Tools and Techniques for Improving Cyber Situational Awareness of Targeted Phishing Attacks

Phil Legg[1] and Tim Blackman[2]
[1]Department of Computer Science and Creative Technologies
[2]Information Technology Services
University of the West of England

Abstract—**Phishing attacks continue to be one of the most common attack vectors used online today to deceive users, such that attackers can obtain unauthorised access or steal sensitive information. Phishing campaigns often vary in their level of sophistication, from mass distribution of generic content, such as delivery notifications, online purchase orders, and claims of winning the lottery, through to bespoke and highly-personalised messages that convincingly impersonate genuine communications (e.g., spearphishing attacks). There is a distinct trade-off here between the scale of an attack versus the effort required to curate content that is likely to convince an individual to carry out an action (typically, clicking a malicious hyperlink). In this short paper, we conduct a preliminary study on a recent real-world incident that strikes a balance between attacking at scale and personalised content. We adopt different visualisation tools and techniques for better assessing the scale and impact of the attack, that can be used both by security professionals to analyse the security incident, but could also be used to inform employees as a form of security awareness and training. We pitched the approach to IT professionals working in information security, who believe this may provide improved awareness of how targeted phishing campaigns can impact an organisation, and could contribute towards a pro-active step of how analysts will examine and mitigate the impact of future attacks across the organisation.**

Index Terms—**Cyber situational awareness, phishing, visualisation, user experience**

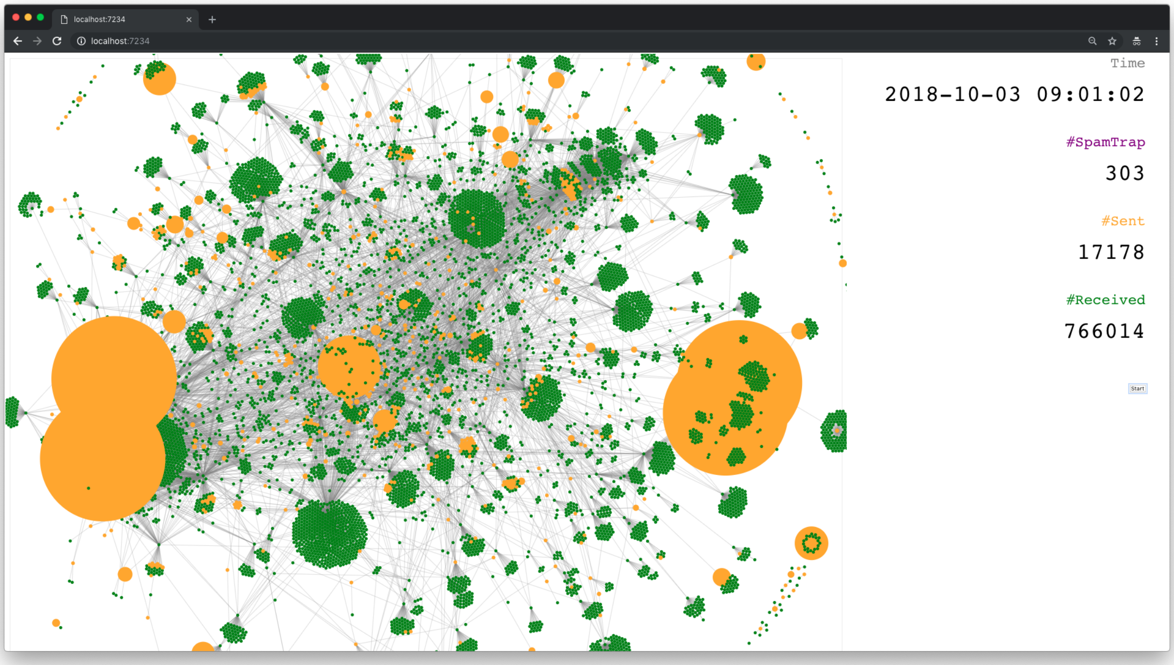
I. Introduction

EMAIL has become a fundamental component of the modern connected world, and is the primary form of communication within, and between, many organisations. Given the wide-spread adoption of e-mail for both business and personal use, it is no surprise that phishing e-mails serve as the most common attack vector for obtaining unauthorised access or stealing information from unsuspecting users. Phishing e-mails typically consist of some message designed to have the victim behave in some manner (which normally means clicking on a malicious hyperlink). Early phishing examples were often very generic and sent in high volume in the hope of attracting even just a small percentage of clicks (e.g., notifications of online orders, winning the lottery, dating messages). In contrast to a high volume attack, spearphishing attacks are highly curated and personalised for the victim. Examples may include personal details gathered from public social media accounts, and make references that entice the victim to believe the communication is genuine. For example, in 2016 Snapchat fell victim to spearphishing when a member of HR handed over sensitive staffing information, having received an e-mail request that was supposedly from the CEO [1]. In a similar incident, Ubiquiti Networks suffered a loss of £33.6m when an employee made a false payment, due to receiving an e-mail that impersonated the CEO and made the request [1]. Spearphishing e-mail are highly bespoke and so often do not spread quite like phishing campaigns, however they are often extremely effective when well executed. This illustrates that just as security researchers try to improve spam detection methods, and end-users become aware of typically phishing e-mails, attackers will continually aim to evade protective measures and lure users.

A number of recent surveys have been published on the impact of phishing within organisations. Wombat Security Technologies [6] states that 76% of organizations say they experienced phishing attacks in 2017. In their 2018 Internet Security Threat Report, Symantec report that by the end of 2017, the average user was receiving 16 malicious emails per month and that fake invoices were the most common disguise for distributing malware [5]. Verizon claim in their 2018 Data Breach Investigations Report [7] that 92.4% of malware is delivered via email. In the Human Factors Report by Proofpoint, they claim that Dropbox phishing lures are the most common, whilst DocuSign lures are the most effective [4]. Finally, the FBI's 2017 Internet Crime Report [3] suggests that business email compromises cost organisations a total of $676 million in 2017. These reports highlight the significance of phishing attacks, and also highlight that despite users becoming more aware of phishing, the problem continues to persist.

A fundamental challenge within security is how to improve awareness among end-users. Technical protections such as spam filters help to reduce the volume of phishing e-mails that users will receive, however attackers are continually looking for novel approaches to overcome such mechanisms. In many cases of a 'worm', an attack is spread by users within an address book or contacts within existing e-mails, and so it becomes more challenging to know whether to accept or to block the communication. In the incident we describe in this paper, a targeted phishing campaign hit a large University where e-mails were passed based on existing contacts and as

[1]Contact E-mail: phil.legg@uwe.ac.uk



Time
2018-10-03 09:01:02

#SpamTrap
303

#Sent
17178

#Received
766014

Fig. 1: Node-link view of the targeted phishing attack. Both Staff and Student with 10508 identities, 16850 linkages, and 17179 emails sent.

replies to previous subject lines, in an effort to convince users that the communication is genuine.

In this short paper, we present tools and techniques for analysing and visualising a targeted phishing attack. The objective is to improve cyber situational awareness of the spread of a phishing campaign across the organisation, which can inform on the impact and severity towards the organisation, and can inform responsive action. Our study is based on a real-world dataset gathered from the University of the West of England, where a phishing campaign was observed by IT staff during a single weekend that had impact on over 10000 user accounts.

II. Problem Context

University of the West of England (UWE Bristol) is the largest University in the South-West region of the UK [2], and it has the second largest IT environment in the region. In total, UWE employs approximately 4,000 staff, and in the 2017/18 academic year, had 28,790 registered students. All staff and students are issued with a UWE e-mail address that is used as the primary form of communication to conduct daily workload activities. As a large, dynamic and complex organisation it is a challenging environment for the Information Security team to monitor, assess, and respond to threats as they appear on the network. Unlike many other organisations, a University environment has a much more open policy for both staff and students, with a mix of corporate and personal devices on the networks, along with legacy systems and non-standard interfaces. Therefore, tools that allows users to obtain a greater view of the overall environment to improve situational awareness are vital.

In September 2018, the University observed a phishing attack across the e-mail network. E-mails were distributed by legitimate users, and were replicated using previous conversation histories as the subject line, making for more convincing lures for recipients. The body of each e-mail was generic, with a button image that read ‘Please click here to read this message’. Whilst the use of a generic e-mail body may alert some users, for many they may well accidentally or intentionally click the button, due to curiosity, lack of awareness, or due to issues of cognitive load. This latter point is a crucial consideration for many modern working environments. We now receive e-mails on all manner of different electronic devices, and all hours of the day. A user who engages with a phishing e-mail whilst otherwise distracted (e.g., travelling on public transport, or whilst in conversation), may be more likely to fall victim to a attack compared to someone who is fully concentrating on the task in a quiet office space.

III. Phishing E-mail Analysis

The first stage of the investigation was to extract the dataset from the University e-mail service, Office365. The extraction of the dataset was performed using the knowledge of the generic e-mail body content, to find all e-mails that consist only of “Please click here to read this message”, during the known time period. Future work would explore how more sophisticated extract methods can be deployed, however this was not the purpose of this initial study. As a combination of

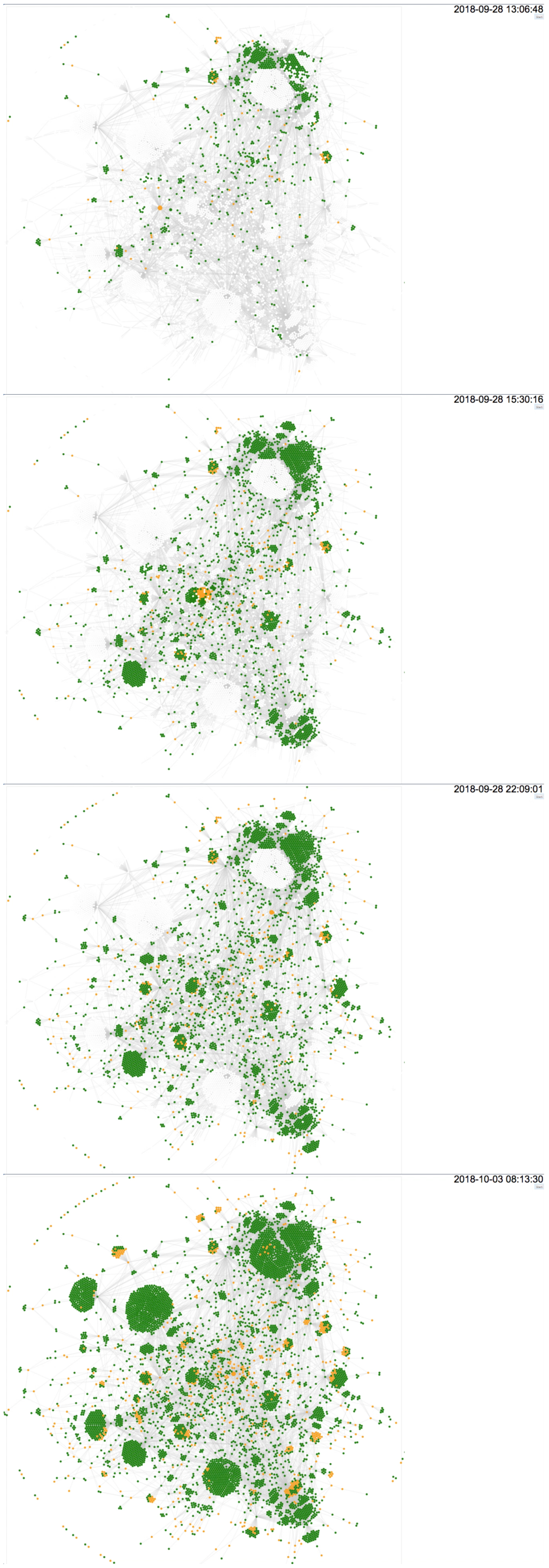


Fig. 2: Sequential view of node-link diagram for the staff e-mail dataset, consisting of 5963 unique users, 11816 connections, and 10097 sent e-mails.

both staff and student accounts, this resulted in 24762 rows of data, 10508 identities, 16850 linkages, and 17179 emails sent. Given that there are over 10,000 users of the University e-mail service affected by this incident illustrates the importance of understanding the characteristics of the attack, how the incident played out, and whether key events can be identified that contributed towards the spread. To achieve this, we use three different visualisations that all contribute towards improved situational awareness at different levels of granularity: a node-link visualisation that conveys the scale of the infection and the individual users within, a multi-temporal time-series plot that shows a summary of the incident at different temporal views, and an activity sequence visualisation that shows a detailed chain of attack for an individual user.

A. Node-link analysis

Figure 1 shows the resultant node-link visualisation after the incident. Given the shear scale of our data, we use a background worker to deploy a force-directed simulation and then render a static representation of the layout. This helps to maintain consistency for comparison of timesteps. Since the data is pre-filtered to only include users who receive the phishing e-mail, the clusters are indicative of users who receive the message from the same sender, which can help identify key events. The animation aspect of the interactive node-link is depicted by keyframes in Figure 2. We use colour-coded nodes to distinguish the current sender (orange) from the recipients (green). When animated, sender nodes are sized based on the number of recipients for the current e-mail, making it more visually apparent which senders contribute towards larger distribution of the malicious e-mail.

B. Time-series analysis

Figure 3 shows the number of sent e-mails (red line) across multiple-aligned temporal scales: minutes, hours, and days. Having an aligned view of the different temporal scales is particularly useful, since we can observe the granular spikes of activity at the minute level, whilst observing the general trend over the day at the hour level, or over the course of the spread at the day level. We also show the cumulative number of e-mails sent (blue line), which in this example is 24762 e-mails over the 7 day period. It is interesting to observe the large initial spike that occurs early on (our first day shows only 4 e-mails, so the majority of activity followed on the second day). The day view shows two peaks in the distribution. On inspection, we observe that the initial spike of activity was on a Friday, with a decrease over the weekend when users are away from their work devices.

Regarding the impact of the attack, the minute level shows peaks in the region of 700-1000 malicious e-mails being sent in a single minute of activity. Assuming that in the majority of cases, these are independent e-mail users, we can begin to comprehend how many users will have been impacted by this attack, along with other important aspects such as the amount of inconvenience, loss of productivity, human resource

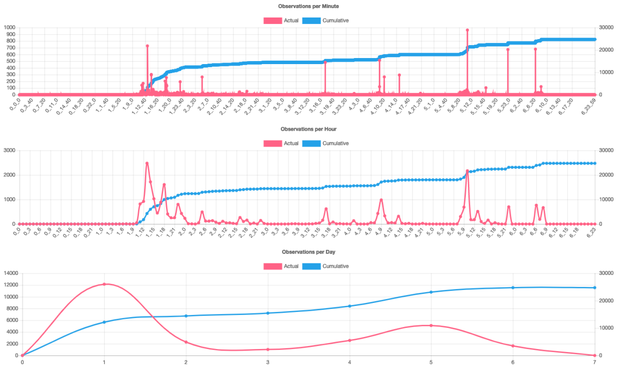


Fig. 3: Aligned time-series plots at different temporal scales to assess the targeted phishing impact, based on number of suspicious e-mail observations. Top: Per minute. Middle: Per hour. Bottom: Per Day

for resolving the outbreak, and other business functions that will have been impacted by this event.

C. Sequence visualisation

The final visualisation technique that we explore is how to observe a trace of activity for an individual, or group of individuals. We refer to this as the activity sequence visualisation. For a given user, we want to observe what e-mails they have received inbound, and what e-mails have been sent outbound. In particular, we are interested to observe the frequency, variety and scale of outbound communications, to see whether the user e-mail account has been sending high volumes of similar messages during a short period of time. It is not uncommon for members of staff to e-mail many recipients at once. However, if many identical messages are repeatedly sent then this may be of concern.

Figure 4 shows the activity sequence visualisation. For each of the examples shown, the individual of interest is shown as a light blue node. In this view, we use orange nodes to illustrate recipients from the individual (with light orange nodes showing recipients of recipients). Purple nodes show senders to the individual (with light purple being those who send messages to the set of senders). In this manner, we can begin to see an activity chain for a given individual to see the contribution they had on the overall scale of the campaign, and analyse the broader 'worm' distribution of the e-mail. Figure 4 shows a high contribution user at the top, where a large number of users receive the e-mail from this individual (and a large number receive it as a result of those recipients). In the bottom-left is a case where no subsequent e-mails were sent by the user, showing an 'endpoint' of the spread. In the bottom-right is a case where a user has a smaller contribution towards the spread of the e-mail. In all cases, we can trace back to investigate the origin of where the e-mail chain begun.

IV. Conclusions

In this short paper, we present an initial study of analysing the spread and impact of phishing e-mails across an organisation. We adopt three visualisation techniques to show how these can help to convey the spread of phishing e-mails effectively, in a manner that may be able to both inform security analysts, and engage end-users to promote greater security awareness. An initial consultation with IT professionals suggested that these techniques would be effective tools within the University security environment for analysing e-mail activity. Furthermore, they could potentially help highlight analysts to activity of interest for further investigation by creating a visually-appealing form of examining e-mail activity. Future work is required to study the effectiveness of how such public visualisations displays could facilitate security awareness and training, and how more sophisticated e-mail threats could be identified by security analysts using visualisation techniques.

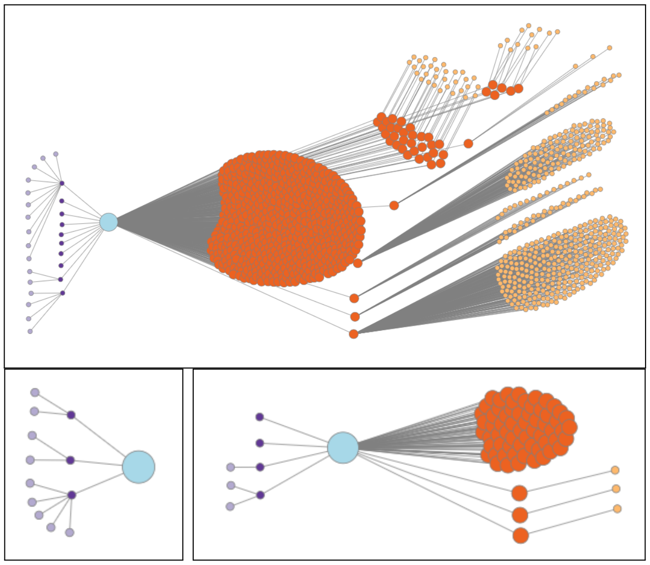


Fig. 4: Sequence plots for three users to illustrate their contribution towards the 'worm' distribution of the targeted phishing campaign. Top: High contribution to the spread. Bottom-Left: No contribution. Bottom-Right: Some contribution.

References

[1] Guardian. Snapchat leaks employee pay data after ceo email scam. https://www.theguardian.com/technology/2016/feb/29/snapchat-leaks-employee-data-ceo-scam-email, 2016. Accessed: 2019-02-22.

[2] H. E. S. A. (HESA). Where do he students study? https://www.hesa.ac.uk/data-and-analysis/students/where-study, 2019. Accessed: 2019-02-22.

[3] F. B. of Investigation. 2017 internet crime report. https://pdf.ic3.gov/2017$_IC3Report.pdf$, 2017. $Accessed$: $2018 - 12 - 05$.

[4] Proofpoint. The human factor 2018 report. https://www.proofpoint.com/us/human-factor-2018, 2018. Accessed: 2018-12-05.

[5] Symantec. 2018 internet security threat report. https://www.symantec.com/security-center/threat-report, 2018. Accessed: 2018-12-05.

[6] W. S. Technologies. State of the phish 2018. https://www.wombatsecurity.com/state-of-the-phish, 2018. Accessed: 2018-12-05.

[7] Verizon. 2018 data breah investigations report. https://enterprise.verizon.com/resources/reports/dbir/, 2018. Accessed: 2018-12-05.