

General understanding towards persuasive elements in phishing email - Research Topic

August 28, 2014

Nurul Akbar

s1192523

University of Twente

Motivation and background

With the enormous growing trends of information technology in modern generation, the evolution of digital era has become more mature in the sense of effectiveness and easiness for societies. Trusted entities such as financial institutions may offer their products and services to the public through the Internet. Consequently, human society has been enthusiastically eager to utilize technology means such as emails, websites, online payment system, social networks to achieve their tasks efficiently, affordable and more relevant. However, the advancement in information and communication technology has been a double-edged sword. It becomes easier to get personal information about someone in the cyber world. 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

One of the well known cyber crimes is called phishing. Phishing attacks aim to gain a financial benefit by masquerading legitimate institutions [21]. Moreover, phishing techniques may associated with fake emails and fake websites, however, the essence of this malicious behavior is the art of social engineering and deception [21][77][10][22][78] i.e. how to trick the potential victims into disclosing their sensitive information. To illustrate one specific example; a phisher impersonates a well-known bank institutions and sends fake email announcing there is system upgrade to its customers, so the phisher would ask the customers to verify their usernames, passwords and credit card numbers in order to complete the system upgrade. In worst case, a clueless bank customer would have no idea and fail to ascertain that legitimate bank institutions would never ask for such things. Consequently, an unsuspecting victim might took the bait

into divulging their sensitive information to the perpetrators. More concrete mechanism to carry out phishing attacks is that when the phishers send a large number of phishing emails which include malicious link or URL which redirect recipient to a fake website. A phishing email might be resembling a legitimate email so that unsuspecting victim might think that it is a genuine. By looking like a genuine email, it would help to circumvent the filtering system [79]. One of phishing email methods is providing a malicious URL within its content, however, filtering systems of phishing email may have access to blacklist database of URLs that are currently being pushed, such as crowd sourced Phishtank. So the phishers might diverge the URL to bypass the matches.

To make phishing email efficient, its context might require potential victim to urgently act upon it, for example an email informs about suspended account in a banking website. Phishing attacks are carried out using fake emails, fake websites and often phishers may exploit the end user's web browser to bypass warnings and URL information. Several countermeasures have been studied to detect phishing emails, such as the technique of looking for bit string that are previously determined as spams and phishing emails [53]. The common characteristics of fake email sent by the phishers would be misleading hyperlinks and misleading header information [60, 61]. Furthermore, one of the non technical approach to defense against cyber fraud such as security awareness education to the public to ignore links within an email, even though the source of the email appear to be legitimate [51]. However, current countermeasures within a phishing email so far do not consider psychoanalytic approach such as persuasion techniques, which it might be important for user decision. Thus, characterization and the relevancy of phishing email properties are needed, as well as the relevancy of persuasion techniques with these structural properties.

We begin our research topic by understanding what is phishing itself, how phishing emerged in term of history.

What is phishing?

As we already aware, the Internet has brought convenience to many people for exchanging information, it also provides opportunities to carry malicious behavior such as online fraud on massive scale with a little cost to the attackers. The attackers can manipulate the Internet users instead of computer system (hardware or software) that significantly increase the barriers of technological crime impact. Such human centered attacks could be done by social engineering. According to Jakobsson, et al. phishing is a form of social engineering that aims to retrieve credential from online users by mimicking trustworthy and legitimate institutions [21]. Phishing has a similar basic principle as 'fishing' in the physical world. Instead of fish, online users are lured by authentic looking communication and hooked by authentic looking websites. Not only that, online users also may be lured by respond to a phishing email, either replying or clicking an obfuscated link within its content. There are diverse definitions of phishing in our literature reviews, therefore, we would like to discuss about its universal definition in later

section. However, one of phishing definitions according to Oxford dictionary:

“A fraudulent practice of sending emails purporting to be from reputable companies in order to induce individuals to reveal personal information, such as passwords and credit card numbers, online” [44].

Several studies suggest that phishing is form of online criminal activity by using social engineering techniques [21][57][78][2]. An individual or a group who uses this technique is called *Phisher(s)*. After successfully gaining a sensitive information from the victim, phishers use this information to access victim's financial accounts or committing credit card frauds. The technique of phishing may vary, but the most common technique of phishing attacks done by using fraudulent emails and websites [22]. A fraudulent website is designed in such a way that it may be identical to its legitimate target. While it may be true, phishing website also could be completely different with its target as there is no level of identicalness. However, Preliminary analysis on what changed or added in phishing website would be conducted in the later section.

The History

The first time the term "phishing" was published by the AOL UseNet Newsgroup on January 2, 1996 and was started to expand in 2004 [47]. Since then, we considered phishing development in cyberspace has been flourishing by phishers to make profit. Total losses due to phishing in 2004 reached more than U.S. \$ 2 billion, it was involving more than 15,000 sites that become victims [14]. We will try to discuss about direct and indirect cost at present days in the later section. Evidently, Jakobsson, et al. [21] mentioned that in the early years of 90's (according to [47] it was around 1995) many hackers would create bogus AOL user accounts with automatically generated fraudulent credit card information. Their intention to give this fake credit card information was to simply pass the validity tests performed by AOL. By the time the tests were passed, AOL was thinking that these accounts were legitimate and resulted to activate them. Consequently, these hackers could freely access AOL resources until AOL tried to actually bill the credit card. AOL realized that these accounts were using invalid billing information, thus deactivated the account.

While creating false AOL user accounts with fake credit card information was not exactly phishing attacks, but AOL's effort to counter against the attacks was leading to development of phishing. This countermeasure includes directly verifying the legitimacy of credit card information and the associated billing identity, forced hackers to pursue alternative way [21]. Hackers were masquerading as AOL's employees asking to other users for credit card information through AOL instant messenger and email system [47]. Jakobsson et al. suggest that phishing attacks were originating from this incident [21]. Since such attack has not been done before, many of users have been victimized by then. Eventually, AOL enforced warning system to the most of its customers

to be vigilant when it comes to sensitive information [47]. At the present day, phishing attacks might not only being motivated by financial gain but also political reason, and they have been emerging not only aim to AOL users, but also any online users. Consequently, large number of legitimate institutions such as PayPal and eBay are being spoofed.

Universal definition

Before we begin to understand deeper about how and why phishing attack works, we will briefly explore common phishing definition. Currently, there is no consensus definition, since almost in every research papers, academic textbook or journals has its own definition of phishing [21, 22, 52, 7, 46, 20, 10]. Phishing is also constantly evolving, so it might be very challenging to define its universal terminology. There is not so much study that specifically addresses the standard of phishing definition. However, one research conducted by Lastdrager [28] addressed to achieve consensual definition of phishing. Before we comply with one consensual phishing terminology, we will take a look at various phishing definitions from other sources:

“Phishing is the act of sending a forged e-mail (using a bulk mailer) to a recipient, falsely mimicking a legitimate establishment in an attempt to scam the recipient into divulging private information such as credit card numbers or bank account passwords” [22]

“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 web sites that impersonate both legitimate E-mail and web sites” [52]

“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” [7]

“In phishing, an automated form of social engineering, criminals use the internet to fraudulently extract sensitive information from businesses and individuals, often by impersonating legitimate web sites” [46]

It is noteworthy that the definition described by James, et al, Tally, et al, and Clayton, et al. [22, 52, 7] specifies that the phishers only use email as a communication channel to trick potential victims. While it might be true because using email would be cost effective method, but we believe that phishing is not only characterized by one particular technological mean, as phishers can also use any other electronic communication to trick potential victims (i.e private message on online social network). This definition is also similar to dictionary libraries [44, 8, 37] that mention email as a medium communication between phishers and users.

We believe that standard definition of phishing should be applicable in most of phishing concept that are presently defined. Consequently, the high level of abstraction and is required to build common definition on phishing. We have convinced that the formal definition of phishing should not focus on the technology that is being used but rather on the technique how the deception is being conducted, the method of an “act” if you will. Therefore, We follow the definition of phishing by Lastdrager [28] which stated:

“Phishing is a scalable act of deception whereby impersonation is used to obtain information from a target”

According to Lastdrager [28], to achieve this universal definition, a systematic review of literature up to August 2013 was conducted along with manual peer review, which resulted in 113 distinct definitions to be analyzed. We thereby agree with Lastdrager [28] that this definition addresses all the essential elements of phishing and we will adopt it as universally accepted terminology throughout our research.

The cost of phishing attacks

It is a challenging task to find a real cost from phishing attacks in term of money or direct cost. This due to financial damage for bank is only known by banks and most institutions do not share this information with the public. Evidently, Jakobsson et al. argue that phishing economy is consistent with black market economy and does not advertise its successes [21]. On this section, a brief explanation of direct and indirect cost on phishing attack will be illustrated based on literature reviews.

According to Jakobsson et al., direct cost is depicted by the value of money or goods that are directly stolen through phishing attack [21]. While indirect cost is the costs that do not represent the money nor goods that are actually stolen, but it is the costs has to be paid by the people who handle these attacks [21] (i.e. time, money and resources spent to reset people password).

As we mentioned earlier, the difficulty of assessing the damage on phishing attacks is caused by banks and institutions that keep this information themselves and the unwillingness of many users to share to acknowledge that they have been victimized by phishing attacks. This happens because of fear of humiliation, financial loses, or legal liability [21]. Evidently, studies estimate the damage ranging from \$61 million [17] to \$3 billion per year [67] of direct losses to victims in the US only [18][68]. In addition, the Gartner Group claimed to estimate of \$1.2 billion direct losses of phishing attack to US banks and credit card companies for the year 2004 [30]. By the 2007, it escalated to more than \$3 billion loss [36]. The estimation also performed by TRUSTe and Ponemon Institute that stated the cost of phishing attack was up to \$500 millions losses in the US for the same year ¹. Recently, RSA FraudAction gives monthly reports

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

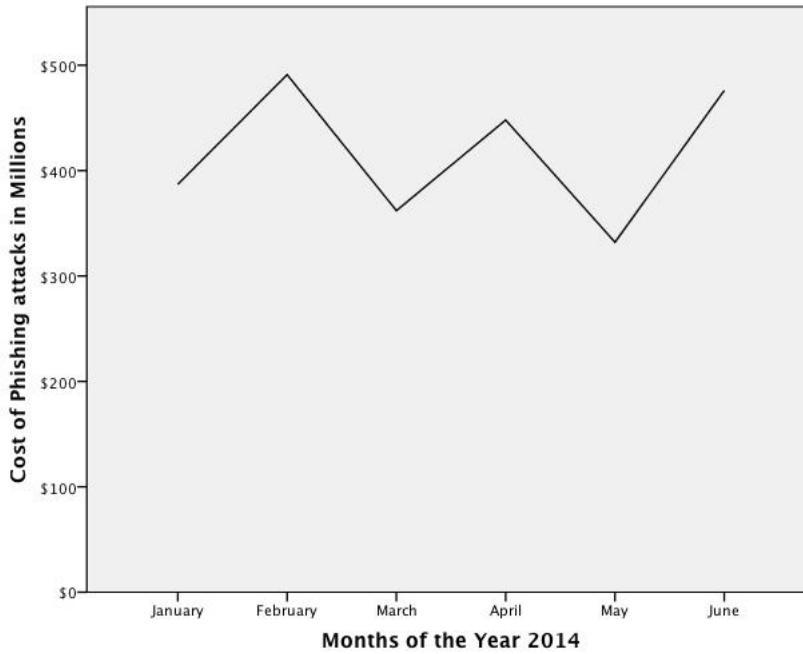


Figure 1: Global phishing cost from January 2014 until June 2014 [69]

of how much the global phishing cost and we compiled them together to make a line graph in Figure 1[69]. However, they do not specify whether this damage is direct cost or indirect cost. In their study, we can see that there are fluctuation of losses in term of money. Furthermore, this total cost is indeed in accordance with the total cost ranging from \$61 million to \$3billion per year [17, 67, 18, 68].

Human factor

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

Kevin Mitnick, who was obtaining millions of dollars by performing social engineering technique, is plausibly the best known person who had used social engineering technique to carry out his attacks. His book that titled “The art of deception: Controlling the Human Element of Security” [38] has defined social engineering as follows:

“Using influence and persuasion to deceive people by convincing them that the attacker is someone he is not, or by manipulation.

As a result, the social engineer is able to take advantage of people to obtain information, or to persuade them to perform an action item, with or without the use of technology.”

From his definition we can learn that people are the main target of the attack, specifies some of the important tools used by the attackers, such as influence and persuasion.

Cialdini suggests that there are six basic principles of persuasion [6], that is, the technique of making people grant to one’s request. These principles include; *reciprocation, consistency, social proof, likeability, authority and scarcity*. Reciprocation is the norm that obligates individuals to repay in kind what they have received, return the favor or adjustment to smaller request [6]. Consistency is a public commitment where people become psychologically become vested in a decision they have made [57][6]. Social proof is when people model the behavior of their peer group, role models, important others or because it is generally "fashionable" [57]. Stajano, et al. suggest people will let their guard down when everybody around them appears to share the same risk [80]. Likeability is when people trust and comply with requests from others who they find attractive or having credibility [57, 6]. While it is our human nature not to question 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 [6, 57]. Stajano, et al suggest that scarcity is related to time principle, that is, when we are under time pressure to make important choice, we tend to have less reasoning to make decision [80]. We will use these principles as our foundation in synthesizing phishing email corpus with human factor.

Human as the “weakest link” in computer security has been exists and exploited for ages. And yet, security designers blame on users and whine “the system I designed would be secure, if only users were less gullible” [80]. Stajano, et al. stated that “a wise security designer would seek a robust solution which acknowledge the existence of these vulnerabilities as unavoidable consequence of human nature and actively build countermeasures that prevent this exploitation” [80]. With this in mind, the exploration of persuasion principles is congruent with our research goal. Cialdini’s six persuasion principles will be the foundation in our research.

Modus operandi

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

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

Table 1: Compilation of phishing phases

J. Hong [18]
<ol style="list-style-type: none"> 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. [15]
<ol style="list-style-type: none"> 1. Planning 2. Email Design 3. Fabricated story 4. Threatening tone/Consequences 5. Spoofed website
Wetzel [55]
<ol style="list-style-type: none"> 1. Planning 2. Setup 3. Attack 4. Collection 5. Fraud 6. Post-attack
Tally, et al. [52]
<ol style="list-style-type: none"> 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 [13]

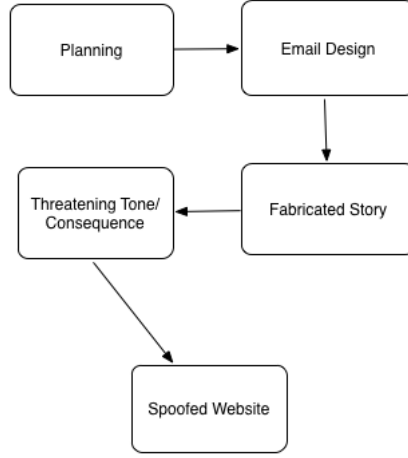


Figure 2: Phishing processes based on Frauenstein[15]

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. [42]

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 [18] argued that there are three major phases. while Frauenstein, et al. [15] suggested that there are five main processes are used to perform phishing attacks based on the perspective of the attacker.

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



Figure 3: Example of fake ING logo in phishing email

look official to the victim. The author called this process as *Email Design*. Figure 3 illustrates the example of fake ING bank logo in a phishing email to create “legitimate” feel². On the third process, the phisher *fabricates* a story to make potential victim think that email is important. To achieve users attention, phisher might build up a story about system upgrade, account hijacked or security enhancement so that the victim would feel obliged to be informed. Evidently, this technique corresponds with Cialdini [6] that suggests there are six principles to persuade people to comply with a request. On the fourth process, a phisher usually include *threatening tone* or explain the urgency and consequences if the potential victim chooses not to take action desired by the phisher (for example; account removal, account blocked, etc.). Consequently, users may fear of their account being deleted. The last process involved with fraudulent website that has been created by the phisher. Users may falsely believe to the message given in the email and may click a URL that is embedded in the email. Subsequently, the URL would redirect users to a *spoofed website* which may prompt users’ sensitive information. Furthermore, the website might be created to be as similar as possible to the target institution’s website, so that potential victim may still believe that it is authentic. We will explain more on Cialdini’s six basic tendencies of human behavior in generating positive response to persuasion [6] in a later section.

Considering that phishing attack is a process, Wetzel [55] suggested a taxonomy to make sense of the complex nature of the problem by mapping out a common attacks lifecycle, and a possible set of activities attackers engage in within each phase. The taxonomy is illustrated in Figure 4. We speculated that Wetzel’s taxonomy is not analogous with Frauenstein’s main phishing processes [15]. The difference is that Frauenstein et al. only focus in the design of the attack while Wetzel has added several phases like *Collection*, *Fraud* and *Post-attack*, therefore, Wetzel taxonomy is more holistic in term of phishing.

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

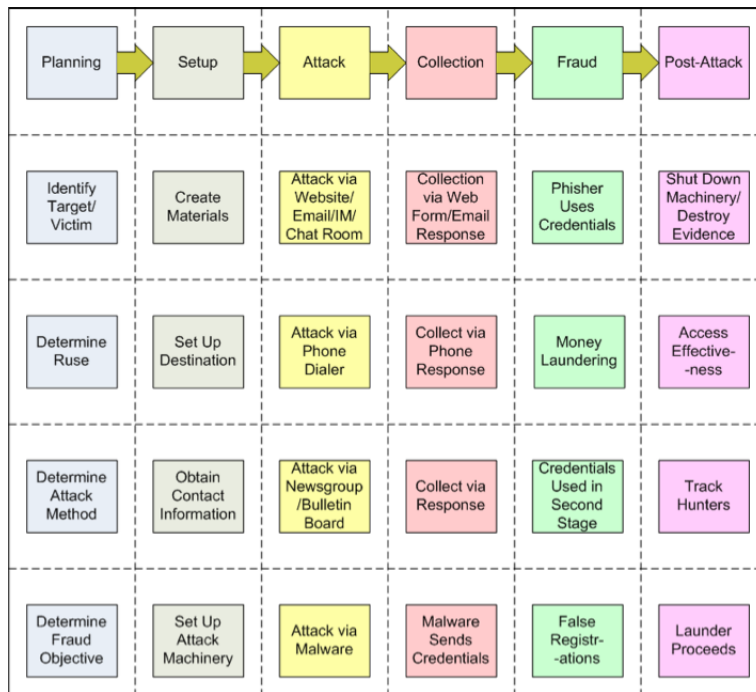


Figure 4: Phishing attack taxonomy and lifecycle[55]

We have listed Wetzel’s taxonomy in Table 1, we explain more of the taxonomy as follows:

1. *Planning*: Preparation carried out by the phisher before continue to the next phase. Example activities include identifying targets and victims, determine the method of the attack, etc.
2. *Setup*: After the target, victim and the method are known, the phisher would craft a platform where the victim’s information could be transmitted and stored, for example: fraudulent website/email.
3. *Attack*: Phisher distributes their fraudulent platform so that it can be delivered to the potential victims with fabricated stories.
4. *Collection*: Phisher collects valuable information via response from the victims
5. *Fraud*: Phisher abuses victim’s information by impersonates the identity of the victim to the target. For example, A has gained B’s personal information to access C so that A can pose as B to access C.
6. *Post-attack*: After the phisher gained profit from the attack and abuse phases, a phisher would not want to be noticed or detected by authority. Thus, phisher might need to destroy evidence of the activities that he/she previously were executed.

Tally, et al. suggest that there are several phases involved in phishing attack based on the attacker’s point of view [52]:

- “1. The attacker obtains E-mail addresses for the intended victims. These could be guessed or obtained from a variety of sources” this fits with planning phase.
- “2. The attacker generates an E-mail that appears legitimate and requests the recipient to perform some action” - fits with design phase.
- “3. The attacker sends the E-mail to the intended victims in a way that appears legitimate and obscures the true source” - fits with delivery and attack phase.
- “4. Depending on the content of the E-mail, the recipient opens a malicious attachment, completes a form, or visits a web site” - fits with attack phase.
- “5. The attacker harvests the victim’s sensitive information and may exploit it in the future” - fits with fraud phase.

Additionally, the phases described by Tally, et al. [52] are comparable with the information flow explained by Emigh[13] represented in Figure 5.

Emigh [13] suggests the information flow of phishing attacks has seven phases. These phases is explained as follow:

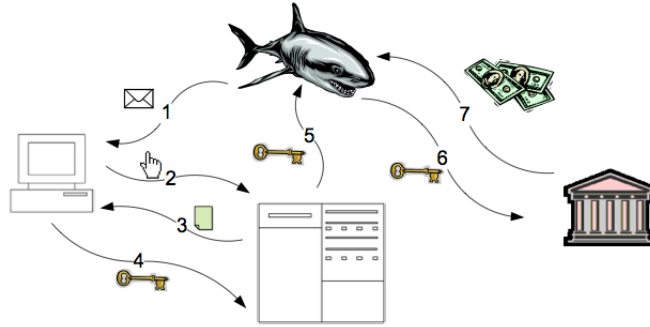


Figure 5: Flow of information in phishing attack [13]

- “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.” [13]

Phishing attack steps performed by the phisher are also addressed by Nero, et al [42]. In their study, a successful phishing attack involves several phases:

- “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” [42]

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. [21]. These key components are include *the lure*,

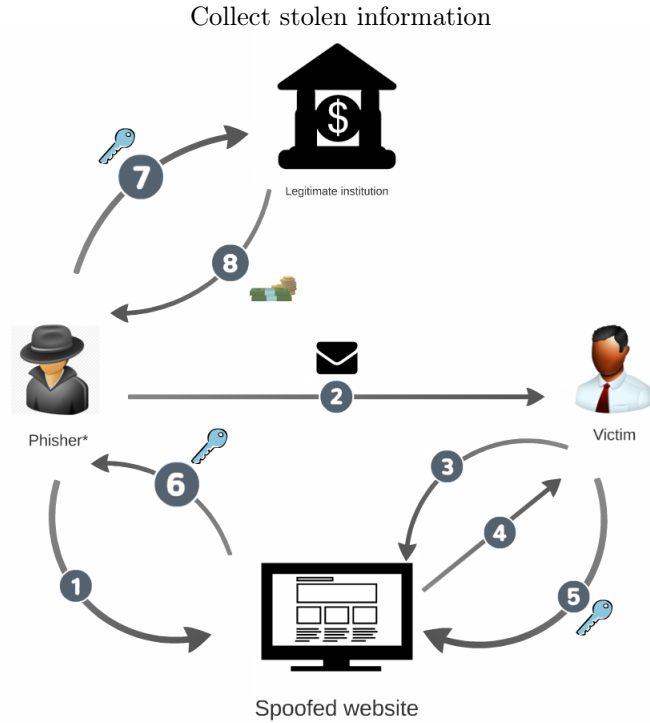


Figure 6: Information flow phishing attack

the hook and *the catch*. As we designed in Figure 6, we synthesized these three components with our phases based on the attacker point of view as follows:

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

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

Types of phishing

In January 2014, 8300 patients data are being compromised in medical company in the US [70]. The data includes names, addresses, date of birth and phone numbers were being stolen. Other than demographic information, clinical information associated with this data was also stolen, including social security numbers. In the April 2014, phishers have successfully stolen US\$163,000 from US public school based on Michigan [71]. It has been said that the email prompted to transfer money is coming from the finance director of the school. In March 2014, Symantec has discovered phishing attack aimed at Google drive users [72]. The attack was carried firstly with incoming email asking for opening document hosted at Google docs. Users that have clicked on the link are taken to fraudulent Google login page prompted Google users credentials. Interestingly, the URL seems very convincing because it hosted on Google secure servers. We hypothesized that even more phishing incidents on financial area as well, but sometimes the news is kept hidden due to creditability reason.

One may ask, what type of phishing are these? What type of phishing commonly used nowadays? Evidently, based on the cost of phishing attacks in , the threat of phishing attacks is still alarming and might be evolving in the future with more sophisticated technique of attacks. For this reason, it might be useful to provide a brief insight on popular variants of phishing that currently exist. We will briefly explain the types of phishing which are the most relevant to our research based on Jakobsson, et al. [21]. These types of phishing are strongly related to the phishing definition that we used, considering phishing is based on the act of deception by the phishers.

Deceptive phishing

There are many variations based on deceptive phishing schemes. Typical scenario of deceptive phishing schemes is to send a large amount of illicit emails containing call to action asking recipients to click embedded links [21]. These variations include cousin domain attack. For example, legitimate PayPal website addressed as `www.paypal.com`, this cousin domain attacks confuse potential victims to believe that `www.paypal-security.com` is a subdivision of the legitimate website due to identical looking addresses. Similarly, homograph attacks create a confusion using similar characters to its addresses. For example, `www.paypal.com` and `www.paypal1.com`, both addresses look the same but on the second link, it uses “1” instead of “l”.

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



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

Another type of deceptive phishing scheme is rock-phish attacks. They held responsible for half a number of reported incidents worldwide in 2005 [39]. These attacks evade email filters by utilizes random text and GIF images which contain the actual message. Rock phish attacks also utilize a toolkit that capable to manage several fraudulent websites in a single domain. Sometimes, deceptive phishing schemes lead to installation of malware when users visit fraudulent website and we will describe malware based phishing scheme in the next section.

Malware-based phishing

Generally, malware based phishing refers to any type of phishing which involves installing malicious piece of software onto users' personal computer [21]. 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.

Bad neighborhoods on phishing

In the physical world, there are parts of certain area that have higher crime rates than others (eg. Bronx in the US) which called hotspots. Evidently, it is statistically more likely that a crime will occur compared to other locations [41]. To better illustrate this analogy, the police department in Belgium [73] has put up statistical information regarding crime rates in the country. Figure 7 shows an example in 2013; there were up to 5525 car theft incidents recorded and we can see there are certain areas that have higher probability that a car got stolen. For example, Antwerp had 415 incidents whereas Berlare had only 1 car theft incident [74]. The data is not necessarily based on cities, it can be based on the residential area within a city. This holds true that much higher crime rates in a concentrated location compared to any other locations. It is called bad neighborhood.

To reduce the crime rates in a bad neighborhood, it makes sense that the authority should put more enforcement in this location. Moreover, the citizen

should avoid this location as much as they can if they want to feel much safer. Evidently, Moura, et al. suggest that the existence of bad neighborhood phenomenon also occurs in the Internet world called “Internet bad neighborhoods” [41]. There are certain networks of Internet infrastructure that contain more malicious activities than other networks. For our preliminary analysis, we will adopt formal definition of Internet bad neighborhoods or Internet Badhoods by [41] which states:

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

Several studies have suggested that the source of the Internet Badhoods tend to be concentrated in certain portions of IP address space [49, 9, 3]. Moura, et al. suggest that Internet Badhoods do not always only based on network prefixes level (e.g. /24, /32, etc..) but it can be aggregated into several levels (ISPs, Countries) [41]. Moreover, Internet Badhoods may vary depending on which application exploited. While spam is most likely distributed all around the world, however, phishing Badhoods are most likely concentrated in developed countries (e.g. US) [41]. This suggests that phishing sites are required to have more reliable hosts in term of availability, while spams are not.

Current countermeasures

There are various types of phishing countermeasures that implemented in different levels. Purkait has conducted an extensive research in reviewing these countermeasures which are available up until 2012 and their effectiveness [48]. He suggests that there is a classification of phishing countermeasures in separate groups and according to Purkait [48], these groups are listed as follow:

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

In addition, Parmar, et al. suggests that phishing detection can be classified into two types; user training approach and software classification approach [45]. He illustrated a diagram and a table that summarizes phishing detection as

countermeasures in a broad view [45]. They also argued the advantages and disadvantages of each category [45]. However, as our research mainly focuses in synthesizing phishing email with cialdini’s six principles of persuasion [6], we will briefly discuss phishing countermeasures such as restriction list group (i.e. Phishtank), machine learning approach (web-based phishing), properties or features in a phishing email, and anti phishing training group (i.e PhishGuru).

Phishing detection

Phishtank

One of the common approaches to detect phishing attacks is the implementation of restriction list. As the name suggest, it prevents users to visit fraudulent websites. One of the efforts to achieve restriction list, is to derive phishing URLs from Phishtank. Phishtank is a blacklisting company specifically for phishing URLs and it is a free community web based where users can report, verify and track phishing URLs [75]. Phishtank stores phishing URLs in its database and is widely available for use by other companies for creating restriction list. Some of the big companies that are using Phishtank’s data includes; Yahoo Mail, McAfee, APWG, Web Of Trust, Kaspersky, Opera and Avira. In this section, we will discuss how the current literatures have to do with phish data provided by Phishtank. The first step to achieve the list of relevant literatures regarding phishtank is by keyword search in Scopus online library. By putting “Phishtank” as a keyword search, it results in 12 literatures. The next step, we read the all the abstracts and conclusions of the resulting keyword search and we decided 11 literatures that are relevant to our research. Lastly, Table 2 summarizes the papers selected and its relevancy with Phishtank

Table 2: Summary phishtank studies

Paper title	First author	Country	Relevancy with phishtank
Evaluating the wisdom of crowds in assessing phishing website [40]	Tyler Moore	United Kingdom	Examine the structure and outcomes of user participation in Phishtank. The authors find that Phishtank is dominated by the most active users, and that participation follows a power law distribution and this makes it particularly susceptible to manipulation.
Re-evaluating the wisdom of crowds in assessing web Security [5]	Pern Hui Chia	Norway	Examine the wisdom of crowds on web of trust that has similarity with Phishtank as a user based system.
Automatic detection of phishing target from phishing webpage [31]	Gang Liu	China	Phishtank database is used to test the phishing target identification accuracy of their method.

A method for the automated detection of phishing websites through both site characteristics and image analysis [56]	Joshua S. White	New york, US.	Phishtank database is used to perform additional validation of their method. They also collect data from twitter using twitter's API to find malicious tweets containing phishing URLs
Intelligent phishing detection and protection scheme for online transaction [1]	P.A. Barraclough	Newcastle, United Kingdom	Phishtank features is used as one of the input of neuro fuzzy technique to detect phishing website. The study suggested 72 features from Phishtank by exploring journal papers and 200 phishing website.
Towards preventing QR code based attacks on android phone using security warning [59]	Huiping Yao	New Mexico, US.	Phishtank API is used for lookup whether the given QR containing phishing URL in the Phishtank database.
A SVM based technique to detect phishing URLs [19]	Huajun Huang	China	Phishtank database is used as validation resulting 99% accuracy by SVM method, plus the top ten brand names in Phishtank archive is used as features in SVM method.
Socio technological phishing prevention [16]	Gaurav Gupta	Australia	Analyze the Phishtank verifiers (individual/organization) to be used as anti phishing model.
An evaluation of lightweight classification methods for identifying malicious URLs [12]	Shaun Egan	Grahamstown, South Africa	Indicating that lightweight classification methods achieves an accuracy of 93% to 96% when trained data from Phishtank.
Phi.sh/\$oCiaL: The phishing landscape through short URLs [4]	Sidharth Chhabra	Delhi, India	Phishtank database is used to analyze suspected phish that is done through short URLs.
Discovering phishing target based on semantic link network [54]	Liu Wenyin	Hong Kong	Phishtank database is used as test dataset to verify their proposed method (Semantic Link Network)

From our literature survey, we know that Phishtank is crowd-sourced platform to manage phishing URLs. For that reason Moore, et al. aims to evaluate the wisdom of crowds platform accommodated by Phishtank [40]. Moore, et al. suggest that the user participation is distributed according to power law. It uses to model data which frequency of an event varies as a power of some attribute of that event [27]. Power law also applies to a system when large is rare and small

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

Table 3: Comparison summary [40]

is common ³. For example, in the case of individual wealth in a country, 80% of the all wealth is controlled by 20% of population in a country. It makes sense that in Phishtank’s verification system, a single highly active user’s action can greatly impact the system’s overall accuracy. Table 2 summarizes the comparison performed by [40] between Phishtank and closed proprietary anti-phishing feeds⁴. Moreover, there are some ways to disrupt Phishtank verification system; submitting invalid reports accusing legitimate website, voting legitimate website as phish, and voting illegitimate website as not phish. While all the scenarios described are for the phishers’ benefit, the last scenario is more direct and the first two actions rather subtle intention to undermine Phishtank credibility.

To put it briefly, the lesson of crowd sourced anti-phishing technology such as Phishtank is that the distribution of user participation matters. It means that if a few high value participants do something wrong, it can greatly impact overall system [40]. Also, there is a high probability that bad users could also extensively participate in submitting or verifying URLs in Phishtank.

Machine learning approach in detecting spoofed website

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

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

⁴The author conceals the identity of the closed proprietary company

and masquerade as legitimate institutions. It does not make sense if phishers created them in a way very distinctive with its target. It may raise suspicions with result of unsuccessful attack. To put it another way, while it might be true, we speculated that most of the phishing websites are mostly identical with its legitimate websites as target to reduce suspiciousness from potential victim.

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

Furthermore, there are also several learning algorithms to classify the data, such as, logistic regression, random forest, neural networks and support vector machine. However, as this particular topic is out of scope of our research, we will not discuss about the learning algorithm that is currently implemented. We will only introduce three features that are used in machine learning based detection.

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

- Lexical features

Lexical features (URL based features) are based on the analysis of URL structure without any external information. Ma, et al. suggest that the structure URL of phishing may “look” different to experts [33]. These features include how many dots exist, the length, how deep the path traversal do the URL has or if there any sensitive words present in a URL. For example the URLs <https://www.paypal.com> and <http://www.paypal.com.example.com/> or <http://login.example.com/www.paypal.com/>, we can see that the domain www.paypal.com positioned differently, with the first one being the benign URL. Figure 8 shows an example analysis of lexical features in a phishing URL [29].

Lexical features analysis may have performance advantage and reduces overhead in term of processing and latency, since it only tells the machine to learn URL structure. 90% accuracy is achieved when utilizing lexical features combined with external features such as WHOIS data [29]. Egan, et al. conducted an evaluation of lightweight classification that includes lexical features and host based features in its model [12]. The study found that the classification based on these features resulted in extremely high accuracy and low overhead. Table 4 lists the existing lexical features that are currently implemented by two different studies [58, 32]. However, Xiang, et al.[58] pointed out that URLs structure could be manipulated with little cost, causing the features to fail. For example, attackers could simply remove embedded domain and sensitive words

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

Figure 8: Example lexical features [29]

Haotian Liu, et al. [32]	Guang Xiang, et al. [58]
<ul style="list-style-type: none"> - 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 	<ul style="list-style-type: none"> - Embedded domain - IP address presence - Number of dots - Suspicious URL - Number of sensitive words - Out of position top level domain (TLD)

Table 4: Existing lexical features [32, 58]

to make their phishing URLs look legitimate. Embedded domain feature examines whether a domain or a hostname is present in the path segment [58], for example, <http://www.example.net/path/to/www.paypal.com>. Suspicious URL feature examine whether the URL has “@” or “-”, the present of “@” is examined in a URL because when the symbol “@” is used, the string to the left will be discarded. Furthermore, according to [58], not many legitimate websites use “-” in their URLs. There are also plenty of legitimate domains presented only with IP address and contains more dots. Nevertheless, lexical analysis would be suitable features to use for first phase analysis in a large data [12].

- Host based features

Since phishers often hosted phishing websites in less reputable hosting services and registrars, host-based features are needed to observe on the external sources (WHOIS information, domain information, etc.). A study suggests host-based features have the ability to describe where phishing websites are hosted, who owns them and how they are managed [33]. Table 5 shows the host-based

Justin Ma, et al.[33, 34]	Haotian Liu, et al. [46][32]	Guang Xiang, et al. [58]
- WHOIS data	- Autonomous system number	- Age of
- IP address information	- IP country	Domain
- Connection speed	- Number of registration information	
- Domain name properties	- Number of resolved IPs	
	- Domain contains valid PTR record	
	- Redirect to new site	
	- All IPs are consistent	

Table 5: Host-based features [33, 34, 32, 58]

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

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

1. WHOIS information: Since phishing websites and hacked domains are often created at relatively young age, this information could provide the registration date, update date and expiration date. Domain ownership would also be included; therefore, a set of malicious websites with the same individual could be identified.
 2. IP address information: Justin Ma, et al. used this information for identify whether or not an IP address is in blacklist [34, 33]. Besides the corresponding IP address, it provides records like nameservers and mail exchange servers. This allows the classifier to be able to flag other IP addresses within the same IP prefix and ASN.
 3. Domain name properties: these include time to live (TTL) of DNS associated with a hostname. PTR record (reverse DNS lookup) of a domain could also be derived whether it is valid or not.
- Site popularity features

Site popularity could be an indicator whether a website is phishy or not. It makes sense if a phishing website has much less traffic or popularity than a legitimate website. According to [58], some of the features indicated in Table 6 are well performed when incorporated with machine learning system.

Guang Xiang, et al. [58]	Haotian Liu, et al. [32]
- Page in top search results	- Number of external links
- PageRank	- Real traffic rank
- Page in top results when searching copyright company name and domain	- Domain in reputable sites list
- Page in top results when searching copyright company name and hostname	

Table 6: Site popularity features [58, 32]

1. Page in top search results: this feature originally used by [61] to find whether or not a website shows up on the top N search result. If it is not the case, the website could be flagged as phishy since phishing websites have less chance of being crawled [58]. We believe this feature is similar to 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 [58], the intuitive rationale of this feature is that phishing websites are often have very low PageRank due to their ephemeral nature and very low incoming links that are redirected to them. This feature similar to Real traffic rank feature employed by [32] 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 [58] and the latter feature can be obtained from amazon.com [32].

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 [2, 82] specify the structure of phishing emails properties as follows:

1. Spoofing of online banks and retailers. Impersonation of legitimate institutions may created in the email level. Phishers may design a fake email to resemble the reputable company to gain users trust.
2. Link in the text is different from the destination. A link(s) contained in the email message usually appear to different than the actual link destination. This method used to trick users to believe that the email is legitimate.

3. Using IP addresses instead of URLs. Sometimes phishers may hide the link in the message by presenting it as IP address instead of URL.
4. Generalization in addressing recipients. As phishing emails are distributed by large number of recipients, the email often is not personalized, unlike the legitimate email that address its recipient by personalized information such as the last four digits of account information.
5. Usage of well-defined situational contexts to lure victims. Situational contexts such as false urgency and threat are a common method to influence the decision making of the recipients.

Moreover, Ma, et al. experimented with seven properties to consider in a phishing emails consist of the total number of links, total numbers of invisible links, whether the link that appears in the message is different than the actual destination, the existence of forms, whether scripts exist within an email, total appearance of blacklisted words in the body and the total appearance of blacklisted words in the subject [79]. Based on this survey, we will be establishing phishing email properties as variables in order to classify our data.

Phishing prevention

Phishing attacks aim to by-pass technological countermeasures by manipulating users' trust and can lead to monetary losses. Therefore, human factors take a big part on the phishing taxonomy, especially in the organizational environment. Human factor in phishing taxonomy comprised of education, training and awareness [15]. Figure 9 illustrates where human factor takes part on phishing threats [15]. User's awareness of phishing has been explored by several studies [22, 15, 13, 26, 23, 11] as preventive measure against phishing attack. According to ISO/IEC 27002 [15][43], it has been shown that information security awareness is important and it has been critical success factors to mitigate security vulnerabilities that attack user's trust. One approach to hopefully prevent phishing attack was by implementing anti phishing warning/indicator. Dhamija, et al suggest that users often ignore security indicators thus makes them ineffective [10]. Even if users notice the security indicators, they often do not understand what they represent.

Moreover, the inconsistency of positioning on different browsers makes them much difficult to identify phishing [24]. Evidently, Schechter, et al. pointed out that 53% of their study participants were still attempting to provide their confidential information, even after their task was interrupted by strong security warning [50]. Therefore, these suggest that an effective phishing education must be added as a complementary strategy to complete technical anti-phishing measure as a strong remedy to detect phishing websites or emails.

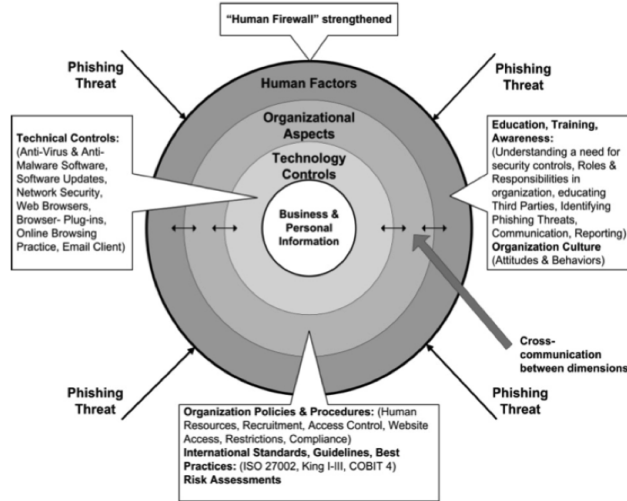
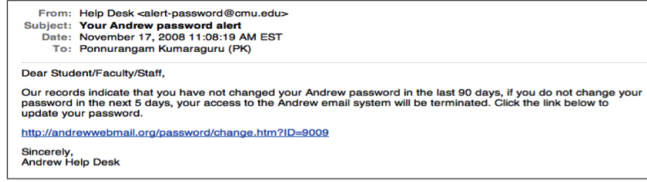
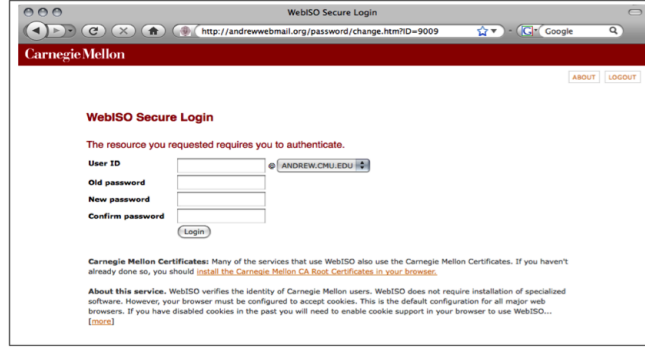


Figure 9: Holistic anti-phishing framework [15]

Phishing education for online users often by instructing not to click links in an email, ensure that SSL is present and to verify that the domain name is correct before giving information, and other similar education. This traditional practice evidently has not always effective [13]. One may ask what makes phishing education effective? A study suggests that in order online users to be aware of phishing threats, is to really engage them to so that they understand how vulnerable they are [35]. To do this, simulated phishing attacks often performed internally in an organization. Figure 10 shows a simulated phishing email and website carried out by Kumaraguru, et al. from PhishGuru [25]. As a result, this scenario puts them in the ultimate teachable moment if they fall for these attacks.



(a) simulated phishing email [25]



(b) simulated phishing website [25]



(c) simulated phishing message [25]

Figure 10: Simulated phishing attack [25]

Phishguru is a security training system operated by Wombat security technology that teaches users not to be deceived by phishing attempts by simulation of phishing attacks[76]. They claimed Phishguru provides more effective training than traditional training as it is designed to be more engaging. Figure 11 illustrates how embedded phishing training was presented by PhishGuru.

Kumaraguru, et al. investigates the effectiveness of embedded training methodology in a real world situation [25]. Evidently, they indicated that even after 28 days after training, users trained by PhishGuru were less likely to click the link presented in the simulated phishing email than those who were not trained. They also find that users who trained twice were less likely to give information to simulated fraudulent website than users who were trained once. Moreover, they argue that the training does not decrease the users' willingness to click on the links from legitimate emails; it means that less likely a trained user did a false positive when he or she requested to give information from true legitimate emails [25]. This suggests that user training strategy as an ef-

fective phishing education in order to improve phishing awareness especially in organizational environment.

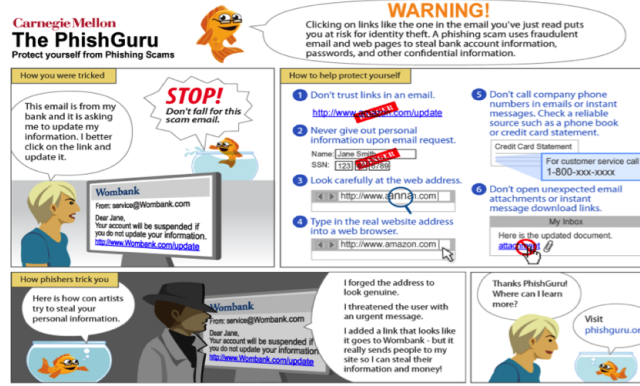


Figure 11: Embedded phishing training [25]

Research questions and Hypotheses

One of the primary objectives of our literature reviews is to summarize the existing literature that explain the general topic about phishing and formulate the research questions in the process. We continue this report by enumerating our research questions and also mentioning our hypotheses related to these questions. We aim to test these hypotheses by the data collected during our study of phishing emails provided by fraudehelpdesk.nl. Firstly, we wanted to know the characteristics of phishing email based on structural properties in our corpus.

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

The characteristics of the reported phishing emails includes:

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

- Relationship between generic properties

Secondly, we wanted to know how relevant are the persuasive principles to the associated phishing email properties.

RQ2: How relevant are the persuasive principles to the generic phishing email properties?

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

In our coding of cialdini's principles and phishing email dataset, we identified phishing emails with fake logos and signatures that may mistakenly regard them as legitimate by average internet users. For example in the context of phishing email, signature such as "Copyright 2013 PayPal, Inc. All rights reserved" or "Administrator Team" and Amazon logo were used to show the "aura of legitimacy". In the real world society, telemarketers and seller has been using authoritative element to increase the chance of potential consumer's compliance [62]. It means that they have to provide information in a confident way. Consumers will have their doubt if sellers unsure and nervous when they offer their product and services to consumers. This principle has been one of the strategies in a social engineering attack to acquire action and response from a target [63].

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

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

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

Similar to authority principle that may trigger reactance, scarce items and shortage may produce immediate compliance from people. In essence, people will react when their freedom is restricted about valuable matter when they think they are capable to make a choice among different options [64]. For example in phishing email context, an email from Royal Bank inform us that we have not been logged into our online banking account for a quite some time, as a security measure, they must suspend our online account and if we would

like to continue to use the online banking facility, we have been asked to click the URL provided. Potential victim may perceive their online banking account as their valuable matter to access facility and information about their savings. Consequently, potential victim may react to the request because of their account could be scarce and restricted. In the real world example, a hard worker bank customer who perceives money is a scarce item may immediately react when his bank inform him that he is in danger of losing his savings due to “security breach”. We therefore hypothesize that

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

As we describe in our decision making consideration section, people tend to trust those they like. In a context of persuasion, perpetrators may find it more difficult to portray physical attractiveness, instead they are relying on emails, websites and phone calling [65]. To exhibit charm or charisma to the potential victims, perpetrators may gain their trust by establishing friendly emails, affectionate websites and soothing voice over the phone. In the phishing email context, Amazon praises our existence in an appealing fashion and extremely values our account security so that no one can break it. Based on this scenario, E-commerce/Retails sector may applied likeability principles to gain potential customers. We therefore hypothesize that

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

Tajfel, et al. argued that people often form their own perception based on their relationship with others in a certain social circles [66]. This lead to affection of something when significant others have something to do with it. Social proof is one of the social engineering attacks based on the behavioral modeling and conformance [57] For example, we tend to comply to a request when a social networking site asks us to visit a website or recommends something and mention that others have been visiting the website as well. Thus, we hypothesize that

H5: Phishing emails which targeting Social networks will likely to have 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 deemed valuable. For example, a perpetrator masquerades as an authority and dressed as police officer halted us on the road, the perpetrator may tell us that we did something wrong and he will held our driving license if we do not pay him the fine. In the phishing email context, an email masquerades as “System Administrator” may tell us that we exceeded our mailbox quota, so the administrator must freeze our email account and we could re-activate it by clicking the URL provided in the email. Based on this

scenario, we know that it has authority principle and also has scarcity principle. Therefore, we hypothesize that

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

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

H7: Phishing emails which have likeability principle will likely to have consistency principle

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

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

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

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

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

We sometimes find it suspicious if a person dressed as police officer that does not have a badge carried with him, unless he is a fake police officer. Consequently, a fake police officer may use a fake badge to build up even more “aura of legitimacy”. Evidently, Cialdini suggests the increment of passerby who have stop and stare at the sky by 350 percent with suit and tie instead of casual dress [6]. Hence, we correlate that a person who wears police uniform and a fake badge in the real world context as authority principle and the presence of image in the phishing mail context. Another example, an email that masquerades Apple company, may clone Apple company logo or trademark to its content to increase the chance of potential victim’s response or increase the believability” if you will. Thus, we hypothesize that

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

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

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

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

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

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

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

Analogous with the hypothesis *H14*, it is make sense if a phishing email which has document related reason such as reviewing contract agreement or

reviewing resolution case, would tend to have a file to be attached. We therefore hypothesize that

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

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

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

Research methodology

The data for this research will be collected from fraudehelpdesk.nl as the “Nationale helpdesk voor vragen en meldingen over fraude” [81] in the Netherlands. We are looking for reported phishing emails in 2013. We will be enumerating our methodology into several steps:

1. Raw data collection
2. Data categorization
3. Variable establishment
5. Data classification
6. Analysis

Data categorization will be determined by several aspects; languages (Dutch, English, others) and whether the email is indeed a phish, spam or legitimate. We predict there will be duplicated phishing emails so that we will only consider only unique email in our dataset.

We have studied the structural properties of phishing emails in the previous section and we will establish our variable based on these properties. Furthermore, the data classification phase will be indicated by our coding of the categorization result into these variables. Lastly, the data will be analyzed from mainly three different viewpoints; properties characteristics, persuasion principles characteristics and their relationships.

Conclusion

In this report, we have tried to summarize all relevant literature. We have outline the basic understanding related to general phishing in respect to its history. We have also reviewed its universal definition from various sources. One study

aggregated and systematically achieved the consensual definition of phishing in which we follow. We have also briefly explain the damage of phishing in term of direct cost. We have learned that human factor is one of the most critical aspects in phishing. Many studies suggest several phases when phishing attack occurred and we tried to synthesized these phases and determined our own phases based on the attacker point of view. We have outlined the general type of phishing and the existence of phishing badhoods. Many studies suggest both technical and social countermeasures such as restriction list, machine learning approach to detect spoofed website and anti phishing training. However, we have found an absence of empirical research conducted to investigate the role of persuasive elements as a human factor in respect to phishing email which can be an aspect to be considered to prevent phishing attack.

We have formulated two research questions and some hypotheses based on our impressions from literature. We have also outlined a methodology to perform the data analysis that hopefully can answer the research questions from its findings.

Bibliography

- [1] P. Barraclough, M. Hossain, M. Tahir, G. Sexton, and N. Aslam, “Intelligent phishing detection and protection scheme for online transactions,” *Expert Systems with Applications*, vol. 40, no. 11, pp. 4697–4706, 2013.
- [2] M. Chandrasekaran, K. Narayanan, and S. Upadhyaya, “Phishing email detection based on structural properties,” in *NYS Cyber Security Conference*, 2006, pp. 1–7.
- [3] Z. Chen and C. Ji, “Optimal worm-scanning method using vulnerable-host distributions,” *International Journal of Security and Networks*, vol. 2, no. 1, pp. 71–80, 2007.
- [4] S. Chhabra, A. Aggarwal, F. Benevenuto, and P. Kumaraguru, “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.
- [5] P. H. Chia and S. J. Knapskog, *Re-evaluating the wisdom of crowds in assessing web security*. Springer, 2012, pp. 299–314.
- [6] R. Cialdini, “The science of persuasion,” *Scientific American Mind*, 2001.
- [7] R. Clayton, D. Dean, M. Jakobsson, S. Myers, S. Stubblebine, and M. Szydlo, “A chat at the old phishin’hole,” *Lecture Notes in Computer Science*, vol. 3570, p. 88, 2005.
- [8] C. E. Dictionary, “Phishing.” [Online]. Available: <http://www.collinsdictionary.com/dictionary/american/phishing>

- [9] M. P. Collins, T. J. Shimeall, S. Faber, J. Janies, R. Weaver, M. De Shon, and J. Kadane, "Using uncleanness to predict future botnet addresses," in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*. ACM, 2007, pp. 93–104.
- [10] R. Dhamija, J. D. Tygar, and M. Hearst, "Why phishing works," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006, pp. 581–590.
- [11] R. C. Dodge Jr, C. Carver, and A. J. Ferguson, "Phishing for user security awareness," *Computers & Security*, vol. 26, no. 1, pp. 73–80, 2007.
- [12] S. Egan and B. Irwin, "An evaluation of lightweight classification methods for identifying malicious urls," in *Information Security South Africa (ISSA), 2011*. IEEE, Conference Proceedings, pp. 1–6.
- [13] A. Emigh, "Online identity theft: Phishing technology, chokepoints and countermeasures," *ITTC Report on Online Identity Theft Technology and Countermeasures*, vol. 3, 2005.
- [14] P. V. Fellman and R. Rodriguez, "The dark side of the internet," in *International Federation for Information Processing, International Meeting, "IT Innovation for Adaptability and Competitiveness"*, Conference Proceedings.
- [15] E. D. Fraumenstein and R. von Solms, *An Enterprise Anti-phishing Framework*. Springer, 2013, pp. 196–203.
- [16] G. Gupta and J. Pieprzyk, "Socio-technological phishing prevention," *Information Security Technical Report*, vol. 16, no. 2, pp. 67–73, 2011.
- [17] C. Herley and D. 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.
- [18] J. Hong, "The state of phishing attacks," *Communications of the ACM*, vol. 55, no. 1, pp. 74–81, 2012.
- [19] H. Huang, L. Qian, and Y. Wang, "A svm-based technique to detect phishing urls," *Information Technology Journal*, vol. 11, no. 7, 2012.
- [20] M. Jakobsson, "Modeling and preventing phishing attacks," in *Financial Cryptography*, vol. 5. Citeseer, Conference Proceedings.
- [21] M. Jakobsson and S. Myers, *Phishing and countermeasures: understanding the increasing problem of electronic identity theft*. John Wiley & Sons, 2006.
- [22] L. James, *Phishing exposed*. Syngress, 2005.
- [23] K. Jansson and R. Von Solms, "Phishing for phishing awareness," *Behaviour & Information Technology*, vol. 32, no. 6, pp. 584–593, 2013.

- [24] I. Kirlappos and M. A. Sasse, “Security education against phishing: A modest proposal for a major rethink,” *IEEE Security and Privacy Magazine*, vol. 10, no. 2, pp. 24–32, 2012.
- [25] P. Kumaraguru, J. Cranshaw, A. Acquisti, L. Cranor, J. Hong, M. A. Blair, and T. Pham, “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.
- [26] P. Kumaraguru, S. Sheng, A. Acquisti, L. F. Cranor, and J. Hong, “Lessons from a real world evaluation of anti-phishing training,” in *eCrime Researchers Summit, 2008*. IEEE, Conference Proceedings, pp. 1–12.
- [27] W. Lai, “Fitting power law distributions to data.”
- [28] E. Lastdrager, “Achieving a consensual definition of phishing based on a systematic review of the literature,” 2014.
- [29] A. Le, A. Markopoulou, and M. Faloutsos, “Phishdef: Url names say it all,” in *INFOCOM, 2011 Proceedings IEEE*. IEEE, Conference Proceedings, pp. 191–195.
- [30] —, “Phishing victims likely will suffer identity theft fraud,” *Gartner Research Note (May 14, 2004)*, 2004.
- [31] G. Liu, B. Qiu, and L. Wenying, “Automatic detection of phishing target from phishing webpage,” in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, Conference Proceedings, pp. 4153–4156.
- [32] H. Liu, X. Pan, and Z. Qu, “Learning based malicious web sites detection using suspicious urls.”
- [33] J. Ma, L. K. Saul, S. Savage, and G. M. Voelker, “Beyond blacklists: learning to detect malicious web sites from suspicious urls,” in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, Conference Proceedings, pp. 1245–1254.
- [34] —, “Identifying suspicious urls: an application of large-scale online learning,” in *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM, Conference Proceedings, pp. 681–688.
- [35] S. Mansfield-Devine, “Interview: Joe ferrara—fighting phishing,” *Computer Fraud & Security*, vol. 2013, no. 7, pp. 17–20, 2013.
- [36] T. McCall, “Gartner survey shows phishing attacks escalated in 2007; more than \$3 billion lost to these attacks,” *Stephane GALLAND*, 2007.
- [37] M. Webster, “Phishing.” [Online]. Available: <http://www.merriam-webster.com/dictionary/phishing>

- [38] K. D. Mitnick and W. L. Simon, *The art of deception: Controlling the human element of security*. Wiley, 2001.
- [39] T. Moore and R. Clayton, “An empirical analysis of the current state of phishing attack and defence,” in *WEIS*. Citeseer, Conference Proceedings.
- [40] —, *Evaluating the wisdom of crowds in assessing phishing websites*. Springer, 2008, pp. 16–30.
- [41] G. C. Moura, *Internet bad neighborhoods*. Giovane Cesar Moreira Moura, 2013.
- [42] P. J. Nero, B. Wardman, H. Copes, and G. Warner, “Phishing: Crime that pays,” in *eCrime Researchers Summit (eCrime), 2011*. IEEE, Conference Proceedings, pp. 1–10.
- [43] O. I. de Normalización, *ISO/IEC 27002: Information Technology, Security Techniques, Code of Practice for Information Security Management*. ISO/IEC, 2005.
- [44] O. Dictionaries, “Phishing.” [Online]. Available: <http://www.oxforddictionaries.com/definition/english/phishing>
- [45] P. Parmar and K. Patel, “Comparison of phishing detection techniques,” in *International Journal of Engineering Research and Technology*, vol. 3. ESRSA Publications, Conference Proceedings.
- [46] B. Parno, C. Kuo, and A. Perrig, *Phoolproof phishing prevention*. Springer, 2006.
- [47] Phishing.org, “History of phishing.” [Online]. Available: <http://www.phishing.org/history-of-phishing/>
- [48] S. Purkait, “Phishing counter measures and their effectiveness - literature review,” *Information Management & Computer Security*, vol. 20, no. 5, pp. 382–420, 2012.
- [49] A. Ramachandran and N. Feamster, “Understanding the network-level behavior of spammers,” in *ACM SIGCOMM Computer Communication Review*, vol. 36. ACM, Conference Proceedings, pp. 291–302.
- [50] S. E. Schechter, R. Dhamija, A. Ozment, and I. Fischer, “The emperor’s new security indicators,” in *Security and Privacy, 2007. SP’07. IEEE Symposium on*. IEEE, Conference Proceedings, pp. 51–65.
- [51] S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs, “Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions,” in *Proceedings of the 28th international conference on Human factors in computing systems*. ACM, Conference Proceedings, pp. 373–382.

- [52] G. Tally, R. Thomas, and T. Van Vleck, "Anti-phishing: Best practices for institutions and consumers," *McAfee Research*, Mar, 2004.
- [53] B. Wardman, G. Shukla, and G. Warner, "Identifying vulnerable websites by analysis of common strings in phishing urls," in *eCrime Researchers Summit, 2009. eCRIME'09*. IEEE, Conference Proceedings, pp. 1–13.
- [54] L. Wenying, N. Fang, X. Quan, B. Qiu, and G. Liu, "Discovering phishing target based on semantic link network," *Future Generation Computer Systems*, vol. 26, no. 3, pp. 381–388, 2010.
- [55] R. Wetzel, "Tackling phishing," *Business Communications Review*, vol. 35, no. 2, pp. 46–49, 2005.
- [56] J. S. White, J. N. Matthews, and J. L. Stacy, "A method for the automated detection phishing websites through both site characteristics and image analysis," in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, Conference Proceedings, pp. 84 080B–84 080B–11.
- [57] M. 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*, vol. 59, no. 4, pp. 662–674, 2008.
- [58] G. Xiang, J. Hong, C. P. Rose, and L. Cranor, "Cantina+: a feature-rich machine learning framework for detecting phishing web sites," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 2, p. 21, 2011.
- [59] H. Yao and D. 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, Conference Proceedings, pp. 341–346.
- [60] Y. Zhang, S. Egelman, L. Cranor, and J. Hong, "Phinding phish: Evaluating anti-phishing tools." ISOC, Conference Proceedings.
- [61] Y. Zhang, J. I. Hong, and L. F. Cranor, "Cantina: a content-based approach to detecting phishing web sites," in *Proceedings of the 16th international conference on World Wide Web*. ACM, Conference Proceedings, pp. 639–648.
- [62] I. Telemarketing, "5 tips for getting past receptionists!" 2013, [Online; accessed 16-July-2014]. [Online]. Available: <http://inspiredtelemarketing.wordpress.com/2013/09/13/5-tips-for-getting-past-receptionists/>
- [63] N. P. D. Network, "Types of social engineering," [Online; accessed 16-July-2014]. [Online]. Available: http://www.npdn.org/social_engineering_types

- [64] J. W. Pennebaker and D. Y. Sanders, "American graffiti: Effects of authority and reactance arousal," *Personality and Social Psychology Bulletin*, vol. 2, no. 3, pp. 264–267, 1976.
- [65] D. P. Dotterweich and K. S. Collins, "The practicality of super bowl advertising for new products and companies," *Journal of Promotion Management*, vol. 11, no. 4, pp. 19–31, 2006.
- [66] H. Tajfel and J. C. Turner, "The social identity theory of intergroup behavior." 2004.
- [67] T. McCall, "Gartner survey shows phishing attacks escalated in 2007; more than \$3 billion lost to these attacks," *Stephane GALLAND*, 2007.
- [68] G. C. Moura and A. Pras, "Scalable detection and isolation of phishing," in *Scalability of Networks and Services*. Springer, 2009, pp. 195–198.
- [69] R. FraudAction, "Rsa monthly fraud report," [Online; accessed 6-August-2014]. [Online]. Available: <http://www.emc.com/security/rsa-identity-protection-and-verification/rsa-fraudaction/rsa-fraudaction-antiphishing-service.htm>
- [70] A. Greenberg, "Medical staffers fall for phishing emails, data on 8,300 compromised," [Online; accessed 13-May-2014]. [Online]. Available: <http://www.scmagazine.com/medical-staffers-fall-for-phishing-emails-data-on-8300-compromised/article/340590/>
- [71] A. Carman, "Phishing scam targets michigan public schools," [Online; accessed 13-May-2014]. [Online]. Available: <http://www.scmagazine.com/phishing-scam-targets-michigan-public-schools/article/343177/>
- [72] T. Robinson, "Phishing scam aimed at google docs, drive users," [Online; accessed 13-May-2014]. [Online]. Available: <http://www.scmagazine.com/phishing-scam-aimed-at-google-docs-drive-users/article/338369/>
- [73] B. P. Department, "Portaal van de belgische politie," [Online; accessed 13-May-2014]. [Online]. Available: <http://www.polfed-fedpol.be/>
- [74] —, "Overzicht per criminele figuur," [Online; accessed 13-May-2014]. [Online]. Available: http://www.polfed-fedpol.be/crim/crim_statistieken/app_crimestat/app_crimestat_dashboard_crimfig_misdrijven_nl.php
- [75] OpenDNS, "Phishtank: Out of the net, into the tank," [Online; accessed 13-May-2014]. [Online]. Available: <http://www.phishtank.com/faq.phpk>
- [76] W. security technology, "Phishguru: Assess and motivate your employees using simulated phishing attacks," [Online; accessed 23-May-2014]. [Online]. Available: <http://www.wombatsecurity.com/phishguru>

- [77] M. Blythe, H. Petrie, and J. 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.
- [78] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, “Social phishing,” *Communications of the ACM*, vol. 50, no. 10, pp. 94–100, 2007.
- [79] L. Ma, B. Ofoghi, P. Watters, and S. Brown, “Detecting phishing emails using hybrid features,” in *Ubiquitous, Autonomic and Trusted Computing, 2009. UIC-ATC’09. Symposia and Workshops on*. IEEE, 2009, pp. 493–497.
- [80] F. Stajano and P. Wilson, “Understanding scam victims: seven principles for systems security,” *Communications of the ACM*, vol. 54, no. 3, pp. 70–75, 2011.
- [81] M. van Veiligheid en Justitie, “Nationale helpdesk voor vragen en meldingen over fraude,” [Online; accessed 13-July-2014]. [Online]. Available: <http://www.fraudehelpdesk.nl/over-ons>
- [82] C. E. Drake, J. J. Oliver, and E. J. Koontz, “Anatomy of a phishing email.” in *CEAS*, 2004.