A Framework to Assist Email Users in the Identification of Phishing Attacks

A. Lötter and L. Futcher

Nelson Mandela Metropolitan University, Port Elizabeth, South Africa
email : {andries.lotter ; lynn.futcher}@nmmu.ac.za

Abstract

This paper proposes a framework to address the problem that email users are not well informed or assisted by their email clients in identifying possible phishing attacks, thereby putting their personal information at risk. Furthermore, it argues that email clients should make use of feedback mechanisms to present security related aspects to the users, so as to make them aware of the characteristics pertaining to such attacks. This paper therefore addresses the human weakness (i.e. the user's lack of knowledge of phishing attacks which causes them to fall victim to such attacks) as well as the software related issue of email clients not visually assisting and guiding the users through the user interface.

Keywords

Email client security, phishing attacks, usable security, user awareness

1. Introduction

A fact that cannot be disputed is that the Internet is an ever growing craze. Every day new users are adopting the Internet for the first time. The global Internet population (as of 2012) represented just over 2.4 billion people compared to the 360 million Internet users in late 2000 (Miniwatts Marketing Group, 2012). Along with this growth of users, the content on the Internet also expands every minute.

Unfortunately, along with any popular phenomenon comes an increase in exploitation thereof. Phishing can be seen as such, and a paper on "Social Phishing" defines phishing as: "*a form of social engineering in which an attacker attempts to fraudulently acquire sensitive information from a victim by impersonating a trustworthy third party*" (Jagatic, Johnson, Jakobsson, & Menczer, 2005). Recent statistics have found that, in the second half of 2011 alone, 83 083 unique phishing domains were registered worldwide. Other findings indicated that 3% of all phishing emails were opened, that eight victims are yielded for every 100 000 targeted users and that the average phishing victim produces around $2 000. Furthermore, 500 million phishing emails appear in user inboxes every day (Orloff, 2012). From this it is discernible that 40 000 people (worldwide) will fall victim to a given phishing attack every day, resulting in daily damages of approximately $80 million.

Phishing attacks are undoubtedly a popular way in which cyber-criminals conduct their crimes. It is argued that part of the blame for why phishing attacks are so successful could be shifted towards email clients. Email clients should therefore implement an effective and secure protection mechanism to protect email users in this regard.

This paper addresses the problem that email users are not well informed or assisted by their email clients in identifying possible phishing attacks, thereby putting their personal information at risk. In addressing this problem, this paper presents a framework to assist email clients and their users in the identification of phishing attacks. A literature study was carried out to determine the characteristics common to phishing attacks and to understand the security mechanisms currently employed in email clients. Furthermore, argumentation and modelling techniques contributed towards the development of the framework. This paper follows on from a paper published at the 2013 ZAWWW Conference (Lötter & Futcher, 2013). The said paper was a work in progress towards the development of the framework presented in this paper.

The results of the literature reviewed are presented in Sections 2 and 3 respectively. Whereas Section 4 presents the framework as a mental model that can assist users in the identification of common forms of phishing attacks, Section 5 discusses how this framework can be implemented in the email client.

2. Email client security

Anyone with an email account is a potential phishing target. Therefore, because of the great reach of phishing emails, it can be deduced that most email users may fall victim to such attacks. However, in order to realistically mitigate phishing attacks, the burden of identifying such attacks should not only lie in the software side; users also require a certain level of awareness. The email client software should therefore be designed and developed in such a way, that it "educates" the users. According to Furnell (2005, p.276), the software should at all times "*provide a visible indication of the security status*" as this is one of the primary causes that leads to the users feeling insecure about the security of their software.

Email clients do implement a reasonable amount of security. At the very least, they implement protection mechanisms such as password protection when accessing one's inbox and make use of spam filters to prevent users from coming into contact with unsolicited email messages. The problem here lies in the fact that these spam filters are not 100% accurate (Spamhaus, 2010). Sometimes legitimate messages get flagged as spam and fraudulent messages pass through the filters. It is at this stage that the user needs to be sufficiently aware of the criteria for identifying fraudulent email, so that they do not fall victim to a potential attack.

Currently, email clients simply place any email message it deems sufficiently suspicious into a "Junk" folder. Thus, it is left to the user's imagination to discern why a certain message was flagged as "Junk". There is no feedback mechanism to identify the portions of the email that caused the email client to believe that the said message is fraudulent. Even when users peruse their "Junk" folder, they may find emails in the said folder that they know does not belong. Often they are left puzzled at the email client's inability to have foreseen that certain messages were in fact genuine. The user interface of an email client should therefore be designed in such a way that it provides feedback to the user. All received email messages should be represented (in a minimalistic manner) according to its level of suspicion. Security dialogs should not be verbose and tedious as to deter the user from learning; however, compact and to-the-point explanations should be available as per the user's request. Therefore, the next time a potential phishing attack bypasses the spam filters, the user should be aware of the criteria to look out for when identifying potential fraudulent email. Thus, the risk that a user will fall victim to a specific phishing attack is further mitigated.

There exist vulnerabilities in email clients which phishers exploit in order for their phishing attacks to succeed. It is thus these vulnerabilities that need to be managed in order to mitigate phishing attacks. What causes a phishing attack to succeed is a combination of the software (email client) that was unable to flag the email as a phishing attack, and the user's gullibility in believing that the email is genuine. A paper on "Why users cannot use security" by (Furnell, 2005, pp. 274-279) states that "*Some clear awareness issues still need to be overcome, and there is unfortunately ample evidence to show that users do not actually understand security very well in the first place"*. From this it is clear that the usable security aspect of email clients must be addressed, as it should be a goal of the email client to prevent users from coming into contact with fraudulent email. It is argued that phishing attacks will only be successfully mitigated, once the average email user has the knowledge to differentiate a legitimate email from its fraudulent counterpart. The user interface in email clients should therefore implement security mechanisms that address the manner in which users perceive and understand security.

3. Understanding Phishing attacks

Phishing can be seen as a type of online identity theft. It is usually conducted by means of sending email messages to (thousands of) potential victims (Ayodele et al., 2012, p. 208). These emails are typically sent out in bulk to act as "bait", claiming to be from individuals or companies that the receiver of the message may trust, asking for confidential and sensitive information. The content of these emails are thus designed to deceive the receiver into divulging their personal details. These details can then be used by the phisher to gain access to the victim's financial accounts.

A variation of this attack, which encompasses much of the same deception techniques, but functions slightly differently, is known as "spear phishing". In a paper specifically focussing on spear phishing, Wang et al. (2012, p. 345) describes spear phishing as being more content specific in comparison to normal phishing attacks. They further explain that spear phishing attacks are perceived to originate from an existing organisation, thereby establishing the sender of the attack as relevant and true. A common use among phishers is to impersonate well-known financial institutions like banks (Chen & Guo, 2006). Spear phishing is effective, because it functions on the statistical fact that a large percentage of the targeted population will have an account with a company with a huge market share. Therefore, spear phishing attacks appear to come from an organisation that the targeted user could possibly have an account with. Phishers can therefore employ this technique by looking up the Chief Executive Officer (CEO) of a company on its website, and send emails to the accounts in the same corporate domain, seemingly from the CEO (Janssen, n.d.).

From the literature studied (Ledford, n.d.; Wang et al., 2012), several characteristics have been identified that can indicate the likelihood of an email message being a potential phishing attack. These characteristics include:

1. **Urgent wording in message:** Phishing attacks in general stress the urgency of the email as to make the victim uneasy and get results quickly.
2. **Request for personal and sensitive information**: Phishing attacks, by definition, aim to deceive victims into trusting the phishers, thereby gaining access to the victim's personal details with which to commit identify theft.
3. **Sender is unknown:** However, spear phishing is, by definition, a more concentrated attack. The phisher often impersonates a co-worker or executive member in the same corporate domain.
4. **Fake (deceiving) hyperlinks embedded**: The hyperlinks usually point to a phishing domain.
5. **Message body is an image:** Spear phishing, on the other hand, is more text-based, and would not necessarily use this evasive technique.
6. **Unrealistic promises:** Although spear phishing does not contain empty promises. They are to the point, to retain credibility.
7. **Poor language and punctuation:** Phishing attacks in general tend to be badly constructed.
8. **Visually represents impersonation:** As mentioned, spear phishing is more text-based, because it "comes from a co-worker" or trusted entity.
9. **Contains malware as attachments:** Generally phishing may try to install malware upon opening attachments.
10. **Emails are sent out at random to large number of random email addresses:** Spear phishing attacks, however, are concentrated, thus the victims are chosen carefully.

Phishing attacks undoubtedly pose a noteworthy problem. It is therefore important to understand the characteristics of these attacks in order to identify them. These characteristics are fundamental to the framework presented in the following section.

4. The Framework as a Mental Model

The framework presented in this section has been developed to simulate the thought process of the user of an email client when determining the legitimacy of a specific email. However, it can easily be adapted to be implemented into email clients (the software) as discussed in Section 5.

The framework depicted in Figure 1 illustrates a sequence of nine steps that the user of an email client should ask him or herself when determining whether an email should be trusted or not. The framework acts as a flowchart in that it guides the user through all nine steps. Only by answering "no" to each question (except for the last), can the positive outcome of "Email should be safe" be reached. The questions posed were determined based on the common characteristics of phishing attacks as described in Section 3.

The questions in this framework have been ordered to range from highly significant to less significant. Thus, a "Yes" answer to the former questions could lead to a higher probability of the email in question being fraudulent. The reason for this particular order is because this framework imitates the thought process of the human mind. Therefore, the most significant characteristics of a phishing attack are considered first. Upon finding that a certain characteristic is present, the framework opts out and classifies the email as a likely phishing attack without considering the other (less significant) characteristics.

This framework can classify a given email in four different ways. If an email contains a highly significant characteristic, it can either be classified as having a high or medium risk of being a phishing attack. Similarly, if the email contains a less significant characteristic, it can be classified as having a low risk of being a phishing attack or as cautious. The cautious classification serves as an intermediate between low risk and medium risk. When an email is classified as such, it advises the user that they should have an elevated sense of caution, since some less significant phishing characteristics are present.

A characteristic that is often present in phishing attacks is the abundance of spelling and grammar errors. However, an email should not be deemed a phishing attack based solely on the presence of such mistakes. One needs to consider that a specific phishing attack may be so meticulously thought out and refined, that it does not contain any spelling and grammar errors. Similarly, a normal, everyday email from one peer to another is often full of spelling and grammar errors, since emails often tend to be sent out in haste. For these reasons, the question asking whether spelling and grammar errors are present is considered with each of the other questions posed. If an email is already deemed suspicious and the email also contains many spelling and grammar errors, the likelihood (risk) of the said email being a phishing attack is increased. Otherwise, if suspicion is never raised about the legitimacy of an email, the spelling and grammar characteristic is never brought into consideration.

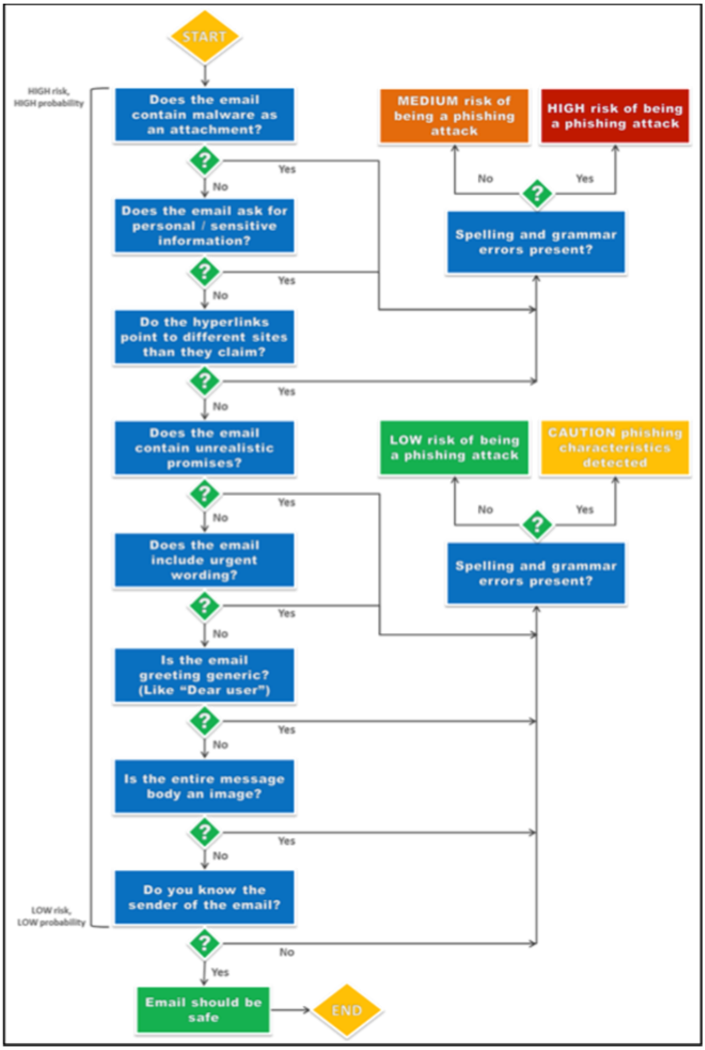

START
HIGH risk, HIGH probability
Does the email contain malware as an attachment?
Yes
No
Does the email ask for personal / sensitive information?
Yes
No
Do the hyperlinks point to different sites than they claim?
Yes
No
Does the email contain unrealistic promises?
Yes
No
Does the email include urgent wording?
Yes
No
Is the email greeting generic? (Like "Dear user")
Yes
No
Is the entire message body an image?
Yes
No
LOW risk, LOW probability
Do you know the sender of the email?
No
Yes
Email should be safe
END
MEDIUM risk of being a phishing attack
HIGH risk of being a phishing attack
No
Yes
Spelling and grammar errors present?
LOW risk of being a phishing attack
CAUTION phishing characteristics detected
No
Yes
Spelling and grammar errors present?

Figure 1: A framework to identify phishing attacks (mental model)

The terminating question, "Do you know the sender of the email?" can be somewhat questioned for phishing emails seldom impersonates a person. Recall that phishing is a "form of social engineering in which an attacker attempts to fraudulently acquire sensitive information from a victim by impersonating a trustworthy third party" (Jagatic et al., 2005). This "trustworthy third party" could thus refer to either a person or a company. Therefore, by answering this question, the user needs to consider all types of phishing attacks. They should thus consider whether they know the company that may have sent them the email. Does it make logical sense for this company to have contacted them (i.e. do they have a connection to this company)? In the case that the sender is a person, they should consider whether this person has merit in contacting them.

Phishing attacks normally visually represent the organisation or company it is trying to impersonate. From a human and software standpoint, it is virtually impossible to identify an email as a phishing attack based on the fact that it *looks like* a legitimate email originating from an organisation. Normally, one would just assume that it is in fact an email from the said organisation. It is thus in combination with the other characteristics – after realizing the email is fraudulent – that one can see how the organisation has been visually impersonated, by means of incorporating a lot of their logos and images. For this reason, this characteristic is not considered in the framework.

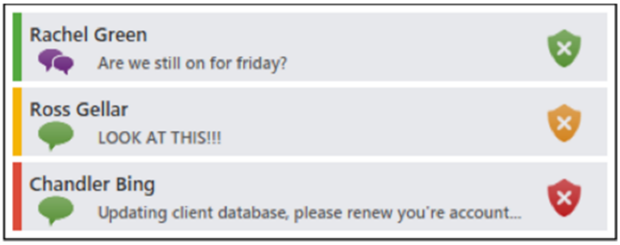
Phishing attacks are normally sent out in bulk to a large number of users. This characteristic, despite not rigidly appearing in Figure 1, has been adapted into "Is the email greeting generic? (Like 'Dear user')". This adaptation seems befitting since an email that is sent out in bulk usually does not address each recipient by name, and therefore makes use of generic greeting lines. Furthermore, phishers normally do not have the recipient's real name because of the manner in which the email addresses are obtained. Therefore, it is logical to deduce that an email may be a potential phishing attack were it to address the recipient in a generic manner.

Lastly, the termination points to this framework make use of "indefinite" statements, such as "Email *should* be safe" or "…*risk* of being a phishing attack". The reason for this is that one can never be completely certain that a specific email poses no security threat whatsoever. An email from a friend may contain an attachment that (unbeknownst to both the sender and receiver) contains a virus. Similarly, a user's email account could have been compromised and is being used to send out malicious emails to all its trusted contacts. For these reasons, one should always consider that an email may still be potentially dangerous, even if all signs point to the contrary.

5. The Framework as a Software Tool

Email clients make use of various techniques in filtering out spam messages, such as rule-based and Bayesian spam filtering. The main purpose of the proposed framework developed is not to improve the existing filtering techniques, but rather to improve the way in which any irregularities present in an email is reported back to the user. Thus, from a software standpoint, the framework can be implemented in the user interface of email clients so as to increase the awareness of users with regard to phishing attacks.

Figure 2 illustrates how the security level of email messages (as they would appear in the inbox) can be conveyed to the user in a minimalistic manner. As seen in this figure, the email items are all associated with a specific colour (as seen by the leftmost border and the rightmost shield icon).



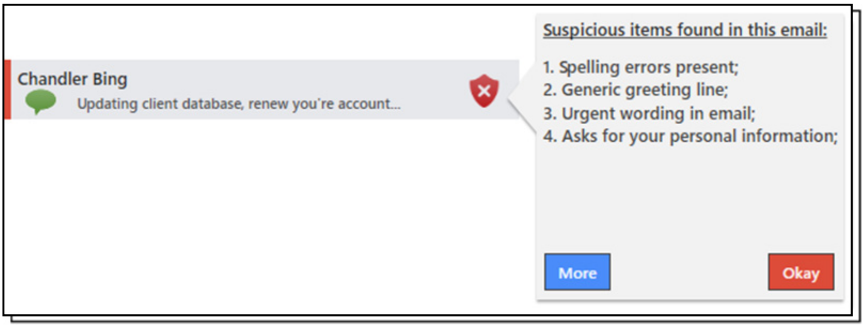
Rachel Green
Are we still on for friday?

Ross Gellar
LOOK AT THIS!!!

Chandler Bing
Updating client database, please renew you're account...

Figure 2: Indicating security level of received emails in a minimalistic manner

These colours (much like traffic lights) instinctively conveys to the users whether an email is considered safe or not, without them having to read a single word. Logically, green would represent a message which is considered safe, orange would indicate that there is some doubt regarding the safety of the message, and red would indicate that the message is most likely a phishing attack. Should the user like to know why an email is considered safe, doubtful or dangerous respectively, they can find the information by clicking on the shield icon. Figure 3 depicts how the information could be presented to the user by means of a context menu.



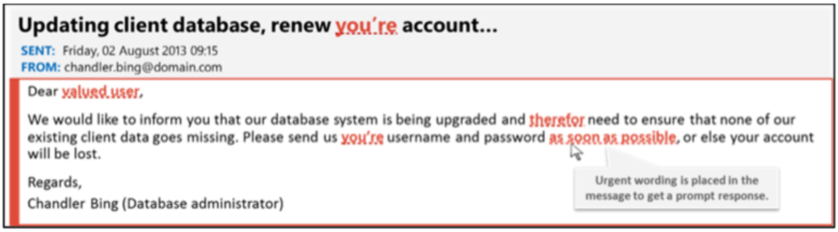
Chandler Bing
Updating client database, renew you're account...

Suspicious items found in this email:

1. Spelling errors present;
2. Generic greeting line;
3. Urgent wording in email;
4. Asks for your personal information;

Figure 3: Additional phishing characteristics identified displayed in a context menu

As can be seen in Figure 3, the user is now presented with a list of suspicious characteristics identified by the framework. Thus, security is placed at the forefront of the user interface. The user does not have to read tedious security dialogs full of jargon and terminology which they do not understand. Users are often not motivated to use security, because of jargon and terminology which they do not understand. As mentioned above, Figure 3 shows the suspicious aspects of a specific email in short, easy to understand terms thus appealing to the user's sense of simplicity. However, detailed explanations should also be provided per the user's request. Figure 4 shows this detailed explanation which can be accessed by the user upon clicking on the "*more*" button seen in Figure 3.



Updating client database, renew you're account...

SENT: Friday, 02 August 2013 09:15
FROM: chandler.bing@domain.com

Dear valued user,

We would like to inform you that our database system is being upgraded and therefor need to ensure that none of our existing client data goes missing. Please send us you're username and password as soon as possible, or else your account will be lost.

Regards,
Chandler Bing (Database administrator)

Urgent wording is placed in the message to get a prompt response.

Figure 4: Detailed explanation of the aspects identified in the suspicious email

As evident in Figure 4, the entire email message is displayed with all the suspicious aspects identified by the framework shaded in red and underlined. The message border is also red, so as to keep displaying the security level to the user. When the user hovers over one of the suspicious aspects, a tooltip is displayed describing the characteristic that was found. Thus, the email is no longer simply placed in a "*junk*" folder without explanation. Through this method, and the ones described previously in this section, the users can be made aware of the characteristics pertaining to phishing attacks.

All of the figures discussed in this section (Figures 2 to 4) rely on the framework developed in order to determine how the user interface of the email client needs to adapt to the security level of a specific email. The email client software should work procedurally through the sequence of questions to see which characteristics are present in the email. If a certain characteristic is found, it should increase the probability of the said email being a phishing attack based on a pre-determined weighting. It is important to note that the weightings for each characteristic should not be equal. An email does not deserve the same penalty for including spelling and grammar errors, than should it contain malware as an attachment. Afterwards, the framework should be followed in moving on to the next question in the sequence and will follow this paradigm until all the characteristics in the framework have been considered. This results in a final score, which is the probability of the email being fraudulent, being presented as output. The user interface of the email client can then be adjusted accordingly based on this score, i.e. the email messages in the inbox can be colour coded as seen in Figure 2.

The colour code that a certain email should be associated with (green, orange or red) can be determined by the probability score. The email client implementing the framework can make it a business decision as to what the ranges are for safe (green), doubtful (orange) and dangerous (red) classifications. It should be noted that an email displayed in green can still have items in its context menu (should the user wish to see it). Figure 5 illustrates a gauge that can be used to determine the ranges for these classifications. As can be seen in this figure, if the resultant probability score is lower than 0.1, it can be deemed as safe. If the score ranges between 0.1 and 0.49, the email may be deemed doubtful. Lastly, if the score is higher than 0.5, the email is deemed dangerous and a potential phishing attack.

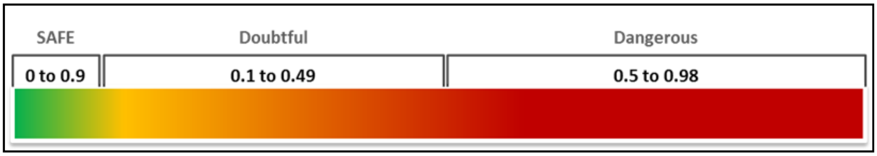


Figure 5: A colour gauge indicating the security level of emails

As stated, determining these ranges can be made a business decision by the email client implementing the framework. Moreover, the email client may even allow the user to define these ranges. As guidance, the email client may have certain default values for these ranges (like the ones specified in Figure 5), but then allow more paranoid or trusting users to redefine these ranges to a level that they are comfortable with.

6. Conclusion

This paper presents a framework that specifically addresses the threat of phishing attacks to email users and is based on the common characteristics found in phishing attacks. Although it was initially developed to be used as a mental model by email users, it can easily be adapted for implementation in email clients. The users of email clients should have a visual indication of security status at all times. Only through user awareness can scams like phishing be successfully mitigated. Through implementation of this framework the user's level of awareness can be raised by presenting to them the aspects identified as being suspicious. Users will therefore be made more aware of the characteristics pertaining to phishing attacks and in so doing this threat could be mitigated.

Future research is required to address other security threats relating to email users in order to ensure that email clients cater for all aspects of security that put email users and their information at risk.

7. References

Ayodele, T., Shoniregun, C., & Akmayeva, G. (2012). Anti-Phishing Prevention Measure for Email Systems. *Internet Security (WorldCIS)*, (pp. 208-211). Guelph.

Chen, J., & Guo, C. (2006). Online Detection and Prevention of Phishing Attacks. *Communications and Networking in China, 2006. ChinaCom '06. First International Conference*, (pp. 1-7). Beijing.

Furnell, S. (2005). Why users cannot use security. *Computers & Security* (24), 274-279.

Jagatic, T., Johnson, N., Jakobsson, M., & Menczer, F. (2005, December 15). Social Phishing. Bloomington.

Janssen, C. (n.d.). *Spear Phishing*. Retrieved April 29, 2013 from Techopedia: http://www.techopedia.com/definition/4121/spear-phishing

Spamhaus. (2010, January). *Whitepapers: Effective filtering*. Retrieved July 16, 2013 from Spamhaus: http://www.spamhaus.org/whitepapers/effective_filtering/