related, product/services and social reasons.

We also find that the next most popular principle for account- related reasons (68.3%) and product/services reasons (20.0%) is the scarcity principle. The illustrations in Figure 13, Figure 14 and Figure 15 perhaps explain why the scarcity principle tends to be used in account-related reasons. Our reasoning is that as recipients we might value our accounts in a certain system more, so that we tend to re-

spond to an email that requests us to act in order to prevents the loss of our valuables within a limited time



Dear Customer,

Thank you for scheduling the following payment to Bill Me Later®, a PayPal service:

Recipient: Bill Me Later®, a PayPal service
Payment source: E-Check
Payment amount: $1394.54 USD
Payment date: Jul 24, 2013 (at 09:59 EST)

Click here to view transaction details

1. To edit or cancel this scheduled payment log in to https://paypal.com.
2. Click **Profile** at the top of the page and select **Bill Me Later®, a PayPal service**
3. Choose the payment you want to change and click **Edit** or **Cancel**

Thanks,

PayPal

PayPal Email ID PP7368

Figure 18: Example of financial incentive and consistency

Based on Table 33, the next most popular principle for financial incentives (26.4%) and document-related reasons (34.8%) is consistency. The email illustrated in Figure 18 states: "Thank you for scheduling the following payment..." This indicates a financial incentive and request for commitment from the recipients that have scheduled payments through PayPal. It is natural for us that we treat financial as a sensitive matter. This may raise our curiosity about when we scheduled a payment through PayPal, and so we click on the link. The recipients who have scheduled a payment through PayPal previously will have even more incentive to click on the URL provided in the email. This explains why consistency is the second most-used principle in terms of financial incentive. Moreover, consistency is also the second most popular principle in terms of document-related reasons. As we observed in Figure 17, the phrases "you have been scheduled..." and "the court notice is attached..." indicate that the email requests for commitment from the recipient for a decision they have previously made, and also portrays a document-related reason. The average recipients who perceive document-related reason as more formal will likely respond to the email. This might explain why consistency is the second most popular principle in terms of document-related reasons.

We found that likeability is the second most popular principle in social reason emails. This can relate to our previous analysis in sub-

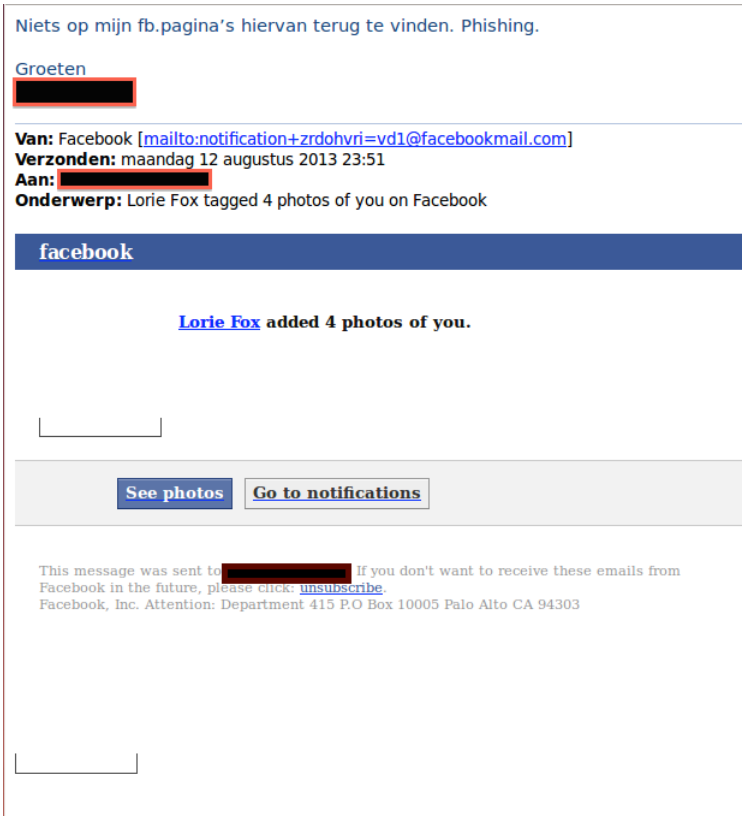


Figure 19: Social reason and likeability principle

section 4.2.1 regarding social media and the likeability principle. As illustrated in Figure 19, by incorporating likeability and social incentives, the recipients might want to respond to the email much more than if other principles were employed. This explains why likeability is the second most used principle in social reason emails.

Table 34: Chi-square tests of Persuasion principles vs Reason types

Relationship	Authority	Scarcity	Likeability	Consistency	Reciprocation	Social Proof
Account related	4.387*	60.535***	1.858	5.801*	1.113	0.718
Financial incentive	2.600	12.140***	0.041	4.038*	2.409	0.017
Document related	0.016	14.417***	4.600*	5.447*	-	0.558
Product/services	0.890	4.058*	0.591	0.088	-	-
Social	7.363**	-	2.059	-	-	4.504*

Note: df = 1, *p < .05, **p < .01, ***p < .001.

When we look at Table 34, we find that both account-related and product/services reasons have significant relationships with the scarcity principle. This supports our previous analysis that finds scarcity is the

second most used principle for both reasons. We also find both financial incentive and document-related reasons have significant relationships with scarcity. However, phi-measurements indicate that both have inverse relationships: phi = -0.242 for financial incentives and phi = -0.264 for document-related reasons. This explains why consistency is the second most used principle for financial- and document-related reasons. We find a significant relationship between social reasons and the authority principle. The phi measurement suggests that they also have an inverse relationship (phi = -0.188). It means that authority principle is most likely not to be used in phishing emails with social reasons. We believe it makes sense if the social reason type does not project fear or negative consequences to persuade potential victims. This also explains why likeability is the second most used principle in social reason emails.

4.2.2.1 *Findings*

Based on our analysis of the relationship between persuasion principles and reason types, we can conclude that depending on the reason types, three persuasion principles (scarcity, consistency and likeability) are still the second most popular persuasion principles used in our dataset.

4.2.3 *Target types and reason types*

It is important for us to explain that target types do not always use the matching reason types. For instance, financial sector targeted emails do not always use financial incentives to trick the victims; they can use an account-related or document-related reason. Similarly, an administrator targeted email does not always use an account-related reason. To illustrate, a phishing email reported on 24 November 2014 7:59 PM claiming to be from an "Administrator" does not mention about an account issue, but instead asks the recipient to download an attachment related to "payroll". Table 35 illustrates the frequency analysis of target types vs. reason types.

Table 35: Frequency analysis target types vs reason types

	Account Related	Financial incentive	Document related	Product / services	Social	**N**
Financial	42	23	9	4	0	78
E-Commerce/Retails	26	9	2	2	1	40
Administrator	24	2	1	2	1	30
Government	0	9	5	0	0	14
Non-existence/Individual	3	3	3	2	2	13
Social media	3	1	1	0	6	11
Postal services	1	1	0	7	0	9
Travel agencies	1	1	2	1	0	5
Industrials	0	3	0	2	0	5
ISP	1	1	0	0	0	2
N	101	53	23	20	10	**207**

We find interesting characteristics regarding the second most used principle by observing the highest occurrence of reason types in terms of the target types. The observation gave the following characteristics:

- Financial - account related = scarcity principle
- E-commerce/retail sector - account related = scarcity principle
- Administrator - account related = scarcity principle
- Government - financial incentive = consistency principle
- Social media - social reason = likeability principle

These characteristics support our findings in subsubsection 4.2.1.1 and subsubsection 4.2.2.1 that indicate that depending on the target types and reason types, three persuasion principles (scarcity, consistency and likeability) are the next most popular principles used in our dataset.

5 DISCUSSION

At the beginning of our research, we stated two research questions that need to be answered. In this section, we discuss our findings to answer these research questions.

5.1 RESEARCH QUESTIONS

What are the characteristics of reported phishing emails?

In chapter 3, we defined seven parameters to characterize the phishing emails in our dataset. Based on our findings, we can conclude the following points:

- Based on Table 9, when attachment(s) are included in a phishing email, they are likely to be ZIP or HTML files.
- Requesting to click a URL(s) is the most prevalent instruction in phishing emails. Table 10 illustrates this finding.
- As we illustrated in Table 11, most of the phishing emails use HTML code and provide URL(s).
- As Table 12 shows, the financial sector is the most common target.
- Table 13 depicts that most of the phishing emails use account-related concerns as a pretext.
- Based on our finding from Table 14, the authority principle is the most-used persuasion technique in phishing emails.
- As we illustrated in Table 27, a phishing email that has an account-related concern as a pretext is likely to include URL(s).
- It can be seen from Table 23 and Table 24 that phishers provide clear instructions on how recipients are meant to act; phishing emails that include attachment(s) are likely to include a request to open it and phishing emails which provide URL(s) are likely to request to click on it.
- Based on our finding in Table 22, the URL(s) in a phishing email are most likely different from the actual destination.
- Table 28 suggests that a government-targeted phish is likely to have a document-related reason.

- As we illustrated in Table 29, phishing emails that have document-related reasons as a pretext are likely to include attachment(s).

Moreover, it is worth pointing out that our finding on the detailed financial sector in Figure 20 indicates many of them are non-Dutch based financial institutions such as Bank of America, Barclays Bank, and Lloyds Bank. However, we found this is not ideal because our dataset came from a Dutch-based organization. Perhaps this is because we conducted the analysis only on 207 unique phishing emails in the English language, which is 2.45% of the total reported emails. We explain why we can only analyze 207 unique phishing emails in the English language in section 5.3. Despite this limitation, we emphasize that our current study can be seen as a precursor to a larger study of persuasion techniques and phishing emails in general. By this reasoning, our method or measure instruments such as algorithms, flowcharts, and variables would not be biasing the method to what we may find in the reported phishing emails in the Dutch language.

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

To answer the second research question, we look at the relationships between the persuasion principles and the generic properties. With this in mind, we have established 10 hypotheses related to these relationships and we look whether the findings are consistent with these hypotheses. Table 36 summarizes the overview of verified hypotheses. Because almost all phishing emails use the authority principle, this implies all phishing email properties related to the authority principle resulted in no significant relationship.

When we look at the 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 the financial sector and scarcity principle, we see that administrator-targeted emails are likely to have scarcity principle. In contrast, non-administrator targeted emails are less likely to have scarcity principle. However, our findings suggest that the strength of association between administrator-targeted emails and the scarcity principle is weak.

The next finding on the relationship between e-commerce/retail sector targeted emails indicates that this sector less employs the likeability principle as both e-commerce/retail sector and non- e-commerce/retail sector targeted emails have a high number featuring the non-likeability principle. Similarly, our data suggests that there are no significant association be-tween social media targeted emails and the social proof principle.

Our data suggests that there is a significant association between likeability and consistency. According to our findings, the higher the likeability, the lower the chance of featuring the consistency principle.

This support our hypothesis that says that the occurrence of likeability will impact the occurrence of consistency. However, we find the strength of association between the variables is 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 reason, it will likely use the scarcity principle as a persuasion technique. Moreover, the result suggests that account-related reasons and the scarcity principle have a strong relationship.

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

Table 36: Overview of verified hypotheses

Hypotheses	Category	Accept	Reject
H1	A		X
H2	A		X
H3	A		X
H4	A		X
H5	A		X
H6	A		X
H7	A	X	
H8	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

Overall, seven hypotheses in respect of persuasion principles are rejected and three of them are accepted. Based on this result and supported by our analysis of the relationship between persuasion principles and target types in subsection 4.2.1 and the relationship between persuasion principles and reason types in subsection 4.2.2, we can answer our second research question with the following underlying perspectives:

- The extensive use of authority as a persuasion technique in phishing emails as opposed to social proof technique. However, our analysis suggests that while the percentages are still high, authority principle is less likely to be used in individual target type and social reason type (see Table 14, Table 31 and Table 33).
- Depending on the target types and the reason types, three persuasion principles – scarcity, consistency and likeability – are the next most popular principles used in our dataset (see subsection 4.2.1 and subsection 4.2.2).
- The scarcity principle will likely be used when phishing emails come from the administrator target type and account reason type (see Table 32 and Table 34).
- The likeability principle affects the usage of HTML-based email and consistency principle (see Table 30 and Table 21).

5.2 CONCLUSION

Our research was aimed at understanding how phishing emails use persuasion techniques. The analysis consists of finding relationships between persuasion techniques and generic properties of phishing emails.

The findings may be influenced by the fact that only one person (the author) has coded the emails. Although we made a flowchart in Figure 11 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 3, we have been able to find the characteristics of phishing emails based on persuasion techniques. Our approach has proven useful in identifying critical characteristics and relationships between generic properties of phishing emails and persuasion techniques. Three important findings of our research are that: (1) authority is the most popular persuasion technique regardless of the target and the reason used; (2) depending on the target types and the reason types, the next most popular persuasion principles are scarcity, consistency and likeability; and (3) scarcity principle has a high involvement with administrator target type and account-related concerns.

When we relate between target types and the reason used in phishing emails, our suggestions for preventing phishing can be described in the following points:

- If we assume that most people are more likely to comply with authority, we suggest a legitimate institution should never use emails as a medium of communication with its customers. Instead, a legitimate institution should have its own secure messaging system to communicate with its customers. This may reduce the risk of costumers believing that phishing emails are real.

- Even if a legitimate institution uses emails, they may use a simple email validation system such as Sender Policy Framework (SPF), which is designed to verify the sender's email server before delivering all legitimate email to the intended recipients. This can prevent a spoofed email being delivered to the intended victim.

- Legitimate institutions such as banks could use what its customers have, such as a phone number registered in the system or a token given by the bank to the customers. A secret code in the email sent by the bank should match the code delivered to the customer's phone or token. This would provide a two-factor authentication and would make it more difficult for phishers to spoof bank emails.

- It might be useful if security experts can create a library that contains the most common words or phrases that signify authority and scarcity principles, so that incoming email could be filtered using the library. Our flowchart can be useful to help the development of the library so that the conventional phishing email detection can be improved.

- Our findings suggest that account-related is the most used reason in phishing emails. Therefore, we suggest that anti phishing filtering systems should also focus on detecting account-related emails to prevent them from being delivered to the intended victim.

- Persuasion awareness in phishing emails is needed to help the end users think before they respond to an email, and to enhance users' ability to identify phishing emails.

Overall, the reflection from this research is that the phishers are not only utilizing a technical strategy to trick the unsuspecting victims, but also persuasion principles to obtain positive responses from the victims. Our research exhibits an important aspect of phishing emails so that future phishing email countermeasures should not only be

developed from a technical perspective but they should also be able to resist from persuasion misuse. Continued research on persuasion techniques in phishing emails is required to stay ahead of the phishers. Our method is a solid starting point in a real world analysis to identify the underlying issue in phishing emails.

5.3 LIMITATION

Although the research produced conclusive results, our findings need to be assimilated in the backdrop of some limitations that arose due to the complex nature of our methodology and the 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 got the data from only one organization. Our study is totally dependent on the information documented by this organization, and we do not know whether the sample data represents the Netherlands overall or represents a certain area or criterion.

The second limitation is that sometimes the emails did not show the complete structures because the reporter forwarded a suspected phishing email as an attachment, which removes essential elements of it such as any 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 the language barrier. Few of the English-based phishing emails forwarded by the reporter had information in Dutch. It might be useful to know to understand the information provided by the reporter. Without understanding non-English language, we could also analyze the structural properties of phishing emails, such as whether the email uses HTML or whether it contain hidden URL. However, since our analysis was aimed at analyzing persuasion techniques in phishing emails, we need to have the language proficiency to know which persuasion techniques were used.

The fourth limitation is that our data classification was done by one person. This 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 persuasion principles could be an issue in terms of accuracy. For example, one person can claim an email is attractive while another person can claim it is not. This introduces the greatest limitation to our research because it significantly impacts our results.

The fifth limitation is the unique dataset of the reported phishing emails (440 phishing emails reduced to 207 unique phishing emails). This resulted in a smaller sample size of data and therefore it was a challenging task to find associations using Pearson chi-square method.

5.4 FUTURE WORK

A follow-up study to analyze phishing emails with Dutch language, which we did not observe, 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 adding resilient validation in data classification in terms of persuasion principles by involving several people to have an objective decision. It is also interesting to identify the authority principle in regular emails to make an objective perspective and compare with phishing emails. By looking at our findings, future study regarding the success of authority principle in phishing emails could be useful if we have data of phishing emails that have already claimed a victim. Thus, the goal for future research would be the success rate of the authority principle in phishing emails. In conjunction, a controlled environment to test persuasion awareness would be helpful to see whether it reduces phishing victimization through emails or not. Finally, as we understand that persuasion principles in phishing email have some influence in user's decisions, it would be interesting if future research can build a simple game in terms of persuasion awareness to grab a user's attention to make the right decision. For instance, the flowchart in Figure 11 can be adapted to a "snakes and ladders" game to alert users of the presence of persuasion principles in an email they receive.

A

APPENDICES

A.1 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 37: Target classification

A.2 REASON TYPES

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

Table 38: Reason classification

A.3 FINANCIAL TARGETED PHISHING EMAILS

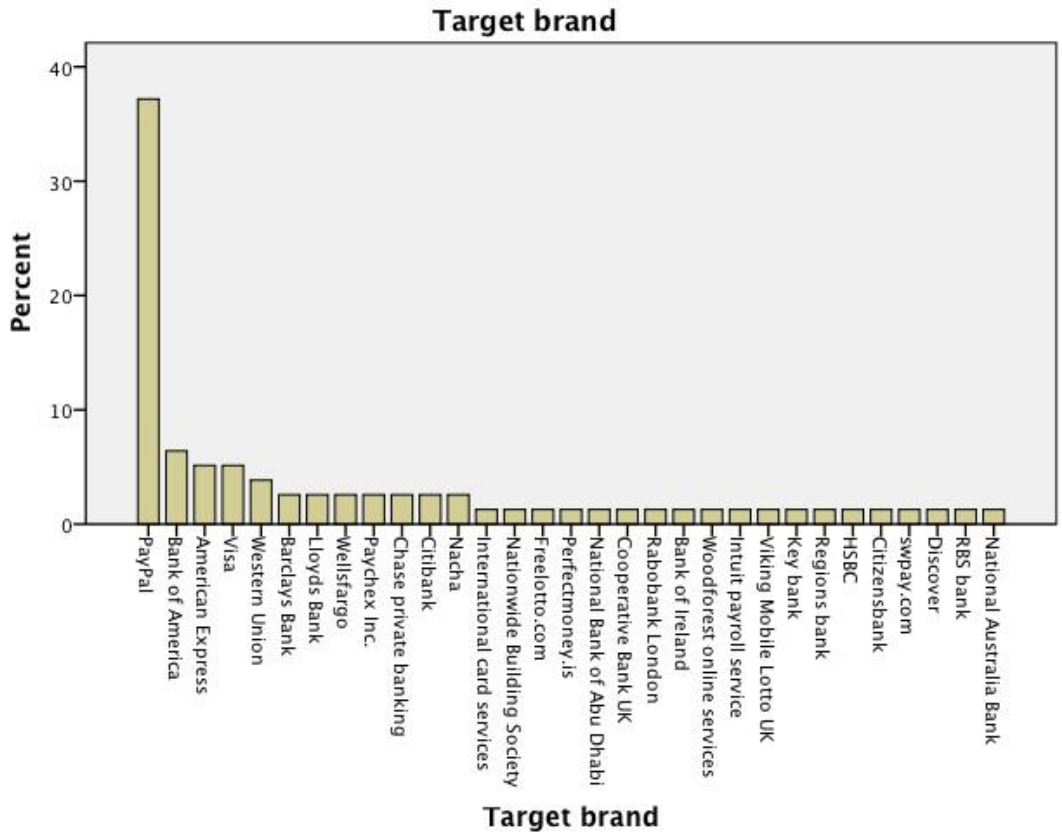


Figure 20: Detailed of financial sectors

B

BIBLIOGRAPHY

[1] Christopher Abad. "The economy of phishing: A survey of the operations of the phishing market." *First Monday* 10.9 (2005) (cit. on p. 10).

[2] PA Barraclough et al. "Intelligent phishing detection and protection scheme for online transactions." *Expert Systems with Applications* 40.11 (2013), pp. 4697–4706 (cit. on p. 23).

[3] 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*. ACM. 2011, pp. 3469–3478 (cit. on pp. 1, 3, 4).

[5] Madhusudhanan Chandrasekaran, Krishnan Narayanan, and Shambhu Upadhyaya. "Phishing email detection based on structural properties." In: *NYS Cyber Security Conference*. 2006, pp. 1–7 (cit. on pp. 2, 9, 28).

[6] Sidharth Chhabra et al. "Phi. sh/$ oCiaL: the phishing landscape through short URLs." In: *Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-Abuse and Spam Conference*. ACM. 2011, pp. 92–101 (cit. on p. 23).

[7] Pern Hui Chia and Svein Johan Knapskog. "Re-evaluating the wisdom of crowds in assessing web security." In: *Financial Cryptography and Data Security*. Springer, 2012, pp. 299–314 (cit. on p. 22).

[8] Robert B Cialdini. "the SCIENCE of Persuasion." *Scientific American* 284.2 (2001), p. 76 (cit. on pp. 2, 3, 5–7, 15, 16, 21, 32, 33, 39, 46, 47).

[9] Richard Clayton et al. "A Chat at the Old Phishin'Hole." In: *Financial Cryptography and Data Security*. Springer, 2005, pp. 88–88 (cit. on p. 11).

[10] Rachna Dhamija, J Doug Tygar, and Marti Hearst. "Why phishing works." In: *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM. 2006, pp. 581–590 (cit. on pp. 1, 11, 30).

[13] Ronald C Dodge Jr, Curtis Carver, and Aaron J Ferguson. "Phishing for user security awareness." *Computers & Security* 26.1 (2007), pp. 73–80 (cit. on p. 29).

[14] Douglas P Dotterweich and Kimberly S Collins. "The practicality of Super Bowl advertising for new products and companies." *Journal of Promotion Management* 11.4 (2006), pp. 19–31 (cit. on p. 37).

[15] Christine E Drake, Jonathan J Oliver, and Eugene J Koontz. "Anatomy of a Phishing Email." In: *First Conference on Email and Anti-Spam*. 2004 (cit. on p. 28).

[16] Shaun Egan and Barry Irwin. "An evaluation of lightweight classification methods for identifying malicious URLs." In: *Information Security South Africa (ISSA), 2011*. IEEE. 2011, pp. 1–6 (cit. on pp. 23, 26).

[18] Edwin Donald Frauenstein and Rossouw von Solms. "An Enterprise Anti-phishing Framework." In: *Information Assurance and Security Education and Training*. Springer, 2013, pp. 196–203 (cit. on pp. 14–16, 29, 30).

[19] A.A. Ghorbani, W. Lu, and M. Tavallaee. "Network attacks." *Advances in Information Security* 47 (2010), pp. 1–25 (cit. on pp. 3, 4).

[21] Gaurav Gupta and Josef Pieprzyk. "Socio-technological phishing prevention." *Information Security Technical Report* 16.2 (2011), pp. 67–73 (cit. on p. 23).

[22] Cormac Herley and Dinei Florêncio. "A profitless endeavor: phishing as tragedy of the commons." In: *Proceedings of the 2008 workshop on New security paradigms*. ACM. 2009, pp. 59–70 (cit. on p. 12).

[23] Amir Herzberg and Ronen Margulies. "Forcing Johnny to login safely." *Journal of Computer Security* 21.3 (2013), pp. 393–424 (cit. on pp. 3, 5).

[24] Jason Hong. "The state of phishing attacks." *Communications of the ACM* 55.1 (2012), pp. 74–81 (cit. on pp. 12, 14).

[25] Huajun Huang, Liang Qian, and Yaojun Wang. "A SVM-based technique to detect phishing URLs." *Information Technology Journal* 11.7 (2012), pp. 921–925 (cit. on p. 23).

[26] Tom N Jagatic et al. "Social phishing." *Communications of the ACM* 50.10 (2007), pp. 94–100 (cit. on pp. 1, 9).

[27] Markus Jakobsson. "Modeling and Preventing Phishing Attacks." In: *Financial Cryptography and Data Security*. Springer, 2005, pp. 89–89 (cit. on p. 11).

[30] K Jansson and Rossouw von Solms. "Phishing for phishing awareness." *Behaviour & Information Technology* 32.6 (2013), pp. 584–593 (cit. on p. 29).

[31] Kim Kaivanto. "The Effect of Decentralized Behavioral Decision Making on System-Level Risk." *Risk Analysis, Forthcoming* (2014) (cit. on p. 3).

[32] Masatoshi Kawakami, Hiroshi Yasuda, and Ryoichi Sasaki. "Development of an E-Learning Content-Making System for Information Security (ELSEC) and Its Application to Anti-Phishing Education." In: *International Conference on e-Education, e-Business, e-Management, and e-Learning, 2010. IC4E'10.* IEEE. 2010, pp. 7–11 (cit. on pp. 3, 5).

[33] Daejoong Kim and Jang Hyun Kim. "Understanding persuasive elements in phishing e-mails: A categorical content and semantic network analysis." *Online Information Review* 37.6 (2013), pp. 835–850 (cit. on pp. 3, 6, 7).

[34] Iacovos Kirlappos and Martina Angela Sasse. "Security Education against Phishing: A Modest Proposal for a Major Rethink." *IEEE Security and Privacy Magazine* 10.2 (2012), pp. 24–32 (cit. on p. 30).

[35] Katharina Krombholz et al. "Social engineering attacks on the knowledge worker." In: *Proceedings of the 6th International Conference on Security of Information and Networks*. ACM. 2013, pp. 28–35 (cit. on pp. 3, 4).

[36] Ponnurangam Kumaraguru et al. "Lessons from a real world evaluation of anti-phishing training." In: *eCrime Researchers Summit, 2008*. IEEE. 2008, pp. 1–12 (cit. on p. 29).

[37] Ponnurangam Kumaraguru et al. "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*. ACM. 2007, pp. 905–914 (cit. on p. 2).

[38] Ponnurangam Kumaraguru et al. "School of phish: a real-world evaluation of anti-phishing training." In: *Proceedings of the 5th Symposium on Usable Privacy and Security*. ACM. 2009, p. 3 (cit. on pp. 30–32).

[40] Elmer EH Lastdrager. "Achieving a consensual definition of phishing based on a systematic review of the literature." *Crime Science* 3.1 (2014), pp. 1–10 (cit. on pp. 11, 12).

[41] Anh Le, Athina Markopoulou, and Michalis Faloutsos. "Phishdef: Url names say it all." In: *INFOCOM, 2011 Proceedings IEEE*. IEEE. 2011, pp. 191–195 (cit. on pp. 25, 26).

[43] Gang Liu, Bite Qiu, and Liu Wenyin. "Automatic detection of phishing target from phishing webpage." In: *Proceedings of International Conference on Pattern Recognition*. IEEE. 2010, pp. 4153–4156 (cit. on p. 22).

[45] Justin Ma et al. "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*. ACM. 2009, pp. 1245–1254 (cit. on pp. 25–27).

[46] Justin Ma et al. "Identifying suspicious URLs: an application of large-scale online learning." In: *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM. 2009, pp. 681–688 (cit. on p. 27).

[47] Liping Ma et al. "Detecting phishing emails using hybrid features." In: *Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing, 2009. UIC-ATC'09.* IEEE. 2009, pp. 493–497 (cit. on p. 29).

[48] Steve Mansfield-Devine. "Interview: Joe Ferrara–fighting phishing." *Computer Fraud & Security* 2013.7 (2013), pp. 17–20 (cit. on p. 30).

[51] Tyler Moore and Richard Clayton. "An Empirical Analysis of the Current State of Phishing Attack and Defence." In: *Proceedings of the 2007 Workshop on the Economics of Information Security*. 2007 (cit. on p. 20).

[52] Tyler Moore and Richard Clayton. "Evaluating the wisdom of crowds in assessing phishing websites." In: *Financial Cryptography and Data Security*. Springer, 2008, pp. 16–30 (cit. on pp. 22–24).

[53] Giovane CM Moura and Aiko Pras. "Scalable Detection and Isolation of Phishing." In: *Scalability of Networks and Services*. Springer, 2009, pp. 195–198 (cit. on p. 12).

[54] Philip J Nero et al. "Phishing: Crime that pays." In: *eCrime Researchers Summit (eCrime), 2011*. IEEE. 2011, pp. 1–10 (cit. on pp. 14, 18).

[58] Parth Parmar and Kalpesh Patel. "Comparison of Phishing Detection Techniques." In: *International Journal of Engineering Research and Technology*. Vol. 3. 3 (March-2014). ESRSA Publications. 2014 (cit. on p. 21).

[59] Bryan Parno, Cynthia Kuo, and Adrian Perrig. *Phoolproof phishing prevention*. Springer, 2006 (cit. on p. 11).

[60] James W Pennebaker and Deborah Yates Sanders. "American graffiti: Effects of authority and reactance arousal." *Personality and Social Psychology Bulletin* 2.3 (1976), pp. 264–267 (cit. on p. 36).

[62] Swapan Purkait. "Phishing counter measures and their effectiveness–literature review." *Information Management & Computer Security* 20.5 (2012), pp. 382–420 (cit. on p. 21).

[64] Stuart E Schechter et al. "The emperor's new security indicators." In: *IEEE Symposium on Security and Privacy*. IEEE. 2007, pp. 51–65 (cit. on p. 30).

[65] Kunal Sharma. "An Anatomy of Phishing Messages as Deceiving Persuasion: A Categorical Content and Semantic Network Study." *Edp Audit, Control, and Security* 42.6 (2010), pp. 1–19 (cit. on p. 3).

[66] Frank Stajano and Paul Wilson. "Understanding scam victims: seven principles for systems security." *Communications of the ACM* 54.3 (2011), pp. 70–75 (cit. on p. 33).

[67] Henri Tajfel and John C Turner. "The Social Identity Theory of Intergroup Behavior." *Psychology of Intergroup Relations* (1986) (cit. on p. 37).

[71] Arun Vishwanath et al. "Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model." *Decision Support Systems* 51.3 (2011), pp. 576–586 (cit. on pp. 3, 5).

[73] Liu Wenyin et al. "Discovering phishing target based on semantic link network." *Future Generation Computer Systems* 26.3 (2010), pp. 381–388 (cit. on p. 23).

[75] 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: *Proceedings of SPIE: The International Society for Optical Engineering*. International Society for Optics and Photonics. 2012 (cit. on p. 22).

[76] 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 (2008), pp. 662–674 (cit. on pp. 2, 3, 5–7, 9, 32, 33, 37, 46, 47).

[77] Ryan T Wright et al. "Research Note-Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance." *Information Systems Research* 25.2 (2014), pp. 385–400 (cit. on pp. 3, 5, 7).

[78] Guang Xiang et al. "CANTINA+: a feature-rich machine learning framework for detecting phishing web sites." *ACM Transactions on Information and System Security (TISSEC)* 14.2 (2011), p. 21 (cit. on pp. 26–28).

[79] 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*. ACM. 2013, pp. 341–346 (cit. on p. 23).

[80] 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*. ACM. 2007, pp. 639–648 (cit. on pp. 2, 28).

[81] Yue Zhang et al. "Phinding phish: Evaluating anti-phishing tools." In: *Proceedings of the 14th Annual Network and Distributed System Security Symposium (NDSS 2007)*. 2007 (cit. on p. 2).

NON PEER-REVIEWED BIBLIOGRAPHY

[4] Ashley Carman. *Phishing scam targets Michigan public schools*. [Online; accessed 13-May-2014]. URL: http://www.scmagazine.com/phishing-scam-targets-michigan-public-schools/article/343177/ (cit. on p. 19).

[11] Oxford Dictionaries. *Phishing*. Web Page. URL: http://www.oxforddictionaries.com/definition/english/phishing (cit. on pp. 9, 11).

[12] Collins English Dictionary. *Phishing*. Web Page. URL: http://www.collinsdictionary.com/dictionary/american/phishing (cit. on p. 11).

[17] Aaron Emigh. "Online identity theft: Phishing technology, chokepoints and countermeasures." *ITTC Report on Online Identity Theft Technology and Countermeasures* 3 (2005) (cit. on pp. 14, 17, 18, 29, 30).

[20] Adam Greenberg. *Medical staffers fall for phishing emails, data on 8,300 compromised*. [Online; accessed 13-May-2014]. URL: http://www.scmagazine.com/medical-staffers-fall-for-phishing-emails-data-on-8300-compromised/article/340590/ (cit. on p. 19).

[28] Markus Jakobsson and Steven Myers. *Phishing and countermeasures: understanding the increasing problem of electronic identity theft*. John Wiley & Sons, 2006 (cit. on pp. 1, 9–12, 18, 20).

[29] Lance James. *Phishing exposed*. Syngress, 2005 (cit. on pp. 1, 10, 11, 29).

[39] Willy Lai. "Fitting Power Law Distributions to Data" () (cit. on p. 23).

[42] Avivah Litan. "Phishing victims likely will suffer identity theft fraud." *Gartner Research Note (May 14, 2004)* (2004) (cit. on p. 12).

[44] Haotian Liu, Xiang Pan, and Zhengyang Qu. "Learning based Malicious Web Sites Detection using Suspicious URLs." *Department of Electrical Engineering and Computer Science, Northwestern University* (2012) (cit. on pp. 26–28).

[49] Tom McCall. *Gartner survey shows phishing attacks escalated in 2007*. 2007 (cit. on p. 12).

[50] Kevin D Mitnick and William L Simon. *The art of deception: Controlling the human element of security*. John Wiley & Sons, 2001 (cit. on pp. 2, 32).

[55] National Plant Diagnostic Network. *Types of Social Engineering*. [Online; accessed 16-July-2014]. URL: http://www.npdn.org/social_engineering_types (cit. on p. 36).

[56] Organización Internacional de Normalización. *ISO/IEC 27002: Information Technology, Security Techniques, Code of Practice for Information Security Management*. ISO/IEC, 2005 (cit. on p. 29).

[57] OpenDNS. *Phishtank: Out of the Net, into the Tank*. [Online; accessed 13-May-2014]. URL: http://www.phishtank.com/faq.phpk (cit. on p. 22).

[61] Phishing.org. *History of Phishing*. Web Page. URL: http://www.phishing.org/history-of-phishing/ (cit. on p. 10).

[63] Teri Robinson. *Phishing scam aimed at Google Docs, Drive users*. [Online; accessed 13-May-2014]. URL: http://www.scmagazine.com/phishing-scam-aimed-at-google-docs-drive-users/article/338369/ (cit. on p. 19).

[68] Gregg Tally, Roshan Thomas, and Tom Van Vleck. "Anti-Phishing: Best Practices for Institutions and Consumers." *McAfee Research, Mar* (2004) (cit. on pp. 11, 14, 17, 18).

[69] Wombat security technology. *PhishGuru: Assess and Motivate Your Employees using Simulated Phishing Attacks*. [Online; accessed 23-May-2014]. URL: http://www.wombatsecurity.com/phishguru (cit. on p. 31).

[70] Inspired Telemarketing. *5 tips for getting past receptionists!* [Online; accessed 16-July-2014]. 2013. URL: http://inspiredtelemarketing.wordpress.com/2013/09/13/5-tips-for-getting-past-receptionists/ (cit. on p. 36).

[72] Merriam Webster. *Phishing*. Web Page. URL: http://www.merriam-webster.com/dictionary/phishing (cit. on p. 11).

[74] Rebecca Wetzel. "Tackling phishing." *Business Communications Review* 35.2 (2005), pp. 46–49 (cit. on pp. 14, 16, 17).

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 October 30, 2014 (`Nurul Akbar` version 1).