Phishing in email and instant messaging

University of Oulu
Faculty of Information Technology and Electrical Engineering / Information Processing Sciences
Bachelor's thesis
Benjamin Mustonen
09.05.2019

Abstract

Phishing is a constantly evolving threat in the world of information security that affects everyone, no matter if you're a retail worker or the head of IT in a large organisation. Because of this, this thesis aims to give the reader a good overview of what phishing is, and due to its prevalence in email and instant messaging, focuses on educating the reader on common signs and techniques used in phishing in the aforementioned forms of communication. The chosen research method is literature review, as it is the ideal choice for presenting an overview of a larger subject. As a result of the research, many common phishing signs and techniques in both email and instant messaging are presented. Some of these signs include strange senders, fake domain names and spellings mistakes. With this thesis, anyone looking to improve their understanding about phishing can do so in a way that is easy to understand. Some suggestions for future research are also presented based on this thesis' shortcomings, namely the lack of studies on phishing in instant messaging.

Keywords

phishing, literature review, email spoofing, fake domain names, information security

Supervisor

Postdoctoral Researcher, PhD Mari Karjalainen

Contents

1. Introduction

This thesis reviews previous literature on phishing and phishing techniques, focusing specifically on email and instant messaging. The motivation for the thesis is to review how previous phishing literature holds up in today's world, as phishing techniques and strategies have become more advanced and harder to spot. The thesis also aims to give a good overview of the subject as a whole. Important prior knowledge on the subject is explained in the second section, and includes topics such as social engineering, spear phishing, malware, cybercrime, and anchor text in hyperlinks.

Phishing is very prevalent in email and IM, which is why the thesis focuses on reviewing literature pertaining to these two cases. Both email and IM phishing attacks very often link to phishing sites (Zhao et al. 2017), which is why phishing sites themselves are covered as well. The main contribution to be derived from the thesis is a look at what the common signs of a phishing attack are and to give a comprehensive picture of the subject for those looking for an overview.

The thesis begins by introducing prior knowledge required to understand the main topics that are covered. These consist mainly of phishing related terminology and techniques used in executing phishing scams. The research method is explained afterwards. Following these sections is the study itself, reviewing literature on phishing in email as well as IM, and providing some examples of phishing to educate the reader on common signs of a phishing scam. The findings of the study are discussed with their implications in the fifth section of the thesis, followed by the conclusion and suggestions for future research.

2. Prior research

This section describes phishing related terminology as well as some of the relevant techniques used in the execution of these attacks, based on prior research.

Social engineering means influencing people to comply with the requests of the attacker, often in an effort to gain access to sensitive information, with the attack involving a computer-related entity (Mouton, Leenen & Venter 2016). There are several types of social engineering attacks, but this thesis only focuses on one of them: Phishing.

Malicious software, also known as malware, is software that is designed to attack other software, such as the computer's operating system, in a manner that causes unintended behaviour. Malware can manifest in many different forms. One such form is referred to as a Trojan horse, a malicious payload of a legitimate program. The program works as the user thinks it's intended, leaving the user believing that they are safe, while the malicious, undesirable functions are executed unnoticed. Self-replicating malware that are intended to spread from one system to another, via the computer network for example, are referred to as worms, and they don't necessarily contain a damaging payload. Malware that modify other software in order to cause damage are referred to as viruses. (Kramer & Bradfield 2010.) Spyware are a type of malware that monitor a user without their consent by tracking their web browsing or recording their keystrokes (Federal Trade Commission 2004). Spyware that record keystrokes are referred to as keyloggers (Jakobsson & Myers 2006). Malware known as ransomware encrypt the victim's files, restricting the usage of the infected device and displaying a message that suggests the victim pay a ransom to decrypt their files or else the files will be gone forever (Mohan & Kumar 2017).

Instant messaging, abbreviated as IM, is a form of online communication via messages that are happening in real-time. Because of this, instant messaging requires both users to be online at the same time. Though most exchanges are text-only, many IM applications support the sending of voice messages as well as file sharing. (Rouse, 2008.)

Phishing is a form of social engineering where a phisher attempts to fraudulently acquire a victim's sensitive credentials, commonly through email or IM. The attacker typically mimics a trusted party to get the victim to visit a fraudulent site or to download some sort of malware. The attacker makes the fraudulent site mimic the look of a real website and hopes that the victim inputs their sensitive information thinking they are on the legitimate site. The word itself is a play on the word fishing, as the victim, or phish, gets lured in by the phisher's bait, the forged message. (Jakobsson & Myers 2006.) Phishing can be automated and done on a large scale, spamming as many email addresses as the attacker has access to, or it can be done in a more focused manner in the form of spear phishing (Aycock 2007).

Spear phishing is a smaller scale targeted phishing attack. It allows the attacker to make the message more customised and more effective, and since it's a unique message tailored to the specific organisation or person, it is less likely to be caught by spam filters. (Aycock 2007.)

Email spoofing is making a sent email appear to be from a different sender than it actually is. This means that an attacker can make their email appear to be from the victim's bank, for example, making the email look more legitimate. (Jakobsson & Myers 2006.)

Anchors, also called hyperlinks, are used to go from one page to another. Anchors contain a URL of the destination page as well as some text, referred to as anchor text. This text can be practically anything, from “Click here” to “A picture of an old tree in London (1998)”, and clicking it will take you to the destination. (Yi & Allan 2010.) The text could also be another URL, meaning that the text says “https://google.com/” but the destination is actually “phishingsite.net/”.

Typosquatting is registering a domain that is a misspelling of a popular domain name. The intent of typosquatting is to catch users that make a typo when typing a domain name in the address bar. There are several ways of typosquatting. The dot between www and the domain name can be removed, resulting in the domain name wwwexample.com rather than www.example.com. A character may also be omitted, swapped with a consecutive character, or be included twice in a row. Typosquatting is used for many purposes, such as displaying advertisements, redirecting to a different domain, or phishing. (Agten, Joosen, Piessens & Nikiforakis 2015.)

3. Research methods

This section starts by explaining the research method that is used in the thesis. After this, the databases and search keywords used in conducting the research are covered.

Literature review is a comprehensive synthesis of a particular topic. Literature reviews can identify what is currently known about a certain subject area as well as help formulate future research questions. They can be written as introductions or foundations in theses, or as their own primary research. A review can give readers a quick and concise overview of a certain topic at various levels of completeness. (Bolderston 2008.) As the intent of this thesis is to provide an overview of phishing in email and instant messaging, the chosen research method is literature review.

A large portion of the research was done using the Scopus database, with additional research done using Google Scholar, IEEXplore, Web of Science and the ACM Digital Library. Google Scholar was particularly useful in cases where the full text wasn't available on Scopus, whereas one could access it via Scholar. The same applied the other way around at least once, though; particularly in the case of Jakobsson and Myers' book "Phishing and countermeasures: Understanding the increasing problem of electronic identity theft." An online version of the book was only available via Scopus.

Table 1. Queries used in research

Name (in text)	Query
Query 1	"TITLE-ABS-KEY ((phishing OR "social engineering" OR scam* OR spam*) AND ("computer security" OR "internet security" OR "information security" OR is OR "cyber crim*" OR cybercrim* OR "computer related crime") AND (method* OR type* OR attack*))"
Query 2	"TITLE-ABS-KEY (("empirical study" OR survey OR experiment) AND phishing) AND (LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "ar"))"
Query 3	"TITLE-ABS-KEY (phishing AND (im OR "instant messag*"))"

Much of the preliminary material for the thesis was found using an extensive search on Scopus, using the lengthy search query of Query 1, as seen in Table 1. This query resulted in about 650 results, which provided a good starting point for research. Once the base concepts of phishing had been established using the general results provided by this search, I started specifying the queries further. Query 2 was used to find empirical studies on phishing. While some of the studies weren't relevant to what I was researching, the studies referenced in their prior research –sections were useful. Some of the sources used in this thesis were also found on Google Scholar using relatively simple search queries such as "phishing email". Surprisingly, the amount of studies on phishing in instant messaging was very low. The very broad query of Query 3 found 33 results on Scopus, compared to the email equivalent of the same query finding over 650 results. This is not to mention that a good portion of these results merely mentioned instant messaging as something that is used in phishing attacks, providing no further insight on the subject. The Malwarebytes Labs blog posts were used to supplement the scientific literature covering phishing in instant messaging, as the amount of scientific literature on the subject was highly limited.

Some of the articles have referred to IM as Instant Messengers, which the portion “instant messag*” of Query 3 would not catch. However, most of the articles did use the term IM, which meant that the query did find them in the end. Despite this, “instant mess*” would have been a more apt choice for the query.

Some of the topics that came up in phishing related studies, such as malware and email spoofing, were searched for separately using very general queries such as “malware” or “email spoofing” both in Scopus and Google Scholar.

4. The study

This section of the thesis is divided into three sections. The first section covers current literature on phishing in email as well as phishing sites. The second section covers literature on phishing in instant messaging, explaining common techniques used by IM phishers. The third section covers some of the proposed counter-phishing techniques found in current literature.

4.1. Phishing in email

Phishing attacks happen most commonly through email. These attacks are often automated and sent to hundreds of thousands of targets at a time. Spear phishing is different in that it is targeted at an individual or an organization. (Jakobsson & Myers 2006.) As phishing attacks can cause copious amounts of harm, it is imperative that people are capable of distinguishing them from real emails.

In phishing attacks, emails are forged to look like they are from a legitimate party, such as a financial institution or a bank. The email usually has a call to action, prompting the victim to enter their details on a fake webpage or to download malware. (Irani, Webb, Giffin & Pu 2008.) Phishers use a combination of several techniques to craft a convincing email. Email spoofing is a central part of a phishing email (Jakobsson & Myers 2006), as you're not likely to click on an email about your bank security from *a9xegh34@gmail.com*. Because of email spoofing, simply looking at the sender is not enough to determine whether an email is phishing or not, but any strange senders such as the previous example are likely phishing or spam.

The email is made more convincing by copying the look and layout of a legitimate email. However, some elements, such as social media links or a clickable logo, might be missing. Figure 1 is from an Airbnb phishing email. The social media images are there, but in reality, they're just part of the background; clicking them does nothing. In a real email from Airbnb, the images link to their respective social media sites. Notice how the text at the bottom also misspells Airbnb's usual "sent with ♥ from Airbnb". Indeed, Bose and Leung (2007) point out that spelling mistakes are a common feature among phishing emails.

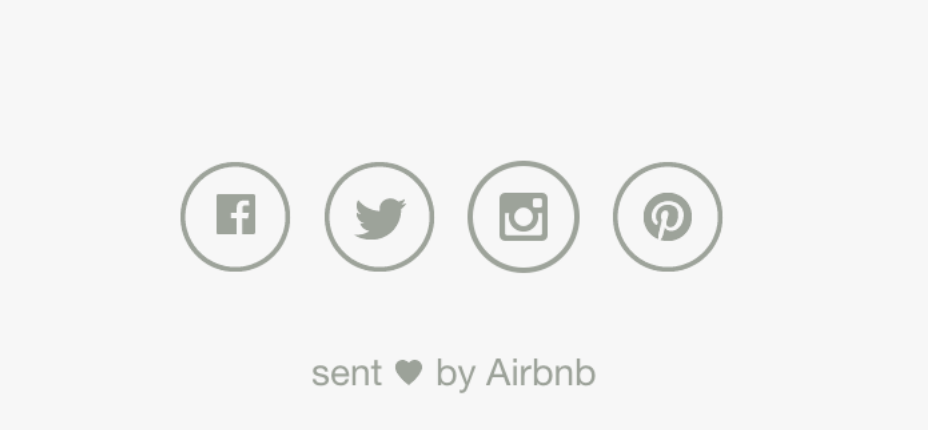

Figure 1. Screen capture from an Airbnb phishing email.

Links contained within emails should not to be trusted at face value. The address of a fake website can be hidden by putting it behind anchor text (Irani et al. 2008). I've demonstrated this in Figure 2, where the anchor text *https://www.amazon.com/login/*'s URL is actually *amazonn.net*, as shown by the browser upon hovering over the anchor text. A fake website might have a domain name very close to the site it is copying, which an unaware victim may not notice (Bose & Leung 2007).

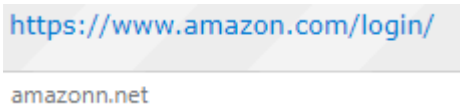

Figure 2. An example of anchor text being used to deceive the reader.

Due to certain characters looking the same, such as 1 (number) and l (letter), it may not even be necessary to conceal a phishing link within anchor text (Chow, Gustave & Vinokurov 2008). Many typosquatting techniques may also have been applied in constructing a legitimate looking domain name. Modern phishing attempts may utilize Cyrillic or Unicode characters that are indistinguishable from the expected Latin characters. This is known as a homoglyph, or homograph, attack. (Hannay & Bolan 2009.) As demonstrated by Hannay and Bolan (2009), the domain names http://www.google.com and http://www.google.com, where the first domain name uses a Cyrillic character 'g' in place of the second g, appear very similar in many fonts. According to Hannay and Bolan (2009), most modern browsers will display a link such as this as Punycode, an ASCII representation of Unicode, in the address bar as well as the status bar upon hovering over the URL, limiting its potential to fool victims. For example, the previously mentioned link with the Cyrillic character would appear as http://www.xn--goole-tmc.com.

It's not unusual for a phishing email to come with malware. This can be in the form of attached files, or the email may link the user to a website containing the malware. (Datta & Wang 2005). Phishers can create very convincing spear phishing emails by looking at the victim's social media pages, for example. Should an employee of an organisation

choose to open the malicious attachment of a spear phishing email at their workplace network, they could compromise the entire company. (Steer 2017.)

Much like emails, phishing websites disguise themselves as the original by copying the layout, font, colours, logos, et cetera (Pan & Ding 2006). According to a study by Dhamija, Tygar and Hearst (2006), this is enough to fool an alarming number of victims. 23% of the participants only looked at the contents of the website to evaluate its legitimacy, ignoring things like the address bar. Pan and Ding (2006) have suggested to pay attention to the anchors of a website. They claim that phishers are not interested in developing multiple web pages, and may simply create the bare minimum that is necessary for their attack. This means that the domains the anchors link to might differ from the current address if the phisher hasn't bothered to create new pages and simply uses ones from the real website. Alternatively, the anchors may link to the same page the victim is already on, or to no page at all. A website missing a certificate is an indicator of a dodgy website. However, some phishing sites may use SSL, and SSL certificates shouldn't be considered a mark of a legitimate site as such. It is therefore recommended that the user verifies that the certificate belongs to the party the site is claiming to be. (Pan & Ding 2006.)

Zhao et al. (2017) classify phishing attacks into three levels: simple, advanced and extreme. They base an attack's level on how similar a website looks and feels to its real counterpart. They use four metrics to define a website's look and feel: appearance, page depth, support to dynamic user interaction and phishing types. A simple phishing site is only somewhat similar to the real website. An advanced phishing site is quite similar to the real website, while an extreme phishing site is similar in every way. The page depth of a simple phishing site is one, meaning that the site does not have multiple web pages linked together. An advanced phishing site has a higher, albeit limited, page depth, increasing the trust of the victim. The page depth of an extreme phishing site is unlimited, and every link is changed. Support to dynamic user interaction refers to support of user interactions like clicking, searching, form submission and dynamic URL creation. This is generally missing from simple and advanced phishing websites but exists in extreme phishing websites. The two types of phishing are traditional and Web Single Sign-On, or Web SSO, phishing. The difference between the two types is that traditional phishing aims to steal a victim's accounts that are made for a specific website, such as a banking website, whereas SSO phishing aims to steal a victim's identity provider accounts, such as Google, which can be used for multiple websites. Simple phishing sites only support traditional phishing, while advanced and extreme phishing sites can support both. The Web SSO phishing on an advanced phishing site is of lower quality than the extreme phishing site. (Zhao et al. 2017).

Zhao et al. (2017) measured 471 existing phishing sites reported on PhishTank and found that most of the phishing websites were simple phishing websites according to their metrics, showing that the claims from Pan and Ding (2006) stand true even ten years later. Only a few were at the level of advanced phishing, and none were at the level of extreme. Zhao et al. designed and implemented a toolkit for extreme phishing to test the feasibility and effectiveness of extreme phishing attacks. Their toolkit constructs unlimited levels of phishing webpages in real time based on the victim's user interactions. Their user study's results show that 91.8% of the 194 participants were not suspicious about the extreme phishing websites, despite a majority being aware of phishing before participating in the study. They contrast this result with the around 10% success rate of existing phishing attacks, as reported in previous studies (Garera et al., 2017; Jakobsson & Ratkiewicz, 2006 as cited in Zhao et al., 2017), concluding that extreme phishing attacks are very effective. Zhao et al. recommend web users to be trained to be aware of extreme phishing,

to pay more attention to the domain name of a URL in the address bar, and to learn to differentiate spoofed Web SSO windows from real ones. (Zhao et al. 2017).

Dhamija et al. weren't the only ones who found alarming results in regard to phishing susceptibility. Mohebzada, Zarka, Bhojani and Darwish (2012) conducted two large scale phishing experiments to find out if there was any correlation between certain demographics, such as gender and age, and vulnerability to phishing attacks. For the first experiment, they sent out an email to ten thousand university members about a false password reset that was spoofed to look like it came from the university IT department. The second experiment was sent soon after, pretending to be a university research group looking for participants for a survey regarding banks. According to a previous study by Sheng, Holbrook, Kumaraguru, Cranor and Downs (2010), it was found that women and younger participants, those between the ages of 18 and 25, were found to be more vulnerable to phishing attacks. Mohebzada et al. (2012), however, found no such correlation. In their sample size of ten thousand participants, there was no conclusive evidence that gender or age were key demographics when it comes to predicting phishing susceptibility. Another key observation from their study was that warning notices against phishing attempts are often ignored. 114 subjects got phished by a password reset scam, even after receiving a warning from the IT department about the ongoing phishing attempt. Almost 10% of the 10 000 participants got phished in one way or another.

Experiments conducted by Kumaraguru, Rhee, Sheng, Hasan, Acquisti, Cranor and Hong (2007) found that using a new method of embedded training was more effective at teaching users about phishing than non-embedded training. Users also retain and transfer more information when trained this way. In the embedded training approach, users are sent fake phishing messages, and if a person falls for one of these messages, they receive an intervention that provides instant feedback on what happened and how the user could protect themselves against phishing. According to Kumaraguru et al., sending non-embedded instructional material through email does not motivate users to spend time reading the instructions, whereas in the embedded approach, users were more motivated. They also suggest that users are more likely to retain their knowledge if they are motivated to read through the training materials. (Kumaraguru et al. 2007.)

4.2. Phishing in instant messaging

Worms are a common threat when it comes to instant messaging. They exploit the trust we have in our friends. Like in emails, the phishing messages sent by worms have some sort of a lure to get the victim to click on its contents. Moore and Clayton (2015) examined a worm that had infected up to 1.67 million users in 2010. The lure of the worm was the text "foto" and a smiley face followed by a link that would infect the victim with the worm. The sender of the phishing message was a friend of the victim who had already been infected. Because of this, the victim proceeds to click on the link thinking their friend might have sent them an image. The victim then automatically forwards the phishing message to all of their friends, causing the worm to spread rapidly. (Moore & Clayton 2015.)

Not all phishing is done through worms. According to McIntire, McIntire and Havig (2010), chatbots, artificial intelligence programs that pose as humans, are used to spam victims with phishing messages. Much like automated email spam, these can be automated to send phishing links and malware to thousands of real users in a short time span.

Unlike email, instant messaging applications don't necessarily support HTML or other markup languages. Because of this, the malicious links can't be hidden behind anchor text. According to a study by Moore and Clayton (2015), phishers either use domain names that seem similar enough to the real website that the victim might not notice, or URL shorteners. If the phisher wants the victim to think that the link leads to a Facebook picture, the domain name could be something like *picture-facebook.com*. According to Moore and Clayton (2015), victims are more likely to click on a link that has a fake domain than on one that is a URL shortener, such as *bit.ly*. They also discovered that victims are more likely to get phished if the lure is in their native language, though language neutral lures like "foto" are effective as well.

Although many browsers will display domains with homoglyphs as Punycode (Hannay & Bolan 2009), these domains could still potentially fool users to click on them in instant messages if the IM client displays the homoglyphs rather than Punycode.

The Internet security company Malwarebytes has reported on their security blog, Malwarebytes Labs, about phishing attempts on the IM service of the popular digital distribution platform Steam. These phishing attempts have a lure, such as a request to look at items that the phisher wants to trade with the victim (Umawing 2014), a message saying that the victim has won a prize and a request to choose the prize from an image (Umawing 2016), or a request to download a phony voice chat application (Umawing 2015). The websites that the phishers link to utilize typosquatting URLs to fool victims into believing they are entering a legitimate site (Umawing 2014). Some of the phishing attempts used an executable with a .scr file extension to trick users into thinking that the file from the phishing site was a picture. The executable loaded a picture for the user to see while also performing malicious activity in the background. (Umawing 2014, 2016). Steam has a security feature called Steam Guard, which prevents unauthorised computers from accessing your account, even if they have your username and password. According to Boyd (2014 a), phishers can bypass Steam Guard by using the victim's SSFN file, a file which makes it so that the user doesn't have to verify their identity through Steam Guard every time they log in. Phishers have attempted to gain access to this file by asking the victim to upload it on a fake Steam phishing site (Boyd 2014 a) and by asking the victim to authenticate using a fake steamguard.exe file, which would then send the SSFN file to the phisher (Boyd 2014 b).

4.3. Counter-phishing techniques

In 2012, Singh, Sarje and Misra proposed a design for a client-side counter phishing application. The application uses an adaptive neuro-fuzzy inference system (ANFIS) to analyse the user's emails. ANFIS utilises neural networks for pattern recognition and to adapt to changes, and fuzzy inference systems for decision making. Hyperlinks within an email are analysed for phishing indicators to determine whether the email is a phishing email. The authors have selected four types of phishing indicators to use as inputs for the ANFIS: number of Dot Decimal type hyperlinks, number of Encoded hyperlinks, number of unequal visual and actual parts of hyperlinks, and number of special characters '@', ',' and '=' occurring in hyperlinks. The ANFIS outputs a value between 0.0 and 1.0, with values higher than 0.8 being categorised as phishing, values between 0.2 and 0.8 categorised as suspicious, and values lower than 0.2 categorised as genuine. In the authors' test set of 100 emails, 50 of which were phishing and 50 genuine, 96 were correctly identified as either phishing or genuine. The remaining four were categorised as suspicious. The authors assigned an error of 0.5% to the suspicious category, resulting in

a 98% detection rate. When comparing with other proposed anti-phishing techniques, they found that their technique had the best results. (Singh, Sarje & Misra, 2012.)

Shyni, Sundar and Ebby (2018) proposed using parse tree validation to detect whether a website is legitimate or not. In this proposed technique, a tree is formed by using the domain name of the website as the root, and the first ten links on the website are used as leaves of the first level. For the next level, the first ten links from one of the previous ten links are used to form the next level, and this is then repeated several times. To parse the tree, the authors use the Depth-First Search algorithm to check whether any child nodes have the same value as the root node. If the occurrence of repetitions is high, there is a high probability of the site being legitimate, and vice versa. They also use text matching to detect phishing websites. If the text of a suspicious website matches that of a known legitimate website, but the domain name is different, the website in question is a phishing website. (Shyni, Sundar & Ebby, 2018.)

Shyni et al. (2018) applied a URL verification method on a thousand legitimate and a thousand phishing websites. Out of these, 123 phishing pages were falsely identified as legitimate, and 143 legitimate pages as phishing. They used these falsely identified pages as input for their parse tree validation method and came out with a false negative rate of 7.3% (15) and a false positive rate of 5.2% (8).

In 2011, Huang, Ma and Chen proposed an authentication service where users receive generated one-time passwords from a secondary channel as opposed to having a single user-defined password. By relying on one-time passwords, users receive three levels of protection: the phisher needs to know the user's account name, the identity of the secondary channel and the password to access that secondary channel. When a user would register for a new account, they would pick an account name, but instead of picking a password, they would tie an IM account to the registration, to act as the secondary channel. When the user attempts to log in, they receive an authentication message with the one-time password via the IM client. An IM bot handles the sending of authentication messages to users. Huang et al. note that the IM provider may be able to understand what is being sent, which may be undesirable. To get around this, they recommend using an IM client with plugin support. A customised plugin may then be used to interpret obfuscated messages for improved security. The authors note that this authentication service could still be vulnerable to cross-site script attacks that steal the user's cookies, as many websites allow the users to maintain their online status using cookies, bypassing the authentication service entirely. (Huang, Ma &Chen, 2011.)

In 2008, Chow, Gustave and Vinokurov proposed using authenticated names to identify companies to avoid phishing attacks. Modelled after trademark systems, their scheme uses a local registry to identify companies according to the local registries the user chooses to use. According to the authors, the local registries would contain companies that the user deals with, such as stores and services local to them geographically. This scheme would deal with phishing sites posing as existing sites, as they would not be authenticated. However, they point out that the scheme could suffer from the same problems that trademarks suffer from, such as impersonation and conflicting names. Internet shopping at unknown global companies would also require the user to import that company's certificate before shopping. Additionally, the authors foresee difficulty knowing how to find the correct remote registries to avoid trusting rogue registries. (Chow, Gustave & Vinokurov, 2008.)

5. Discussion & implications

The purpose of this study was to review previous literature on phishing, specifically in email and instant messaging, and to present the information as a comprehensive overview. One key point was to identify common signs and techniques used by phishers. These findings can be found summarized in Table 2 for email, in Table 3 for IM and Table 4 for the phishing sites that the phishers may link to.

Phishing attacks both in email and instant messaging aim to direct the victim to a malicious website or to download attached malware. As a result, there is some overlap in the techniques used between the two approaches to attacks as well as the sites themselves. In both attack types, fake domains are used. One notable difference is the usage of anchor text to hide the suspicious fake domain in emails, whereas due to its unavailability, IM attacks utilize URL shorteners and typosquatting techniques instead.

Table 2. Common phishing signs and techniques in email

Phishing sign or technique	Explanation	Reference
Strange senders	The sender of the email may differ from whom the sender claims to be. For example, *security-bankofamerica@gmail.com* versus *security@bankofamerica.com*.	Jakobsson & Myers 2006
Email spoofing	The sender's email address may be spoofed to appear to be from a legitimate party.	Jakobsson & Myers 2006
Calls to action	A phishing email usually has an urgent call to action. For example, the email might say that the user must deny a payment within 24 hours or the money is taken from their account, prompting them to deny it from the link in the email.	Irani, Webb, Giffin & Pu 2008
Hiding a phishing URL behind anchor text	A phishing email may contain a link to a phishing site that is obscured by anchor text. For example, the anchor text *google.com* may hide the URL *phishingsite.net*.	Irani, Webb, Giffin & Pu 2008
Fake domain names	The domain name of a phishing website is usually close to the one of a real website. For example, *picture-facebook.com* versus *facebook.com*. Additionally, typosquatting techniques and homoglyphs may be used.	Bose & Leung 2007; Moore & Clayton 2015; Hannay & Bolan 2009
Spelling mistakes	Spelling mistakes are common in phishing emails. Look for anything that may be spelt wrong, such as the slogan of the company, for example.	Bose & Leung 2007
Malicious attachments	A phisher might include an attachment that contains malware in the phishing email.	Datta & Wang 2005
Copying the look of the original email	A phishing email will try to copy the look of the legitimate entity they are posing as.	Pan & Ding 2006

Table 3. Common phishing signs and techniques in instant messaging

Phishing sign or technique	Explanation	Reference
Propagation through worms	Phishing in instant messaging often propagates via worms that spam the phishing message to the friends of the infected victim. As the message comes from a friend, the victim is more likely to have their guard down.	Moore & Clayton 2015
Lures	The lure in a phishing message is often either language neutral (to target as many victims as possible) or in the target's native language (to improve chances that the target falls for the phish). A message with a simple expression like "haha" followed by a link may be a sign of a phishing attempt.	Moore & Clayton 2015; Umawing 2014, 2015, 2016
Fake domain names and URL shorteners,	As IM clients rarely give the users as much freedom with their markup languages as emails do, the phishers may not be able to disguise their links with anchor text. As a result, they make the domain similar to that of the real website, such as *picture-facebook.com* in a Facebook phish, or they hide the actual domain name by using a URL shortener, such as *bit.ly*. If the IM client does not display homoglyphs as Punycode, such URLs may be used by phishers as well.	Moore & Clayton 2015; Hannay & Bolan 2009
Chatbots	Chatbots may be employed to spam users of IM clients with phishing messages. If a new friend immediately sends you a link, there is a chance that they are a phishing chatbot.	McIntire, McIntire & Havig 2010; Umawing 2014, 2016
Executable files posing as pictures	Phishers may try to get victims to download malicious .exe files under the guise of pictures.	Umawing 2014, 2016

The use of anchor text or URL shorteners to hide the fake domain names seems to bank on the victims not paying attention to the address bar. This isn't an outlandish concept, as proven by a study from Dhamija et al. (2006), where almost one in four participants ignored parts other than the content of the web page. Both attack types also utilize a lure to get the victim to open the contents of the phishing email or message. In email, the lure is the entirety of the email that is sent to the victim, created mimicking a legitimate email. However, in IM, the lure can just be a simple sentence, something to convince the victim to click the link or to download the file. The lures work best when they're in the recipient's native language. I would get quite suspicious if my Finnish speaking friend sent me a message in a language other than Finnish, or if a stranger sent me a message in a language I wouldn't understand, so this makes sense. Another overlap is that both attacks can also be automated, whether it be by using bots or a worm to automatically propagate the phish.

Table 4. Common phishing signs and techniques in phishing websites

Phishing sign or technique	Explanation	Reference
Copying the look of the original website	Phishers will try to make the website look like the site they are posing as to make them seem legitimate.	Pan & Ding 2006
Typosquatting, fake domain names	The domain name of a phishing website will attempt to look very close to the original to fool the victim.	Moore & Clayton 2015; Hannay & Bolan 2009
Anchors that go nowhere or to a different site	The phishers may not have bothered to make more pages for the site, and as such, anchors on the site may point to the current site, no site at all, or to the real site that the phishing site is copying.	Pan & Ding 2006
SSL certificates	A missing certificate may be a sign of a phishing site but should not be relied upon. Phishers can get SSL certificates for their phishing sites.	Pan & Ding 2006
Web SSO phishing	Phishers may use fake Web Single Sign-On login windows. Always make sure that you're on the verified Web SSO login page when logging in, rather than a fake window from a phishing site.	Zhao et al. 2017

The phishing websites want to look legitimate, much like phishing emails. However, much like in IM, the websites cannot hide their domain name in the address bar with anchor text and must therefore use typosquatting or otherwise similar looking names to fool their victims.

During my research, I noticed that there were very few studies that focused on phishing in instant messaging. A lot of the focus seems to be on email and phishing websites. With how popular the digital distribution platform Steam and platforms like it have become, some more research could be done on phishing on their IM features. With Steam marketplace, individual Steam accounts can have inventories that are worth thousands, making them ideal targets for phishers, as shown by Umawing and Boyd.

The statistics from empirical phishing studies also seem to highlight the need for improved phishing awareness. This thesis should be useful for those looking for an overview of common phishing techniques and signs of phishing to better understand the subject and to protect themselves. Embedded phishing training, as shown by Kumaraguru et al. (2007), could be used to effectively train the employees of an organisation against phishing attacks.

6. Conclusion

This thesis has covered existing literature on phishing, with a primary focus on email and instant messaging. As a result, an overview of common phishing techniques and signs of phishing in email, IM, and the phishing sites themselves, has been presented. The thesis has also provided a look into various proposed counter-phishing technologies.

The literature on email is quite comprehensive, and so is the list of identified signs and techniques. One such sign is strange senders, such as a sender whose address is different from those whom they claim to be, whether it is by making the address look close to that of the original or something completely random. A technique enabled by the Simple Mail Transfer Protocol is email spoofing, where the sender can spoof their address to be that of a legitimate party's. It is also identified that phishing emails contain calls to action in order to lure the victim to click on their malicious links or to download their attached malware. Phishers may use anchor text to hide a malicious link behind what seems like a perfectly legitimate domain name. The malicious domains are often very close to that of the real website, containing only a single changed letter or potentially an added word that is hyphenated. Spelling mistakes are a common sign of a phish, though not all phishing emails contain them. Similarly, phishers may forget to add links to various images used throughout an official email, such as company logos and social media images. As a result, these elements are merely parts of the background and therefore unclickable.

Literature regarding phishing in IM is a lot scarcer. Despite this, some common techniques and signs were able to be identified. There are two main techniques for spreading phishing in IM: Worms and chatbots. Worms are capable of propagating the phishing message rapidly, as anyone who gets phished automatically sends the message forwards to all of their contacts. In addition to this, as the phishing message appears to be coming from a friend, the chances that the victim clicks on the link or attachment are higher. The second technique is using chatbots to spam users of IM clients with phishing messages. The lures in phishing messages are generally either language neutral or in the target's native language. This is because the success rate is understandably lower when victims receive a message they cannot understand. The malicious links included in phishing messages are either fake versions that look close to the legitimate domain name they are pretending to be, or URL shorteners. URL shorteners are used as IM clients rarely support markup languages that could be used to hide the URL within anchor text. In some cases, executable files posing as pictures have been used to trick victims into executing the malicious software.

Various counter-phishing technologies have been proposed in the current literature. Some of these deal with analysing whether a site is a phishing site, such as by using parse tree validation and text matching, or by analysing the links to the sites themselves using ANFIS. One solution was to use a system modelled after trademark systems, where local registries would determine whether a site is trusted or not. An authentication service using one-time passwords was also proposed. This service would use a separate IM account to send one-time passwords in place of a single, static password, meaning that the phisher would need to have access to the victim's IM account in addition to the account they are trying to access in the first place. The service would not be too dissimilar to the many multiple-factor authentication systems in use today.

The lack of studies on phishing in instant messaging has caused some limitations in the scope of the thesis. While lures and propagation were covered, no scientific literature was found on how phishers pick their targets. Because of this, future research would ideally

cover phishing in modern instant messaging. One such research subject could be the digital distribution platform Steam, which is also the world's largest online gaming platform. Steam is unique in that the users' accounts have value in items that can be sold on its marketplace, proving phishing a potentially lucrative endeavour. The platform's web API allows phishers to search for users with valuable inventories, making finding worthwhile targets effective.

References

Agten, P., Joosen, W., Piessens, F., & Nikiforakis, N. (2015, February). Seven months' worth of mistakes: A longitudinal study of typosquatting abuse. In *Proceedings of the 22nd Network and Distributed System Security Symposium (NDSS 2015)*. Internet Society.

Aycock, J. (2007). A design for an anti-spear-phishing system. Paper presented at the *Virus Bulletin Conference,* 290-293.

Bolderston, A. (2008). Writing an effective literature review. *Journal of Medical Imaging and Radiation Sciences, 39*(2), 86-92.

Bose, I., & Leung, A. C. M. (2007). Unveiling the mask of phishing: Threats, preventive measures, and responsibilities. *Communications of the Association for Information Systems, 19*(1), 24.

Boyd, C. (2014a, April 17). Phishers Bypass Steam Guard Protection [Blog Post]. Retrieved May 6, 2019, from https://blog.malwarebytes.com/cybercrime/2014/04/phishers-bypass-steam-guard-protection/

Boyd, C. (2014b, June 25). Phishy Steam Guard File Steals SSFN [Blog Post]. Retrieved May 6, 2019, from https://blog.malwarebytes.com/cybercrime/2014/06/phishy-steam-guard-file-steals-ssfn/

Chow, S., Gustave, C., & Vinokurov, D. (2008, July). Authenticated names. In *Proceedings of the 2007 Workshop on New Security Paradigms*, 3-32. ACM.

Datta, S., & Wang, H. (2005, May). The effectiveness of vaccinations on the spread of email-borne computer viruses. In *Canadian Conference on Electrical and Computer Engineering, 2005*, 219-223. IEEE.

Dhamija, R., Tygar, J. D., & Hearst, M. (2006). Why phishing works. Paper presented at the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 581-590.

Federal Trade Commission. (2004). FTC Consumer Alert: Spyware [Brochure]. Retrieved November 23, 2017, from http://www.invisus.com/news_archive/old/spyware/Spyware.htm

Hannay, P., & Bolan, C. (2009, December). Assessment of Internationalised Domain Name Homograph Attack Mitigation. In *Australian Information Security Management Conference*, 13.

Huang, C. Y., Ma, S. P., & Chen, K. T. (2011). Using one-time passwords to prevent password phishing attacks. *Journal of Network and Computer Applications*, *34*(4), 1292-1301.

Irani, D., Webb, S., Giffin, J., & Pu, C. (2008). Evolutionary study of phishing. Paper presented at the *ECrime Researchers Summit, 2008,* 1-10.

Jakobsson, M., & Myers, S. (2006). Phishing and countermeasures: Understanding the increasing problem of electronic identity theft. *Phishing and countermeasures: Understanding the increasing problem of electronic identity theft*, 1-699.

Kramer, S., & Bradfield, J. C. (2010). A general definition of malware. *Journal in Computer Virology, 6*(2), 105-114.

Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L. F., & Hong, J. (2007). Getting users to pay attention to anti-phishing education: Evaluation of retention and transfer. Paper presented at the *Proceedings of the Anti-Phishing Working Groups 2nd Annual eCrime Researchers Summit,* 70-81.

McIntire, J. P., McIntire, L. K., & Havig, P. R. (2010). Methods for chatbot detection in distributed text-based communications. Paper presented at the *2010 International Symposium on Collaborative Technologies and Systems, CTS 2010,* 463-472.

Mohan, J. C., & Kumar, R. (2017). On the Efficacy of Android Ransomware Detection Techniques: A Survey. *International Journal of Pure and Applied Mathematics,* 115(8), 115-120.

Mohebzada, J. G., El Zarka, A., BHojani, A. H., & Darwish, A. (2012). Phishing in a university community: Two large scale phishing experiments. Paper presented at the *Innovations in Information Technology (IIT), 2012 International Conference On,* 249-254.

Moore, T., & Clayton, R. (2015, May). Which malware lures work best? Measurements from a large instant messaging worm. In *2015 APWG Symposium on Electronic Crime Research (eCrime)*, 110. IEEE.

Mouton, F., Leenen, L., & Venter, H. S. (2016). Social engineering attack examples, templates and scenarios. *Computers and Security, 59*, 186-209.

Rouse, M. (2008). What is instant messaging (IM or IM-ing or AIM). TechTarget SearchUnifiedCommunications. Retrieved November 23, 2017, from http://searchunifiedcommunications.techtarget.com/definition/instant-messaging

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., & Downs, J. (2010). Who falls for phish?: A demographic analysis of phishing susceptibility and effectiveness of interventions. Paper presented at the *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 373-382.

Shyni, C. E., Sundar, A. D., & Ebby, G. E. (2018, February). Phishing Detection in Websites using Parse Tree Validation. In *2018 Recent Advances on Engineering, Technology and Computational Sciences (RAETCS),* 1-4. IEEE.

Singh, S., Sarje, A. K., & Misra, M. (2012, November). Client-side counter phishing application using adaptive neuro-fuzzy inference system. In *2012 Fourth*

International Conference on Computational Intelligence and Communication Networks, 788-792. IEEE.

Steer, J. (2017). Defending against spear-phishing. *Computer Fraud and Security, 2017*(8), 18-20.

Umawing, J. (2014, September 16). Steam Threats: What They Are and What You Can Do to Protect Your Account [Blog Post]. Retrieved May 6, 2019, from https://blog.malwarebytes.com/101/2014/09/steam-threats-what-they-are-and-what-you-can-do-to-protect-your-account/

Umawing, J. (2015, March 31). Phony My Team Voice App Being Spread on Steam Chat [Blog Post]. Retrieved May 6, 2019, from https://blog.malwarebytes.com/cybercrime/2015/03/phony-my-team-voice-app-being-spread-on-steam-chat/

Umawing, J. (2016, March 31). Latest Steam Malware Shows Signs of RAT Activity [Blog Post]. Retrieved May 6, 2019, from https://blog.malwarebytes.com/cybercrime/2016/03/latest-steam-malware-shows-sign-of-rat-activity/

Pan, Y., & Ding, X. (2006, December). Anomaly based web phishing page detection. In *2006 22nd Annual Computer Security Applications Conference (ACSAC'06)*, 381-392. IEEE.

Yi, X., & Allan, J. (2010). A content based approach for discovering missing anchor text for web search. Paper presented at the *SIGIR 2010 Proceedings - 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval,* 427-434.

Zhao, R., John, S., Karas, S., Bussell, C., Roberts, J., Six, D., Gavett, B. & Yue, C. (2017). Design and evaluation of the highly insidious extreme phishing attacks. *Computers & Security*, *70*, 634-647.