Phishing – challenges and solutions

Article *in* Computer Fraud & Security · January 2018

DOI: 10.1016/S1361-3723(18)30007-1

2 authors, including:

Sathish A.P Kumar
Cleveland State University
71 PUBLICATIONS **984** CITATIONS

SEE PROFILE

Phishing – challenges and solutions

Ike Vayansky and Sathish Kumar, Coastal Carolina University


Ike Vayansky


Sathish Kumar

Phishing is a major threat to all Internet users and is difficult to trace or defend against since it does not present itself as obviously malicious in nature. In today's society, everything is put online and the safety of personal credentials is at risk. Phishing can be seen as one of the oldest and easiest ways of stealing information from people and it is used for obtaining a wide range of personal details. It also has a fairly simple approach – send an email, email sends victim to a site, site steals information.

In reality, phishing has become a complex and escalating threat to everyone's Internet security. By gathering even a small amount of information about a victim, the attacker can produce a personalised and believable email. These phishers are not easy to catch either, as most of them can hide the location of their servers and work in almost complete anonymity. Even a user with excellent security software can fall victim to a phishing attack, because for the most part they depend entirely on information typed into a form, not malware infection of a computer.

However there are many ways to protect against phishing attempts. In this article we'll discuss three different approaches, each at varying stages of the attack. An attack can be detected before it reaches the user, once the user has reached the phishing site, or the user can be trained to be aware of the attack. Using these approaches together, a secure and safe environment for a user can be created. We'll also examine the current methods phishing uses to deceive its victims and how much it has evolved over the years.

Phishing has become so advanced and stealthy that, according to Forbes, it results in about $500m in losses per year to US businesses alone.[1] And the advances in targeted spear-phishing attacks – and how easily they can find data just by searching publicly available sources – are astonishing.

Background

Phishing is one of the most organised crimes of the 21st century. It is defined as a type of malware or a term for where someone sends out a spoofed email to random victims to try to get personal information about them. More specifically in computing, phishing is a criminal activity using social engineering techniques to fraudulently acquire sensitive information such as usernames and passwords by attempting to trick users of popular websites by emailing them fake versions of the website to provide their credentials to.

This may seem easy to avoid but the advances in the phishing community are making phishing scams harder and harder to identify from the victims' standpoint. The term phishing has evolved from being almost a poorly constructed instant messaging attack into spoofing entire websites to fool users into providing personal information. An example of a phishing attempt made to some of the email users at the authors' institution in 2017 is shown in Figure 1.

For this article, we surveyed the literature to study the current state of phishing and existing solutions. To address the many new developments in phishing, such as spear-phishing, pharming and social phishing, and the way that phishers are also developing more and more convincing sites and emails to deceive users, we have designed a three-step approach to prevent and control phishing. Based on the proposed phishing solution framework, the attack can be stopped before it reaches the user, once the user is at the phishing site, or by training users to avoid it by themselves.

The objective of this article is to provide a clear analysis of the current state of phishing and recommend practical solutions. The following sections provide a description of phishing in general, the history of phishing and common problems associated with the practice.

"Pharming leverages malicious code such as viruses, worms, trojans and spyware to carry out sophisticated attacks including host file modification, DNS cache poisoning and so on"

Phishing has been a major issue for security for a long time without a good solution in place. The problem with phishing is that a holistic solution that works to protect users securely from being phished does not exist. As the defences against phishing have evolved, so have the current phishing methods. As a result, the need for more advanced methods of security to identify phishing scams is important.

Literature review

'Phishing and Pharming – The Deadly Duo': In this paper, the author covers the increasing trends in phishing and how attacks have evolved over time.[2] The goals of most phishing attacks are identified as broad or personalised. A


**********From: **********
**********Sent: **********
**********Subject: RE: [] ****** Sign-in Alert
**********To: undisclosed-recipient:
**********Subject: Sign-in Alert
********** Sign-in Alert
**********Dear Office365 Email User,
**********We noticed a login to your Microsoft account from an unrecognized device on 9:37 AM **
**********Was this you? If so, please disregard the rest of this email.
**********If this wasn't you, please follow the links below to keep your E-Mail account safe and
**********provide required information to keep your account ACTIVE.
**********CLICK HERE
**********Thanks,
********** Admin Services
**********©2017 All rights reserved.

Figure 1. An example of a phishing attempt by email.

breakdown of the phishing process is described in a stage-by-stage process that helps to create the concept of a typical phishing attack. We can visualise the damaging effects from phishing on victims using three different angles: enterprise, customer and government authority. Another malicious type of mass phishing is pharming. The paper looks at the pharming attack process and how it differs from phishing. Pharming leverages malicious code such as viruses, worms, trojans and spyware to carry out sophisticated attacks including host file modification, DNS cache poisoning and so on, and the user will not be aware of it. There are few solutions that are proposed to prevent phishing attacks.

'Online Frauds in Banks with Phishing': This presents a more detailed look into the implications of phishing frauds in online banking. This study explains the most common frauds with online banks and how these are associated with phishing. Many different definitions that can be used for phishing are outlined here. Further, this work outlines the various phishing techniques that attackers may use. This paper describes the reasons for the increased prevalence of phishing attacks and outlines a few examples of actual organisations affected by these malicious actions. It provides a detailed action plan on how to combat bank fraud by phishing. Various tables of data include the top hosting methods for phishing sites and reports of attacks increasing over the years.

'The Current State of Phishing Attacks': In this work, the author analyses the state of phishing attacks.[4] The paper describes how attacks will trick victims with multiple types of malware. The anatomy of a phishing attack is explained through the concepts of fake phishing emails, setting up fake websites and monetising stolen information. The psychology behind why phishing attacks work is explained briefly. Phishing causes damage to organisations and costs lots of money every year. Countermeasure recommendations are covered, with different approaches to keep someone from becoming a victim of phishing scams.

'Classification of Phishing Email Using Random Forest Machine Learning Technique': The primary focus of this research work is the application of machine learning to identify phishing emails.[5] First, it introduces the concept of phishing and problems that are associated with it. The concept of machine learning is also described for its use in discovering phishing emails. Most email filtering methods have not evolved with the phishing techniques which is why machine learning for discovering patterns in phishing emails is important. The classification of a phishing email used in the machine learning detection system is described based on a set of rules. It approaches phishing using an experiment based on an algorithm for detecting phishing emails. The results were encouraging for this technique, with few false positives.

'Spear-phishing: how to spot and mitigate the menace': Spear-phishing can be defined as the preliminary stage of an advanced persistent threat (APT) attack, to create a point of entry into the organisation.[6] This article is mainly focused around the concept of spear-phishing and how it works differently from just generic phishing. It covers how phishing affects victims and what is gained by the attacker using these methods. A brief description of how to avoid spear-phishing is also provided.

'Social Phishing': In this paper, the authors define phishing in terms of social network approaches.[7] The many ways that a victim's personal information can be mined and exploited using data that can easily be found online are discussed. An attacker can use various social networking sites to produce a focused and much more effective phishing tactic. The paper describes research that was carried out by browsing and documenting relationships freely available to the public on such sites and using that information to launch a mock phishing attack on their subjects. During this procedure, the authors describe the methods social phishing attacks employ to steal user credentials. Finally the paper discusses the results of the experiment and the demographics of the subjects that fell for the attacks.

'Protecting People from Phishing: the design and evaluation of an embedded training email system': This work discusses a possible form of phishing prevention in an email system format.[8] Instead of focusing on an automated web browser system or algorithms that detect phishing attacks, this research was focused on the approach from a user interface standpoint. The authors have designed an embedded training system into their email servers to train users how to identify and protect themselves from phishing attempts through the use of basic reasoning and simple tips. In the study, the users of the email server were sent false phishing emails with embedded links that if clicked on would send them to a page

demonstrating that the link was not trustworthy. It would also provide an explanation via a comic, which described the best way to identify and avoid phishing emails. The result of the system demonstrated that the system was more effective at reducing successful phishing attacks than the basic awareness emails.

'TrustBar: protecting (even naïve) web users from spoofing and phishing attacks': This paper describes yet another method of defending against phishing attempts through the application of a user interface system.[9] The system proposed by the authors is designed to help users who are not familiar with computers and current anti-phishing protections. First they present the security principles that a user interface system should follow and briefly go over similar projects that have been proposed by other researchers. Current server authentication using the Secure Socket Layer (SSL) and Transport Layer Security (TSL) is described, along with its shortfalls. The paper also describes the nature of phishing and spoofing (a key component of a phishing attack) and how they exploit the vulnerabilities of the SSL/TLS protocol as used by web browsers. The criteria that the design follows in order to prevent spoofing is listed with the user in mind. Finally, the authors present their system for identifying protected and trusted sites in a clear and visible manner.

'Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions': The experiment outlined in this paper demonstrates which users appear to be most susceptible to a phishing scam.[10] The experiment had a large test group take the role of students at a fictitious university and gave them an email log to look through and determine what action they would take with each from a given list of actions. The composition of the groups by gender, education, whether they originated from the US, if they were a student, the average years of experience on the Internet, and the average emails per day was documented for each experimental group. The groups had varying levels of anti-phishing training using pre-existing programmes. They then discussed the results of the experiment in relation to key demographics and the effect that the different forms of training had on susceptibility.

"Phishers have become more skilled at forging websites to appear identical to the expected location, even including logos and graphics in the phishing emails to make them more convincing"

'Modelling and Preventing Phishing Attacks': In this paper, the author has provided a series of visual aids for understanding how phishing attacks are carried out.[11] By use of a graph-based model, the various components and factors of an attack are represented. The way that these factors are represented is also explained by the authors. An exemplar phishing scenario is described and then visualised using these models. The different ways that one attack can be carried out given the attacker's knowledge of the victim is shown. Another, more advanced, form of phishing known as 'context aware phishing' is defined. In order to provide a better understanding of how this style of attack works, examples are provided and modelled as well. The different methods of victim selection and data collection and linking are noted. Finally, the paper concludes with an analysis of the example attacks described, and possible defences against such attacks.

Problem and challenges

The problem with phishing is that attackers constantly look for new and creative ways to fool users into believing their actions involve a legitimate website or email. Phishers have become more skilled at forging websites to appear identical to the expected location, even including logos and graphics in the phishing emails to make them more convincing.

There are dangerous new advanced phishing methods that utilise personal information that is easily available to the public in order to produce plausible and believable attacks that directly target victims. Methods such as social phishing and context aware phishing are perfect examples of attacks utilising the massive amount of public information to increase the effectiveness of their scams. One study shows that victims are 4.5 times more likely to fall for a phishing attempt if it is from a personal contact or personally relates to them.

"Phishers have also started to develop a psychology behind their emails that plays off urgency, greed or trust. Combined with the legitimate look and feel of the spoofed websites, even more cautious and aware users can fall victim to their attacks"

These methods all fall within the classification of spear-phishing, where the attacks directly target specific victims with something in common that they can exploit. Spear-phishing requires some information about the victims – their bank, where they work, what sites they've ordered from recently – to produce a targeted attack, and much of this data can easily be found by combing profiles, blogs and other websites. Some phishing attacks even incorporate malware such as worms or trojans into the emails they send, which then directly compromise the security of the victim's computer and create another tool from which they can select victims and send out attacks. Phishers have also started to develop a psychology behind their emails that plays off urgency, greed or trust. Combined with the legitimate look and feel of the spoofed websites, even more cautious and aware users can fall victim to their attacks.

Phishing by its nature is also widespread: in the final quarter of 2009, the

Anti-Phishing Working Group (APWG) found over 90,000 unique phishing emails and over 130,000 unique phishing websites. The estimates for the annual monetary losses associated with phishing are varied because of the lack of data from banks and other financial institutions, but are reported to be anywhere between $100m and $3bn just from victims in the US. Financial and banking services find themselves the focus of most attacks, making up almost 93% of reported attacks.

Phishing affects people globally and is conducted internationally, making it difficult to track and prosecute the criminals behind it. One common technique that phishers have utilised is called 'fast flux', where a large pool of proxies and URLs is used to keep the true location of the phishing site hidden. By doing this, it is harder to blacklist the site and the server being used takes more work to find. The attackers have also begun to produce networks, where each part of the attack is carried out by a different person. For instance, one person who is good at producing a forged site might produce a toolkit for other phishers to use, only requiring them to select a site to copy and where to send the information. These toolkit users would then only need to select victims and send emails. Interestingly, as many as a third of these toolkits would actually send the stolen data somewhere else. This way the person who created the toolkit has essentially recruited inexperienced phishers to do all the work and take the blame but reap none of the rewards. In this way, the true phisher could get away without detection.

Solution approach

We propose that there are three ways in which the solution to phishing can be approached: detect phishing attacks before they reach the user, detect once the user has reached the phishing site, or train users to detect or prevent them by themselves. Each option has its own benefits and downsides, but the best method is an approach utilising a mix of all three. Phishing is evolving every day to avoid detection and bypass these defences, so by taking on all three we increase the chances that they will be found and stopped. Figure 2 shows our approach and the proposed anti-phishing solution framework.

Step 1 – Prevent phishing: Phishing can be stopped before it reaches the user either by blacklisting or blocking phishing sites or by filtering out phishing emails. The first method is carried out by looking at the URLs and the sites that they claim to be, either manually or automated through the use of machine learning. Although this may catch some sites, there is little hope of catching all of them, since a phisher can easily just make a new site once one is taken down.

The second method can be seen as more effective, because if successfully carried out it will stop the user from ever being exposed to the link for the phishing sites. There are many successful spam filters used by email servers, but few phishing filters due to its more complex nature. Filters for phishing are being designed using machine learning techniques as well. In 'Classification of Phishing Email Using Random Forest Machine Learning Technique' the authors discuss the characteristics used for classifying phishing emails. Some examples of these are the use of URLs containing an IP address, non-matching 'href' attributes and link text, the number of dots contained within a domain name and checking the domain names against the email sender. There are also a few simple keywords that the program looks for, such as 'urgent', 'update', 'suspend' and 'verify'. The result of their experiment showed an accuracy of 99.7% with a very small false positive rate of about 0.06%. This indicates this method is a very effective method of combating phishing, even more so since the machine learning technique can evolve with the evolving phishing attacks.

Step 2 – Detect phishing: Since attackers use sophisticated methods to ensure that phishing emails and websites reach vulnerable users, a method is sought to either identify possible phishing sites or indicate to the user to avoid malicious sites (or avoid giving malicious information in these emails or sites) even if they have received (and opened) a malicious email. Many web browsers already have defences in place against phishing sites, which will either have a passive indicator or an active indicator. Active indicators will have pop-up windows with a warning that the site they are on is a suspected forgery or that it is not considered safe, while passive indicators do not interrupt the user's task.

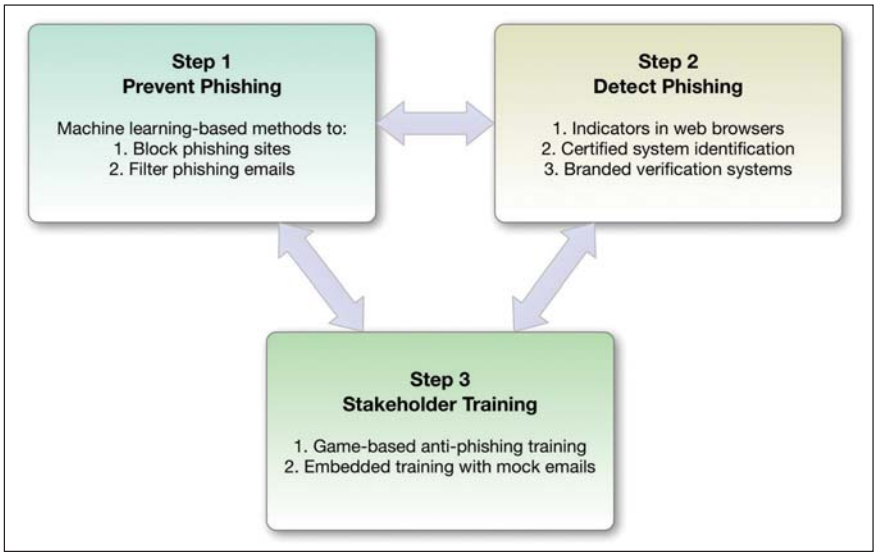

Figure 2: Proposed solution for phishing challenges.

As expected, many users would ignore or simply not notice the passive indicator and active indicators were much more effective. However, some users trust that the sites they are going to are what they expect because they were originally sites they trusted. To combat this, applying a verification system for sites that are trusted and secure can be helpful. If users see that verification every time they visit the genuine site, they are more likely to notice its absence on the fake website. The provision of the certified identification and branding attracts the eye and helps assure the user that they are on the correct site.

Step 3 – Stakeholder training: Training users to avoid falling for phishing scams is the third approach in our solution methodology. Most existing general phishing training is broad and does not combat the current more advanced phishing attacks, plus it depends on users actually engaging with and reading the material. Emailing warnings or material about phishing generally does not work because most users have been conditioned to disregard such emails and believe that they know how to protect themselves.

In our solution, we propose anti-phishing training methods using games or embedding training systems into an email server. Researchers are working on such games. For example, one of the most successful examples of the game format is 'Anti-Phishing Phil', a micro game that helps teach users to identify suspicious URLs and other components of phishing scams. This approach is both engaging for the user as well as informative, but users must still go to this program for themselves.

The other method, embedded training, can be useful because by sending mock phishing emails, users who are not trained in avoiding phishing scams will be trained by default. In the report 'Protecting People from Phishing: the design and evaluation of an embedded training email system', the authors outline a system that would send mock phishing emails which, if users opened and followed the link, would direct them to a page notifying them what was wrong with the email and what they should look for in the future to avoid being phished. Another approach was to use a comic to outline some key tips to help users avoid compromising their personal information. Both groups performed better than the control group, which only received security notice emails. This is a useful method because if the user is using the email server and clicking on the bait emails, then they will encounter the training email and become more aware of the risks, turning a premium phishing victim into an educated user.

Recommendations for future work

Phishing is increasing in complexity and is becoming harder to identify for cyber-security professionals. On the other hand, phishing is also becoming more complicated for attackers due to the increase in online security in recent years. Phishing is also getting more complicated for victims because new methods of attack make it harder for the lay person to distinguish phishing activity from normal activity.

We believe that the best defence to protect against phishing on a widespread scale would be to incorporate the proposed anti-phishing solution framework described in the previous section into email provider servers, such as Gmail, Yahoo and Hotmail. This would ensure that even those not experienced with computers and phishing risks are still protected at the most basic level. It would be even more effective if these servers also incorporated embedded training systems into their email services, because then the users will become more educated on how to protect themselves in the future, which will lead to a society of aware users and make it very difficult for phishers to successfully launch attacks.

Conclusion

Phishing is becoming an ever-growing threat to users as the attacks evolve and become more difficult to distinguish. The criminals who carry out these attacks are increasingly hard to catch. To combat these challenges, we have proposed a three-pronged approach. The use of a filtration system helps lessen the number of phishing emails that reach the user, decreasing the chances that they will be phished. The user interface model provides users with warnings when the site they are visiting is not trusted, therefore defending against the chance that a convincing email has led them to a phishing site. Finally, by engaging users with educative games or embedded training, the users themselves can start to practise methods of preventing phishing.

Even though attackers keep updating phishing tactics and it's becoming a more complex task to prevent and detect phishing, staying up to date with

Continued on page 13...

...Continued from page 13
machine learning-based automated defences in these three categories in our proposed solution approach will be able to help keep phishing under control.

About the authors

Ike Vayansky is currently a graduate student in the Information Systems Technology programme at the Coastal Carolina University, Conway, South Carolina, US. He earned his BS degree in Information Systems 2016. His current research interests are in cyber-security, data science and machine learning. He has worked on several independent programming projects in his free time. He has been programming and coding since the age of 10 and has experience of hosting servers for games. He can be reached at irvayans@coastal.edu.

Dr Sathish AP Kumar is currently an Assistant Professor in the Department of Computing Sciences at the Coastal Carolina University. He earned his PhD degree in Computer Science and Engineering from the University of Louisville, Kentucky in 2007. His current research and teaching interests are in cyber-security, data science, big data analytics and distributed systems. He has published more than 30 technical papers. He is also a senior member of IEEE. He can be reached at skumar@coastal.edu.

References

1. Matthews, L. 'Phishing Scams Cost American Businesses Half A Billion Dollars A Year'. Forbes, 5 May 2017. Accessed Jan 2018. www.forbes.com/sites/leemathews/2017/05/05/phishing-scams-cost-american-businesses-half-a-billion-dollars-a-year/#3c420cc73fa1.
2. Srivastava, T. 'Phishing and Pharming – The Deadly Duo'. SANS Institute, 2007. Accessed Jan 2018. www.sans.org/reading-room/whitepapers/privacy/phishing-pharming-evil-twins-1731.
3. Singh, NP . 'Online frauds in banks with phishing'. The Journal of Internet Banking and Commerce, vol.12, no.2, pp.1–27, 2007.
4. Hong, J. 'The Current State of Phishing Attacks'. Communication of the ACM, vol.55, no.1, pp.74–81, 2012.
5. Akinyelu, A; Adewumi, AO. 'Classification of Phishing Email Using Random Forest Machine Learning Technique'. Journal of Applied Mathematics, vol.2014, pp.1–7, Apr 2014.
6. Caldwell, T. 'Spear-phishing: how to spot and mitigate the menace'. Computer Fraud & Security, Jan 2013, pp.11–16. Accessed Jan 2018. www.sciencedirect.com/science/article/pii/S1361372313700071.
7. Jagatic, T; Jakobsson, M. 'Social Phishing'. In Communications of the ACM 50, no.10 (2007): 94–100.
8. Kumaraguru, P; Ree, Y; Aquisti, A; Cranor, LF; Hong, J. 'Protecting People from Phishing: the design and evaluation of an embedded training email system'. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp.905–914. ACM, 2007.
9. Herzberg, A; Gbara, A. 'TrustBar: Protecting (even) naive web users from spoofing and phishing attacks'. Bar Ilan University Technical Report, 2004.
10. Sheng, S; Holbrook, M; Kumaraguru, P; Cranor, LF; Downs, J. 'Who Falls for Phish? A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions'. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp.373–382. ACM, 2010.
11. Jakobsson, M. 'Modelling and Preventing Phishing Attacks'. In Financial Cryptography, vol.5. 2005.

EVENTS

2–4 February 2018
REcon Brussels
Brussels, Belgium
https://recon.cx

7–8 February 2018
Manusec Europe
Munich, Germany
www.manusecevent.com/europe/

9 February 2018
Hackron
Canary Islands, Spain
www.hackron.com

16–18 February 2018
Munich Security Conference
Munich, Germany
www.securityconference.de/en/

20 February 2018
European Information Security Summit
London, UK
https://biztechevents.co.uk/teiss/

22–24 February 2018
International Conference on Information Systems Security & Privacy
Funchal, Portugal
http://www.icissp.org/

22–23 February 2018
DevSecCon Singapore
Singapore
www.devseccon.com/singapore-2018/

27 February – 3 March 2018
NullCon
Goa, India
http://nullcon.net/website/

2–4 March 2018
Hacktech
Pasadena, CA, US
http://hacktech.io

6–8 March 2018
National Privacy & Data Governance Congress
Calgary, Canada
http://pacc-ccap.ca/congress/