Protection from Credential Loss Through In-house Phishing Campaign Profiling

By

Jeremy Koster

A thesis submitted to Macquarie University

For the degree of Master of Research

Department of Computing

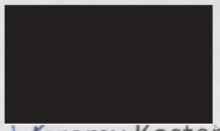
June 2020

MACQUARIE
University

Examiner's Copy

Declaration

This work has not previously been submitted for a degree or diploma in any University. To the best of my knowledge, the theses contains no material previously published or written by another person except where due reference is made in the thesis itself.
Approval has been provided from the Macquarie University Human Research Ethics Committee HREC Humanities & Social Sciences Committee with Reference No:52020515014436

Signed: 
Jeremy Koster

Date: 19/6/2020

Abstract

Phishing that leads to credential loss is a problem for organisations and individuals globally. High-profile data breaches regularly feature phishing and credential loss as major contributing factors. Previous research efforts have focussed on the identification of phishing emails and websites for all organisations. The focus of this study is to take datasets that are the product of one organisation's in-house account protection efforts and determine if successful phishing campaigns can be profiled. Findings include common characteristics of high-risk phishing campaigns and the profiling of IP addresses that were involved in phishing and credential loss specific to the target organisation.

Table of Contents

1. Introduction

The theft of usernames and passwords (account credentials) is an enduring and growing problem for organisations and individuals globally. A significant contributor to this problem is the prevalence of phishing campaigns that entice users into divulging their account credentials. Phishing is the imitation of a legitimate or trusted entity through digital means for malicious or criminal benefit. The issue of phishing continues to plague users of email, web sites, social media and the broader internet (APWG, 2019). The sustained success of phishing as an attack method is apparent in industry reports, appearing as a regular feature in major breaches. According to the Verizon DBIR, 32% of data breaches involve phishing (Verizon 2019). The Office of the Australian Information Commissioner (OAIC) states that 36% of reported privacy breaches from July to December 2019 reported to them involved compromised credentials through phishing (OAIC 2019).

Conducting phishing over email continues to be a popular means of targeting victims both at work and at home. The existence of phishing might be considered the result of an open communication medium, where any email user can be sent a message with minimal or no cost, coupled with a lack of sender authentication. The successful avoidance of a phishing email that has reached a targeted user relies on the identification of the email as an impersonation. While there are a range of non-technical indicators of a phishing email, such as knowledge of the sender, formatting and the style of language used, a knowledge of the underlying email technology can be a decisive factor for identifying well-crafted phishing emails. Users without the technical knowledge of the underlying email technology are often, and understandably, incapable of distinguishing a well-crafted phishing email from a legitimate email. Therefore, the modern IT Support staff member is tasked with this responsibility of perpetually looking for ways to block phishing emails before they arrive, engendering an awareness of phishing emails in staff members, and quickly identify occurrences of credential theft.

The issue of credential theft is a threat to the confidentiality, integrity and availability of digitally stored information. This has increased in scope due in large part to the expanding amount of personal and organisational information collected and stored in a digital format. As the value and size of the information asset increases so does the risk posed by exposure or damage to that information.

The protection of an organisation from credential theft involves a constant effort of monitoring, blocking and remediation on multiple fronts. The modern IT Support staff member tasked with this responsibility is perpetually looking for ways to block phishing emails before they arrive, identify credential theft quickly and create awareness of phishing emails in staff members.

No single phishing protection mechanism is currently available to solve the global phishing problem. Organisations require layered automated and manual responses now and into the future to continue to protect themselves from phishing attacks that lead to credential theft. These protection mechanisms can be costly to implement with an efficacy that is difficult to validate.

In this research, the prospect of deriving threat intelligence from the attacks that target an organisation, from that organisation's own systems, is investigated. This paper analyses specific data and metrics derived from the activities of an IT Support team protecting their organisation's Office 365 implementation from phishing attacks. The aim is to identify techniques to focus and hone protection measures already employed by an organisation to enhance security performance. If existing tools in use by an IT Support team can better identify or predict phishing campaigns and credential theft, more practicable phishing protection measures that assist in better detection and response may be found.

This paper focusses on a number of datasets that result from real-world Office 365 credential loss as a result of an anonymous organisation (Organisation X) being a target of credential theft. The two main ways in which credentials are lost for Organisation X are through successful phishing campaigns and credential stuffing. The data gathered includes the characteristics of successful phishing campaigns and the IP addresses that perpetrate credential stuffing.

1.1. Organisation

This thesis is divided into six sections. Following this introduction section is section two, a background literature review that discusses the existing research work leading to and surrounding the thesis topic. Section three provides problem statements that articulate the issues the research is attempting to solve, and section four summaries the key research contributions that the paper is aiming to make.

Section five discusses the datasets and methodology used for the research and the reasons why they were chosen. Section six contains five sub-sections that discuss the results that were derived from the analysis of those datasets, as well as a summary of findings. Section seven includes suggestions for possible future directions that were considered during the research and in review of analysis conducted. The final section, section eight, is a conclusion that summarises the research project.

2. Background and literature review

2.1. The issue of trust

The decisions that email users make to protect this growing repository of digital information assets is built on trust. Trust in the physical world is based largely on what people can see and hear. For example, a request for sensitive information can be assessed for legitimacy against a range of criteria and cues. Is the face of the requesting person familiar, do I know what this person does, or is the person looking suspicious? In general, these questions are easy to answer, making authentication of individual requests second nature. In contrast, trust decisions in an internet-enabled world, are made without being able to directly assess, the physical characteristics or features of the individual making the request. Without these cues present, people can understandingly mistake a fraudulent request for sensitive information, such as a username and password, for a legitimate request. Criminals consistently make use of misplaced trust to obtain an individual's username and password (credentials) to access digital systems.

Research conducted by Thomas et al., (2017a) provides insight into the numerous ways that credentials are be obtained by criminals. Hacking incidents that lead to credential leaks, keyloggers that infect computers and capture passwords, and phishing kits that trick

users into entering their passwords in fake sites. Many of these methods makes use of an impersonation or social engineering technique to gain unauthorised access to system or information that would otherwise remain secure. While these methods are known to be successful for obtaining credentials, phishing campaigns in particular are distinct in that they continue to target the lack of awareness in individuals and organisations without a clear solution for adequate protection. Even if organisations implement sophisticated technical measures that vastly limits the activity of attackers and the use of technology by staff, there will still be the issue of digital impersonation. The correct assessment of an email's legitimacy is still a human issue that requires internet and email users to have enough technical knowledge to identify well-crafted phishing campaigns. This is made even more difficult when a phishing email comes from a trusted organisation, a trusted individual or even from a colleague within the organisation about a recent subject. When compromised accounts are used to send phishing emails to previous contacts, an expectation that a user will necessarily detect the phishing is unreasonable. As long as there are parties looking to profit from digital communication channels, there will always be a need for a person to make a trust decision.

2.2. Phishing on other platforms

Phishing is not just an email-borne problem. Other technology platforms susceptible to phishing are continuously evolving and emerging.

An example of this is the technology infrastructure used to establish Wireless network connections. As demonstrated by the WiFiPhisher tool (WiFiPhisher, 2020), security researchers have devised techniques to determine the operating system of a connecting WiFi client and with a man-in -the middle attack, mimic the "connect to wireless network" dialogue box specific to that operating system. The victim places their WiFi password into the password dialogue prompt that is presented by the fake page generated by the WiFiPhisher tool. While this attack requires the attacker to be physically local to the target, it is a good example of the successful imitation of legitimate digital interactions to obtain a user's credentials.

Even on tightly controlled environments such as smartphone applications, there are opportunities for imitating legitimate digital interactions. For example, recent research

(Aonzo et al., 2018) has demonstrated the plausibility of phishing attacks utilising Android-based password manager apps. Such apps attempt to improve the user experience by authenticating user credentials to Android apps with minimal input, causing less user friction. However, Aonzo et al. note that password managers rely on non-definitive factors (heuristics, app name and other metadata) to determine a smartphone app's legitimacy. This allows for the possibility of an app to be imitated and hence invoke the password being supplied to a rogue app via the password manager. As an illustration, a malicious party may attempt to publish a fake Facebook app to the relevant Android app marketplace or use other methods for installation. The fake Facebook app would mimic the artefacts of the legitimate Facebook app and thereby trick the password manager into supplying credentials when the fake app is opened.

The platforms that are used by individuals on a day to day basis are changing rapidly. However, consumers will often choose convenience and openness over security to avoid a negative user experience, or loss of productivity. Therefore, it is clear that vulnerabilities posed by phishing and app impersonation will continue to shadow individual interactions with digital technology as it develops. Solutions to current and future issues are required for the safety of the growing population who use and trust online services with their information. The overarching challenge will be to identify solutions that are also appealing to the consumers that drive sales and revenue for companies that produce the technology. Understanding phishing techniques within the traditional email platforms will also assist new platforms in building protection measures into the emerging platforms.

2.3. Common phishing elements

Email-based phishing campaigns targeting credential theft typically comprise of several elements. There is an email that delivers the original impersonated request along with a URL or file that directs a user to a website. The website is often an impersonation of a cloud service such as Office 365 requesting a username and password to complete or satisfy the original request. The request can take the form of an urgent account issue, such as unauthorised access or sharing of a document through a cloud service, such as Dropbox. All elements, the email, the website and accompanying files are designed to trick the user in to

clicking, opening and providing their username and password. As opposed to imitating a person or organisation in the physical world, these digital elements can be crafted to look almost identical to the original and legitimate artefacts. The Office 365 login page can be visually replicated, as well as internal emails that have previously been obtained through an earlier compromised email account. A trust decision is needed to be made by the potential victim which can be challenging since they may not have the technical knowledge to determine the origin of the email or the legitimacy of the website's URL. Several studies have investigated how people respond to malware warnings and phishing emails (Neupane et al., 2015, Harrison et al. 2016). Overall, these studies have identified that users do not spend enough time looking at key phishing indicators and that personality traits, such as attention control and elaboration, impact a person's phishing detection accuracy. While the nature of phishing detection by people is a worthy study for the development of training material to protect individuals from phishing campaigns, it leaves us with the practical problem that it's impossible to train all users. Knowledge of phishing attacks fade over time and the sophistication and method of phishing campaigns continues to evolve. While we must continue to educate individuals to better detect phishing campaigns, technological solutions can provide viable security solutions. Previous efforts such as email content filtering, real-time block lists and sender reputation schemes have addressed numerous issues related to malicious emails. It is hoped that utilising existing technological controls to reduce the impact of phishing on the broader internet community is still a viable and useful endeavour.

2.4. Phishing email and website detection

The phishing problem is well researched and documented. Solutions to phishing, both implemented and proposed, are numerous. Two particular areas of study have been covered well by researchers when identifying phishing campaigns: First, detecting and blocking phishing emails, and; second, detecting and blocking phishing websites.

Email has enabled the almost instant communication between individuals across the globe without those individuals requiring an existing knowledge of each other. While there are industry-wide efforts to verify a sender of an email (such as Domain Keys Identified Mail and Sender Policy Framework) it is still relatively simple to convincingly impersonate a person or an organisation through email. Similarly, for website hosting, there are mechanisms such

as certificates and transport layer security to verify the identity of a website. But it's still possible to impersonate the website of a target organisation that is convincing to a proportion of the population.

This means that detection of both email and website phishing campaign elements remains a challenge for computer programs. Fraudsters are able utilise the open nature of email communication and website hosting to create and deliver content that looks legitimate to users but still conform to the rules of the protocols that transport the content.

2.4.1. Protection from spear phishing

Protection against sophisticated phishing campaigns is especially difficult due to the similarities shared with legitimate emails. Spear phishing campaigns are often well constructed and will target an individual and their normal business processes, often impersonating a trusted person. Since this kind of phishing is likely more time consuming to assemble for the phishing operator, the instances of phishing are less numerous, but the expected gain is higher and likely targeted towards a specific organisation. A recent research paper explains that spear phishing attacks have led to a number of high-profile breaches that have featured in the news (Ho et al., 2017). These include those suffered by the US State Department, the White House, Google and RSA. Spear phishing emails are difficult to identify because they are often smaller in number and do not contain the content that a spam filter might be equipped to identify.

This paper proposes a method of identifying spear phishing emails that draws on a dataset that contains SMTP (email server) logs, NIDS (Network Intrusion Detection System) logs and LDAP (authentication) logs. The proposed method is named Direct Anomaly Scoring (DAS) which identifies suspicious anomalies based on all other events in the datasets that are predominately legitimate and safe. This real-time detector sustained a very low false positive rate of 0.004%, equating to an average of 10 false alerts a day. While analysing 370 million emails, all but two spear phishing emails were identified, with two previously undetected phishing attacks identified as well. Considering the damage that spear phishing can lead to, this type of anomaly detection based on a broad dataset shows potential for protecting users from a particularly difficult form of phishing campaign.

2.4.2. Phishing website detection

Detecting websites used in phishing campaigns and blocking them at an organisational level can also be an effective means of protecting users from most phishing attacks. However, the task of phishing website identification remains a challenge for the entire industry. Research has been conducted into the construction of websites used in phishing campaigns and asserts that 90% of phishing websites are replicas of previous phishing websites, for the simple reason that common phishing kits are used (Cui et al., 2017). Automated downloading of 21,303 phishing websites was facilitated by a web crawler. Analysis of the structure of the downloaded website was then performed. The research investigates methods for identifying classes of phishing websites by comparing website code tag structure with the code from other identified phishing campaign websites. This method shows promise for protecting an organisation by providing tools that block access to suspicious domains, such as a browser plugin or as part of a web proxy that filters web access. The pervasive use of TLS (Transport Layer Encryption) and of free TLS services such as "Let's Encrypt" will make it difficult for network devices such as web proxies to perform inline inspection of websites. However, such protective tools can constitute another instrument in the array of controls needed for the ever-evolving phishing landscape. As with other phishing protection mechanisms, layering these controls will provide a good level of protection, but eventually, due to a requirement for email and internet access to be functional, a small proportion of phishing emails or phishing websites will still likely make it through.

Another technique that has been researched by a team from Department of Computer Science, Virginia Tech is the possibility of identifying domains registered to appear like legitimate domains (Tian et al., 2018). These are commonly called squatting domains, as they are registered in the hope of possibly selling the domain to a large company without actually using the domain for any functional purpose. In the case of phishing domains, they are used to reduce the chance that a victim will realise they are malicious by closely mimicking the legitimate website domain. This research has developed a technique using machine learning on text obtained from optical character recognition (OCR) on website screenshots to identify phishing websites. This enables analysis of a website as a user would see it, and obtain text from the phishing website to determine if it is genuine. This is needed because a significant amount of phishing websites obfuscate the identifying text of their website within the HTML

of the site itself. This technique of hiding the website text in code has the result of hiding the website text from a scanner that only relies on reading HTML code, not rendering as a browser would. Much like other tools used to identify phishing websites, this tool would help in identifying a good proportion of websites that other instruments may not identify. It is suggested by the research team that the tool can be used by organisations such as PayPal, that are commonly the target of phishing campaigns, to protect their customers. This technique can lead to an increase in protection from phishing campaigns, not as a single solution, but as part of a layered approach.

2.5. Phishing research challenges

The research of phishing activity has several inherent challenges. One of which is that real-life phishing is the attempt, and in some cases the result, of conducting illegal activity. It is also difficult to simulate phishing as the illegal activity cannot be identically replicated in an enclosed lab environment. Many of the techniques involved in protecting individuals from phishing are only effective in a live email stream or coming into contact with a live phishing campaign. This also means that researching phishing activity has a number of ethical considerations. Since real-world phishing campaigns actively engage in criminal behaviour there is a potential to allow or cause harm to an individual through information theft or privacy concerns. The researcher ideally wants to identify and capture real-world activity while still protecting potential victims from harm. In one study, researchers performed an investigation of real-world phishing activity by allowing phishers to install phishing kits on honeypot systems that appeared vulnerable (Han et al., 2016), but were actually contained. This allowed the research team to measure the lifespan of a phishing kit and further the understanding of phishing kit efficacy. This then allows designers of phishing protection measures to better understand the artefacts of phishing kits and how to identify them. Against the persistent threat of phishing, organisations will also need to fundamentally understand why recipients of phishing emails continue to fall for phishing techniques.

To understand more completely how victims fall for phishing techniques a research team from Auburn University Montgomery, Alabama, USA, have developed a browser plugin called ChromePic (Vadrevu et al., 2017). ChromePic can capture how a user reacts and is susceptible to phishing attacks while they interact with a live phishing website. This browser

plugin can capture user behaviour and detect credentials being handed over to the phishing campaign. This enables those tasked with protecting individuals from phishing campaigns to replay the way in which the phishing campaign was successful and to understand the cues that caused a victim to respond as they did. These insights may then also lead to more targeted training for the individual that has fallen victim to the phishing attack. In addition, knowing if credentials have been provided by the victim to the phishing website is also helpful, since a password reset can be conducted in this case. The researches note that it is even possible to capture the actual credentials that were provided into the phishing website. Care must be taken that the plugin stores any sensitive information in a secure manner and does not place user credentials at risk. The handling of another individual's username and password in a tool such as this would need careful consideration and may have other privacy implications.

2.6. Profiling phishing actors and campaigns

Once the phishing message has been identified and the phishing website has been analysed, another step that can be taken is to profile the phishing actor based on the assets (email, website, login system interaction) used in the course of the phishing campaign. It is possible that one can lure a phishing campaign operator into interacting with systems that can be monitored by the target organisation. If fake credentials are provided into the phishing website, then it is likely that the bogus username can be monitored, and the resulting interaction tracked or blocked. Without any provision of access, the target organisation can gain valuable information about the attacker from the attempted system access. This can include an IP address and the user agent values from the browser which often contains information about operating system and application versions. To extend this information further, methods for fingerprinting individual browsers are discussed in a research paper provided by Laperdrix et al., published in 2016. The paper identifies 17 individual browser attributes including headers, language and plugins that allow a website operator to fingerprint a particular computer environment. Some of these techniques use existing HTTP information, but also information requested of the browser through javascript and built-in functions (Laperdrix et al., 2016). It is possible that these techniques could be added to the content hosted on target services such as an organisations Office 365 login page to request

the additional information from a visiting website. With the identification of false credentials, the computer environment of a malicious actor can be tracked and blocked from accessing accounts where legitimate credentials were provided by a victim of phishing. Even merely knowing the IP address of a malicious actor can allow an organisation to monitor for additional suspicious activity from that IP address and resolve any compromised accounts and where necessary, block traffic.

To take this concept even further, Vastel et al. (2018) studied the possibility of increasing the accuracy of browser fingerprinting over time as software versions are frequently updated and changes to the HTTP headers will occur. Their tool PF-STALKER allows a website to track a particular browser instance for 51 days, extending the previous timeframe of 36 days from previous techniques (Vastel et al., 2018). To take this concept even further, a research team comprised of some of the same researchers from the previously discussed paper, studied the possibility of increasing the accuracy of browser fingerprinting over time as software versions are frequently updated and changes to the HTTP headers will occur. Their tool PF-STALKER allows a website to track a particular browser instance for 51 days, extending the previous timeframe of 36 days from previous techniques (Vastel et al., 2018).

The techniques of tracking a malicious actor so that an organisation can identify compromised accounts and block login activity is another useful tool in the layered defence that can be used by modern organisations to combat phishing attacks. If the use of fake login and password is used to track a phishing campaign operator's activity, they may also be used as a helpful security intervention to overwhelm the phishing campaign operator. Through the provision of many fake credentials, the operator will be forced to sort through them, consuming their time and resources in order to identify legitimate credentials. An origination wishing to seed a phishing website with fake credentials may need to identify a source of fake passwords that appear like passwords and are likely to be used by an account owner. Simply using known weak passwords like *letmein*, *querty* or *password* may be easily identified and filtered by the phishing campaign operator. Using varying length fake passwords consisting of randomised characters may also be easy to distinguish from real passwords as they are less likely to be chosen by an account owner. Evidence of password choice can be seen in publicly available compilations of breached passwords.

(PasswordRandom.com, 2020). A possible good source of bulk passwords that are difficult to detect, may be the many passwords dumps that have emerged from data breaches though the years. However, information obtained by illicit means has the potential to be harmful to individuals if used in a damaging manner. Thomas et al., (2017b) discuss the ethical issues of using password dumps and other leaked information for research purposes. Research Ethics Boards are required to provide guidance to researchers to ensure no harm is incurred to individuals during the course of research. When using information such as dumps of previously used passwords, ethical boards should be informed of the dangers of such a dataset, even if it is not considered research that involves direct interaction with individuals. Providing information to attackers that they may not already have may assist them in gaining unauthorised access and may increase the risk to the growing number of internet users.

2.7. Credential loss and Office 365

The number of internet-connected individuals has reached over 4.4 billion (Internet World Stats, 2019). Internet-based services and applications are also increasing in number and size. As discussed in recent research (Thomas et al., 2017a) the digital footprint of users expands to encompass social networks, financial records and data stored in the cloud. In many cases, this information is protected by a single "root" email account used for account and password resets. Once an attacker has gained entry to this account, access to other services and the information is possible. Services such as Office 365 or Gmail serve this purpose for many people and are increasingly becoming the aggregated doorway to large repositories of information.

In the past, organisations have built and operated their own mail servers within their own networks. This has meant that a fraudster needs to target individual company networks separately to identify web-enabled email access systems, such as Outlook Web Access. It requires additional effort to target these systems as email systems are set up differently and to the specification of the target organisation. Domain names, versions and login screens for each organisation are likely to all be different. However, with Office 365 there is one standard authentication page allowing phishers to target a single access point.. The phisher can simply, and indiscriminately enter lists of emails into the Office 365 login screen, and if the

account resides on Office 365, they can use pre-built scripts to access and abuse stolen credentials. This has led to Office 365 being the first-line access point for phishers and the battleground for a rapidly growing number of organisations trying to keep their information assets secure.

The automated submission of mass collections of lost credentials is known in the industry as "credential stuffing". This term was reportedly coined by Sumit Agarwal in 2011 while serving as Deputy Assistant Secretary of Defense at the Pentagon (Shape 2017). The credentials used in credential stuffing are predominately obtained by criminals through breaches of bulk account information from well-known and popular websites. Collections of breached credentials have grown to massive proportions, with a single compilation reaching 2.2 billion sets of usernames and passwords (Greenburg 2019). Research has shown that 6.9% of these credential sets remain valid, even years after being exposed (Thomas et al 2017a). This is due to the common practice of password re-use across multiple sites. This presents a problem for companies that have their information assets protected by usernames (commonly email addresses) and passwords that have been used for both corporate and personal sites by their staff. The sheer number of sites that the average person is required to maintain over time makes it difficult to remember a distinctly different password for each site. A recent paper presents a privacy preserving protocol that allows a user to be notified if the credentials they are using have appeared in a credential breach (Thomas et al., 2019). Through the use of browser plugin that makes use of the protocol, 21,177,237 lookup requests were made by 667,716 users. This equates to roughly 1.5% of credentials used by those using the plugin were exposed in a credential breach.

To compound the issue, platforms that are a target of credential stuffing and phishing such as Office 365, have become a significant repository of confidential personal and organisational information. Not only serving as an email system, mainstream cloud products include a growing number of services such as document storage, collaboration and data processing applications. As well as containing more types and a larger quantity of information, Office 365's user base is growing rapidly. In 2019 Office 365 active users grew by approximately 3 million per month. In October 2019 it reached a user base of 200 million users (Spencer 2019). As Office 365 continues to gain popularity it will continue to increase

as a target of credential stuffing and phishing campaigns. The protection of Office 365 also has a collective impact. Partner or associated organisations that lose Office 365 credentials, place other organisations at risk that have an existing relationship. Phishing exploits the existing trust relationships and business processes involving business emails to acquire credentials and commit fraud.

3. Problem Statement

The automated identification and blocking of phishing campaigns is a difficult problem. The less lofty objective of manually identifying and responding to phishing campaigns in a timeframe that limits damage, is also a difficult problem. Phishing campaigns are only limited by the ingenuity and imagination of their maker. The fraudsters will continue to exploit regular business practices and channels of trust to gain access to email accounts.

For the organisations and individuals around the world to safely use cloud products, such as Office 365, cost-effective and within-reach protection techniques need to be identified that can be implemented for well and moderately resourced organisations alike.

This paper aims to identify creative techniques to detect and respond to phishing attacks and credential loss thereby helping organisations protect their information assets.

4. Contributions

By collecting, enriching and analysing datasets from locally sourced systems, this paper aims to provide insight into the profiling of phishing emails and credential loss. In summary the contributions of this study are:

- Identification and analysis of datasets that are the product of protecting an organisation from phishing emails and credential loss.
- Identification of which methods of IP address enrichment provide value in profiling IP address that are likely to pose a risk to an organisation.

- Identification of common phishing techniques that represent the greatest risk to an organisation.
- Profile of the characteristics of phishing campaigns that have led to credential loss to determine if profiling information can be used to detect and respond to phishing campaigns.

5. Methodology

The main three toolsets available to Organisation X's IT support team include an antispam solution, the Office 365 platform and a log aggregation and analysis platform. Data collected from these platforms in the process of protecting real world phishing campaigns was analysed for patterns in the pursuit of identifying techniques to better respond to phishing emails.

These tools allow for the exporting of information, supplemented by data retained from manual tracking of phishing campaigns. The three main datasets that are the focus of this study is data related to phishing campaigns that were successful at penetrating the antispam defences, the IP addresses identified as malicious by way of failed logins (credential stuffing) and credentials that have been identified as lost or exposed.

In each dataset, the IP addresses involved were enriched with location, network provider and service type details.

A python script was written to lookup each IP address in the publicly available MaxMind GeoLite2 City database (MaxMind 2020). This attributes country code, city location and estimated GPS coordinates. For the purposes of this paper only the country code was used for analysis. Using a similar python script, the IP addresses were also matched against Autonomous System Number (ASN) and Autonomous System Organization Name (APNIC 2020) using the MaxMind GeoLite2 ASN database. These IP address details allow for the identification of the network operator that the IP address belongs to. Using this information, it is then possible to determine if an IP Address belongs to a consumer-style service or a server or datacentre-style service. This is done by performing searches on the ASN and the

organisation name in a number of network information services (PeeringDB 2020,) or even directly reviewing the network providers web site.

This is important for analysis as consumer-style services are typically used by end user devices such as personal computers and mobile phones. The use of these services for malicious purposes can indicate direct access from the attacker or the use of a device that is compromised by malware. The user of server-style services can indicate the use of anonymous services such as VPNs, TOR network nodes or virtual private servers (VPNs).

The enrichment of IP address lists with this data enables the profiling of attack locations and service types. This has formed a significant portion of the analysis activity included in this paper.

Three datasets are used as the basis for this research. The basic characteristics are shown in Table 1 below and described further in the following three subsections.

Dataset	Title	Time span	Records
1	Successful phishing campaigns	18 months	13,822 emails (date, sender, recipient, subject, sending IP address)
2	Credential stuffing	10 months	5,536 IP addresses (date, IP address, number of failed auth)
3	Credential loss	18 months	117 occurrences (date, sender recipient, subject, access attempt IP address)

Table 1

5.1. Dataset 1 - Successful phishing campaigns

The protective measures in place to stop phishing for Organisation X (predominately the anti-spam solution) blocks thousands of unwanted emails a day. This includes phishing, spam, hoax and virus emails. Only a small proportion make it through to Organisation X's staff mailboxes to be viewed by mailbox owners. Of those viewed by mailbox owners, a small proportion account for credential loss.

A successful phishing campaign for this dataset is considered a campaign that was able to send significant amounts of phishing emails through the anti-spam defences and resulted in a significant response from recipients of phishing emails or the IT Support team to remediate. Over a period of 18 months, the recipients of 21 discrete successful phishing campaigns were catalogued. Phishing campaigns were separated by sender. The data collection occurred during the day-to-day operations of the Organisation-X IT Support team. As successful phishing campaigns were detected, they needed to be investigated and responded to. In each case of a significant successful phishing campaign, the phishing campaign metadata was captured. This included the email send date, recipients, subject and sending IP address. The dataset includes 13,822 individual emails, 597 different subjects and 10,753 individual recipients.

The IP addresses captured in this data set were enriched with the IP location, network provider and service type details. This was then used to profile the phishing campaigns to identify patterns in the captured data. Microsoft Excel was predominately used to compose graphs illustrating the analysis if the captured data. To illustrate the relationship between recipients of campaigns, a Sankey diagram using the d3-sankey diagram library (D3 2020) was used.

5.2. Dataset 2 - Credential stuffing

As discussed earlier in this paper, credential stuffing is the practice of submitting large collections of usernames and passwords into services such as Office 365. This is perpetrated by malicious parties attempting to gain access to cloud accounts by trying the many username and password combinations at their disposal. Lists of usernames and passwords could be derived from a number of places. These were referred to in the previous section as *external means.* Organisation X begun monitoring for credential stuffing in its Office 365 platform in April 2019.

Failed access logs were obtained through the Office 365 Azure AD API. Office 365 allows for two methods of obtaining Office 365 authentication logs; Azure Active Directory API and the Microsoft Graph API. Both APIs offer different event formats and information.

Office 365 authentication logs obtained through Azure Active Directory API have less structure to the data format (field data is less processed) and provides event logs for login attempts of invalid accounts within the Office 365 tenant. Log events for invalid user accounts are particularly useful in this dataset as it provides more occurrences of failed login attempts. Cyber criminals have access to large amounts of credentials, but the quality and accuracy of the dataset can be low. This makes it easier to identify malicious behaviour, rather than generating false positives.

MS Graph API has a more structured format and more processed event fields, but only provides logs for existing legitimate Office 365 accounts.

The Azure AD API was used for building the list of malicious IP addresses as it provides details for failed login attempts for usernames that don't exist. This is crucial in identifying malicious IP address as the fraudsters have no real way to know the validity of a credential unless they try it.

The scheduled python script was used to collect the Azure AD API logs every 5 minutes accounting for event message delays. A second python script was written to run a query every 24 hours on all Azure AD login events collected in the previous 10 days. If an IP address failed access attempts of four different Organisation X accounts (valid or invalid) over a 10-day period, it would be added to the list of credential stuffing IP addresses. The list contained the IP address, the date the IP address was added and the number failed login events for accounts that the IP address had made in the 10-day period. This last field was a measure of aggressiveness. If the IP address had failed login attempts against only 4 account in the 10-day period, it was considered low in aggressiveness. A higher number would mean the IP address was being used more aggressively to stuff credentials.

10 months of credential stuffing activity was collected by Organisation X to identify credential stuffing IP addresses. 5536 IP addresses were collected over the 10-month period.

The IP address were associated with Geolocation from the freely available MaxMind GeoIP Lite 2 databases. The fields added were country code, city name and network name. A python script was used to perform the lookup and add the extra fields.

The IP addresses captured in this data set were enriched with the IP location, network provider and service type details. This was then used to profile the credential stuffing IP addresses to identify patterns in the captured data. Microsoft Excel was predominately used to compose graphs illustrating the analysis if the captured data.

5.3. Dataset 3 - Credential loss

For the purposed of this paper, credential loss is defined as the divulging of credentials by an Organisation X account holder to either an unauthorised party. Due to protection measures beyond the discussion of this paper, a credential loss does not necessarily equate to a successful unauthorised access of a mailbox and its contents. While attempts at access are made, even with correct credentials, credential loss is detected by other automated and manual processes such as IP address blocking, account activity monitoring and phishing campaign reporting by staff.

18 months of credential loss events were collected and analysed as phishing attempts were detected and responded to by Organisation X's IT support team. A credential loss is caused by either a successful phishing campaign or external means. *External means* is defined as a credential loss that occurred by a non-phishing method, such as a keylogging virus, use of the same password on another site, use of a guessable password or the compromise of a password management account such as Google Chrome with a personal Google account that contains the Organisation X account credentials. For each credential loss identified, the date, email address, phishing email sender, subject and IP address of the access attempt was recorded.

The successful phishing campaigns were categorised into sender type, email hook type, and website type.

Sender types included untrusted external, trusted organisation or trusted individual. *Untrusted external* is an email sender that is from an unknown organisation. *Trusted organisation* is a sender that is from a known organisation, but the sender name is not known by the recipient. *Trusted individual* is a sender with a name that is known by the recipient or previous legitimate correspondence.

Email hook types describe the hook included in the email to entice the recipient to click the email. Examples might be “shared file” or “account lockout threat”.

The IP addresses of the access attempt was captured and was the first malicious IP addressed used with credentials lost, when multiple where used. In the case of a credential loss from eternal means, this was the first IP address that used the correct credentials. This is known, as the logs obtained through the Office 365 API differentiates from incorrect password and blocked IP address errors. When a password is correct and the IP address is blocked, a different error is shown from when a password is incorrect.

The IP addresses captured in this data set were enriched with the IP location, network provider and service type details. This was then used to profile the credential loss IP addresses to identify patterns in the captured data. Microsoft Excel was used to compose graphs illustrating the analysis if the captured data.

5.4. Methodology choice

The method of using existing toolsets and processes to gather datasets relating to phishing campaigns has been chosen to explore the value of at-hand information. Phishing is a problem experienced by most large and medium organisations with a range of resources available to them. By utilising the dataset that are produced during the process of defending a particular organisation from phishing campaigns and credential loss, the analysis will be directly reflective of those presenting the greatest risk to the subject organisation.

An alternative method might include implementing dedicated processes or systems to gather more detailed data on phishing campaigns that the IT Support team would not necessarily require for the operational task of protecting an organisation from phishing campaigns. This method would likely bring greater insight into phishing campaigns but might lead focus away from the vital few campaign types and actors that represent the greatest risk to the organisation. The benefit of using real-world data from an operational IT support team brings common ground to this method for other organisations endeavouring to build their own understanding of the major phishing threats posed to them.

Another alternative is to gather data from other organisations to build a common dataset that allows for the protection from many existing and future phishing campaigns that

may target any given organisation. Phishing campaigns are wide and varied as are legitimate emails that an organisation receives. The implementation of email filters and IP address blocks needs to be balanced so as not to block legitimate email and access entering the organisation's systems. An email system that blocks all emails is secure, but not very useful to the organisation that operates it. Taking a dataset from a broad range of organisations may see an inefficiency in protection as maintenance is increases with the number of filters and processes.

A final alternative might be to gather data from fictitious organisations in an attempt to coax phishers to target that organisation. While this method might work to net some phishing campaigns, it would likely be difficult to build such a public profile that would attract phishing campaigns that would be relevant to other organisations. While a good dataset might be achievable that is relevant for that fictitious organisation it would most likely be irrelevant for an organisation hoping to protect itself from the phishing campaigns targeting it.

6. Analysis and Results

6.1. Successful phishing campaigns

The 21 successful phishing campaigns occurred over 18 months. The distribution of the phishing campaigns is shown in the two graphs below. Figure 1 shows the emails sent by individual campaigns during the data collection period. Figure 2 shows the cumulative emails sent during the data collection period. The rate of phishing campaign emails making it through the anti-spam defences has been significant in the early months (February to June) of 2019 and 2020. The later months of the year (August to December) in 2018 and 2019 are relatively quiet.

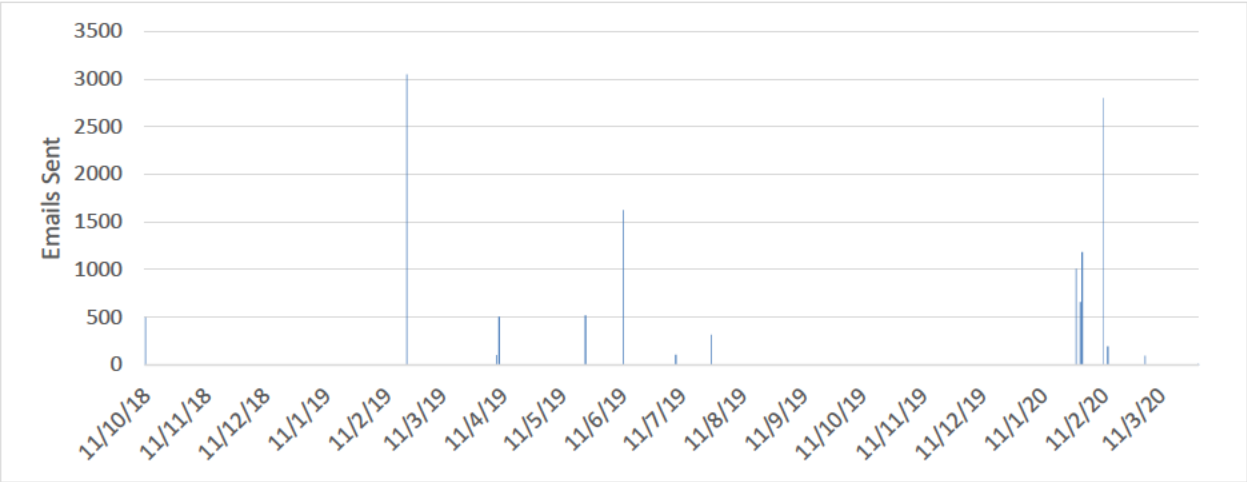


Figure 1

This is shown clearly in Figure 2 as the cumulative number of emails sent in February to June of 2019 and February to March in 2020 exceeds those in other months of the year.

This may be because the actors targeting Organisation X, have a natural cycle for credential theft and the subsequent use of credentials for fraud. Understanding the patterns and of increased phishing activity can assist an organisation in preparing for and resourcing additional efforts to combat peak phishing periods.

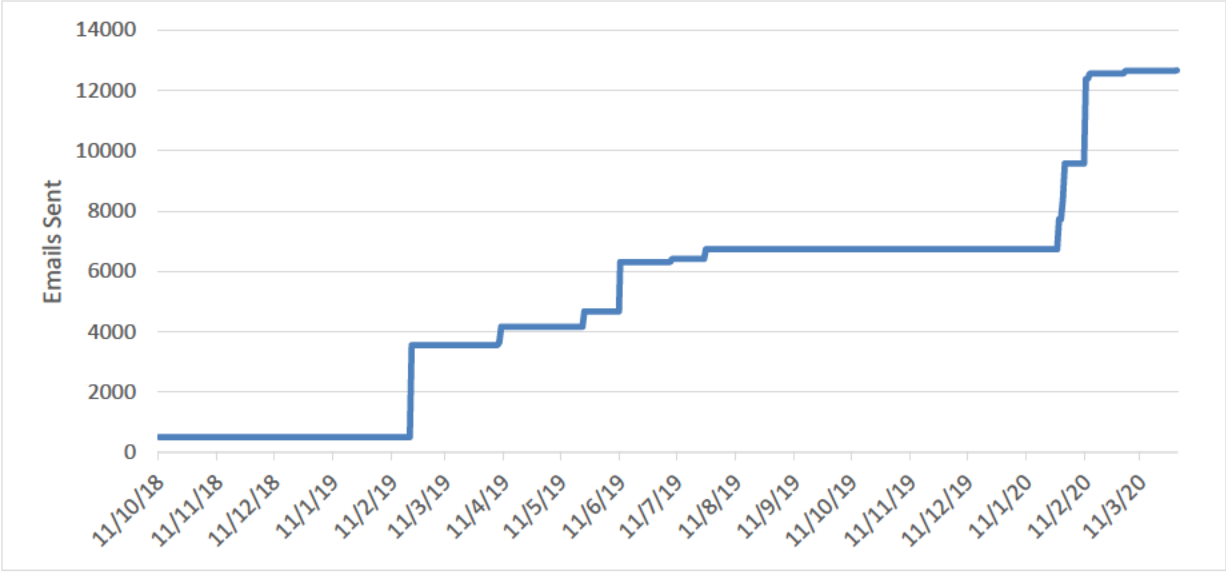


Figure 2

6.2. IP Address Analysis

Phishing campaigns were perpetrated from IP addresses from a number of countries across the world. The graph in Figure 3 below shows the distribution of IP addresses used for sending each campaign (shown as C1 through to C21 in the graph), their country location and the network provider name.

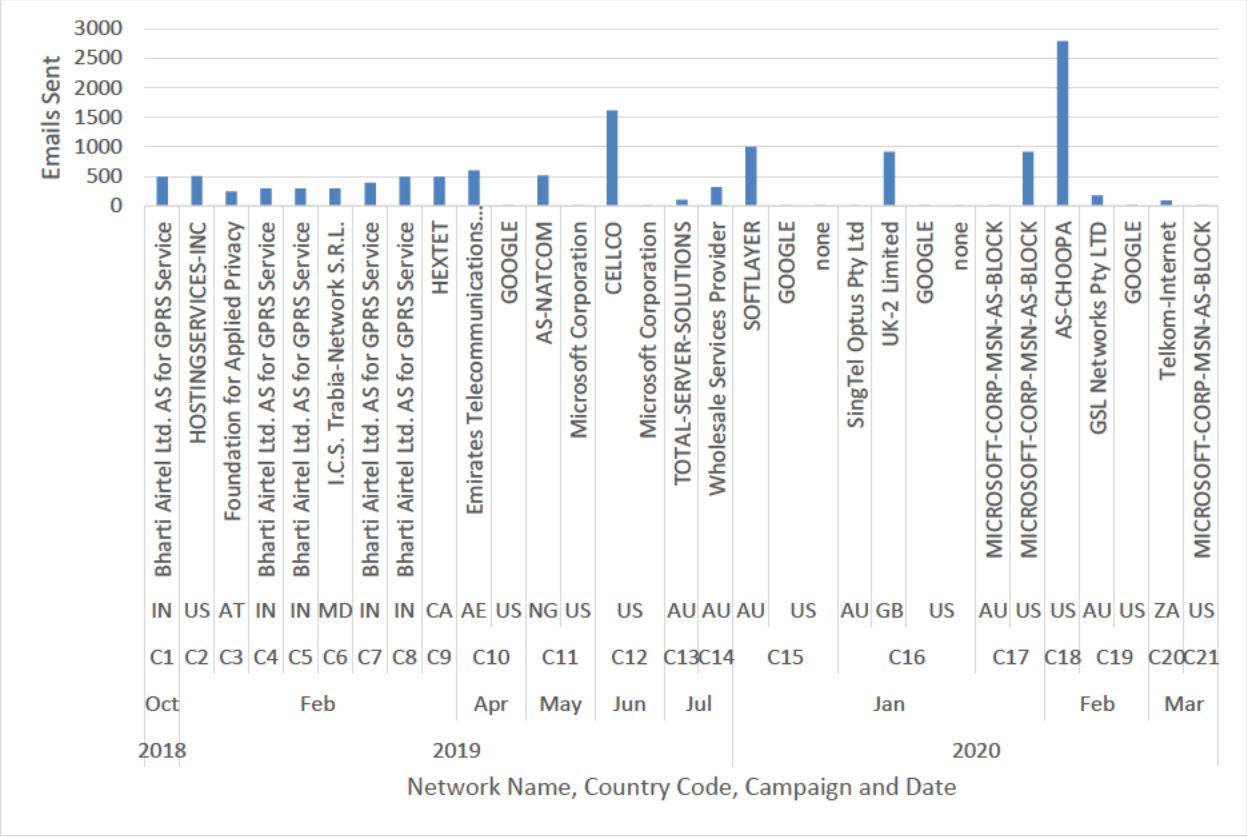


Figure 3

The prominent sources for the first 9 campaigns are India and the US, with the majority of these campaigns either being sent directly from India or non-consumer service IP addresses (such as HEXTET and HOSTINGSERVICES-INC). It is conceivable that the first nine campaigns were from an actor located in India, using non-consumer services occasionally to hide their location.

Campaigns C10 to C21 where either perpetrated from Nigeria, South Africa, The United Arab Emirates or non-consumer service. The use of non-consumer services becoming more regular with these phishing campaigns.

To focus on the consumer vs non-consumer service IP addresses, Figure 4 shows the split of these services with "No" indicating a non-consumer services and "Yes", indicating a consumer service.

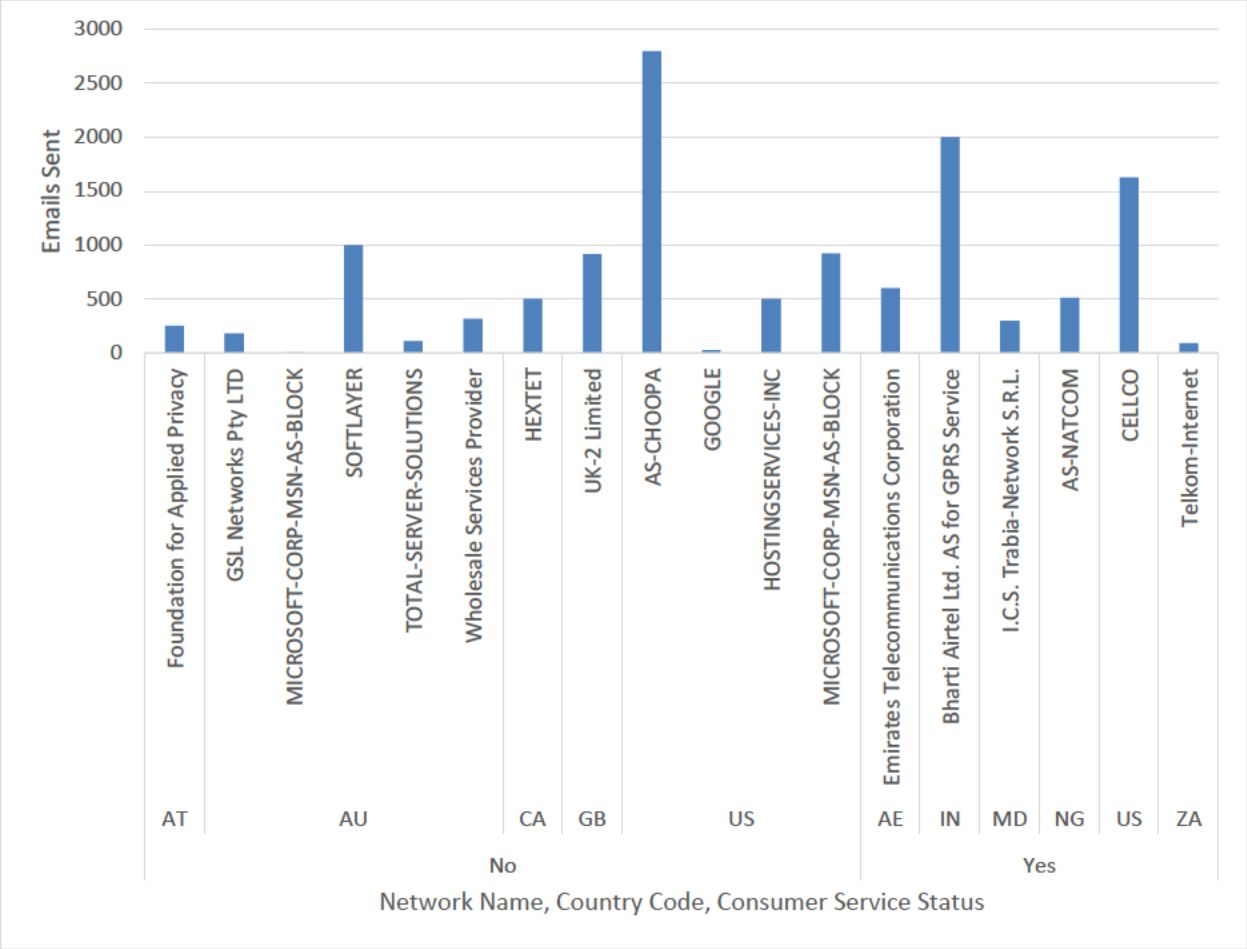


Figure 4

Purely relying on the geolocation of IP addresses that send phishing emails to block phishing can be easily subverted by phishing campaign owners with the use of non-consumer services IP addresses hosted in other countries. In particular it can be seen the non-consumer services hosted in Australia were regularly used. This may be further complicated by the legitimate use of VPN services by staff member access. Staff members

may use these services in an attempt to provide more security for their online activity, but instead end up making it more difficult to identify malicious access attempts.

An organisation may address this issue by providing a VPN solution to staff that can be identified by the IP address that is visible to Office 365. This would allow IT support staff to differentiate between legitimate VPN usage and malicious VPN usage. While some staff may still use anonymous VPNs, a firm approach of blocking access from these services would certainly limit the detection avoidance techniques in use by the malicious actors.

6.2.1. Campaign technique analysis

The relationship of recipients between successful phishing campaigns may allow for the profiling of specific phishing campaign types to assist in detection and response. Of the 21 individual campaigns analysed there were six main techniques utilised by the phishing campaigns.

1. **Personal information gathering** this technique used the offer of low-cost loans to gather personal information.
2. **Shared file – generic** this technique used a generic file sharing notification that led the target to a fake Office 365 login page or, in one campaign, a fake DropBox login page.
3. **Update contact details** this technique used the fear of false contact information to entice the target to click on a false Office 365 login
4. **Shared file – previous subject** this technique used a previously used email subject with a link to an Office document that led to a fake Office 365 login page.
5. **Message clipped** this technique used the target's name in the email subject and a "message clipped" message that linked to an Office 365 login page. This mimicked the behaviour of some well-known smartphone mail applications.
6. **None** this appeared to be a misfire as the email subject was "Important", but no link was included in the campaign email.

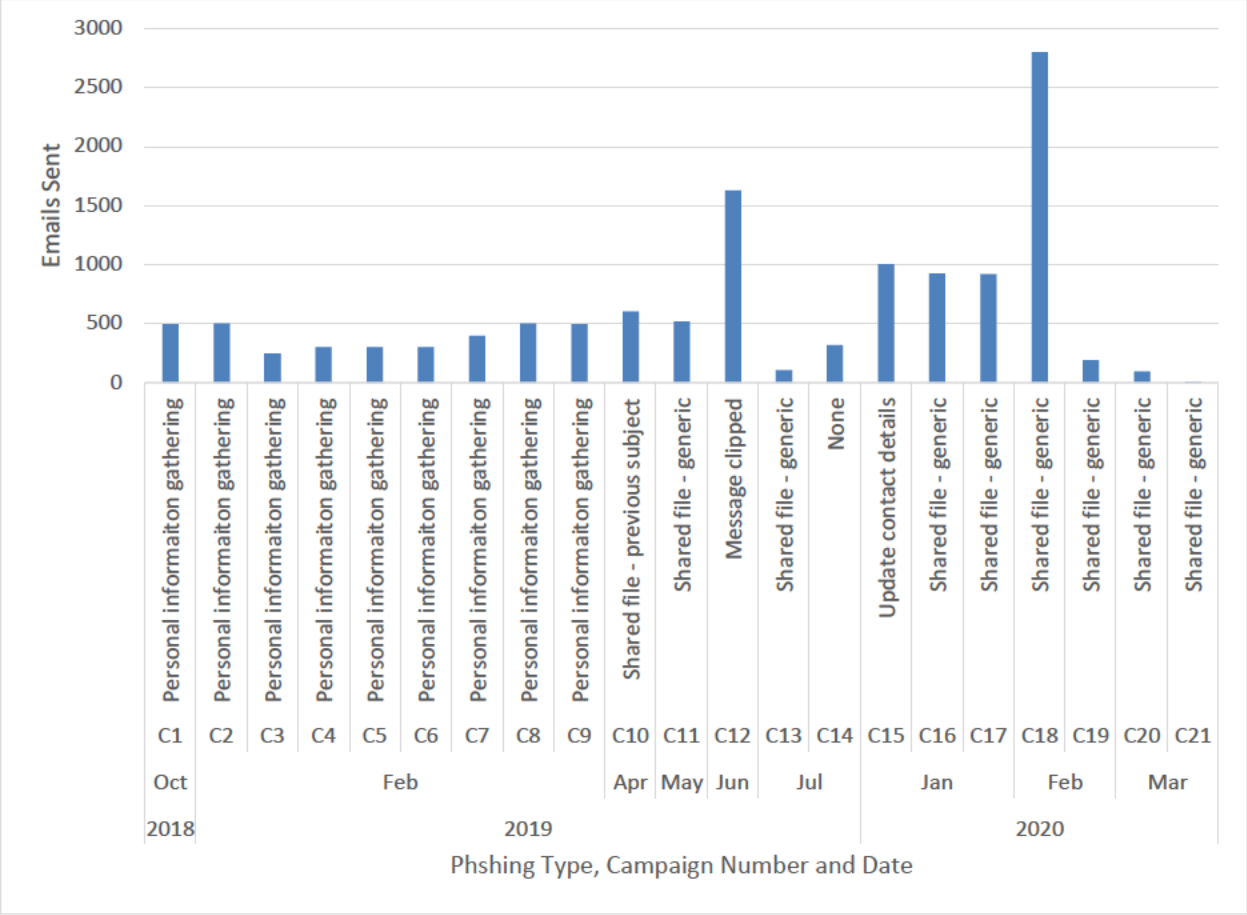


Figure 5

The graph in Figure 5 shows the change in phishing techniques over time, the more successful phishing campaign appears to have become "Shared file generic". The first nine campaigns shared the "personal information" technique. This graph also shows that successful phishing campaigns have generally reached more recipients over time with a sharp decline in the February and March 2020.

IT Support staff can be made aware of campaign types to quickly block access to phishing websites as phishing emails are reported by staff. Taking this profiling approach to phishing campaign defence that is specific for the targeted organisation may allow an organisation to be more efficient in protecting an organisation from specific phishing campaigns, rather than taking a whole of industry approach. For instance, warning staff about a delivery package notification that is prevalent in the broader industry but does not affect a specific organisation may waste time and resources in educating against a campaign type that does not affect that particular organisation. Analysing data from an organisation's

own defence systems may be a more efficient way to protect an organisation from plausible attacks.

Spending time analysing phishing campaigns that successfully make it through anti-spam measures may allow an organisation to better target phishing campaigns specific to their organisation for blocking and training purposes. For Organisation X there were only a few techniques used by the majority of successful phishing campaigns that made it through anti-spam defences. Simply relying on the anti-spam vendor to take data from the broader installation base of their product to build anti-phishing measures may leave the individual organisations exposed to phishing attacks that are specific to their organisation or industry.

6.2.2. Recipient Analysis

Considering that the campaigns may be related, the similarity in recipients for each campaign was analysed. Figure 6 shows the overlap of recipients that the campaigns shared. Campaigns 1-9 shared occasionally, with campaigns 1-3 borrowing heavily from each other. Campaigns 10-20 shared recipients regularly with each other. This supports the possibility that the campaigns 1-9 where from a single actor and campaigns 10-20 also shared a single actor. While not useful in applying a blocking measure to an antispam platform, as blocking emails with similar recipients would cause many false positives, it does provide an insight into how phishing campaign recipients groups differ and may assist phishing campaign operators in being profiled.

It also shows that a phishing campaign may be biased in their targeted recipients using a particular group or list of emails that is available to them. It's clear that campaign 2 and 3 are related as campaign 3 recipients were entirely included in the campaign 3 recipient list. This being the case, individuals within an organisation can be targeted for education and awareness if they are a regular recipient of phishing emails.

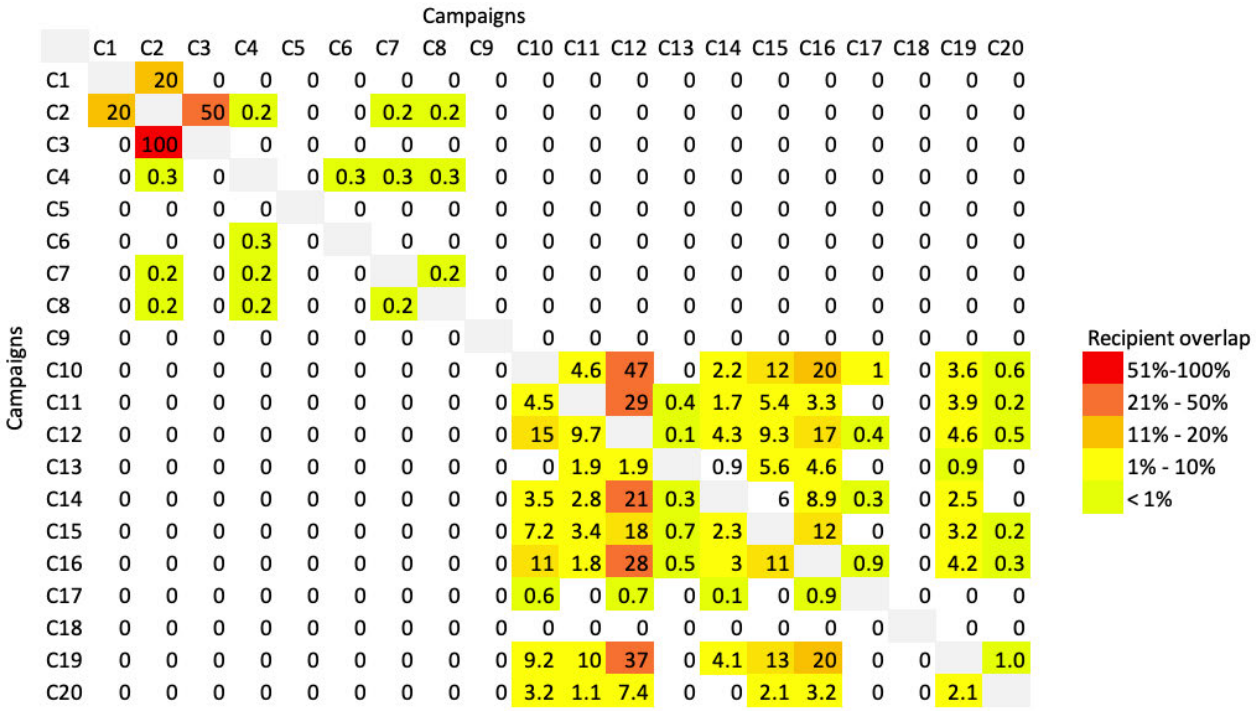


Figure 6

6.3. Credential stuffing

The 5535 IP addresses that were involved in credential stuffing over the 10-month period were found in 108 different countries. The distribution across the top 10 countries are shown in Figure 7 below. During the data capture period, IP addresses located in China accounted for 21% of the credential stuffing IP addresses.

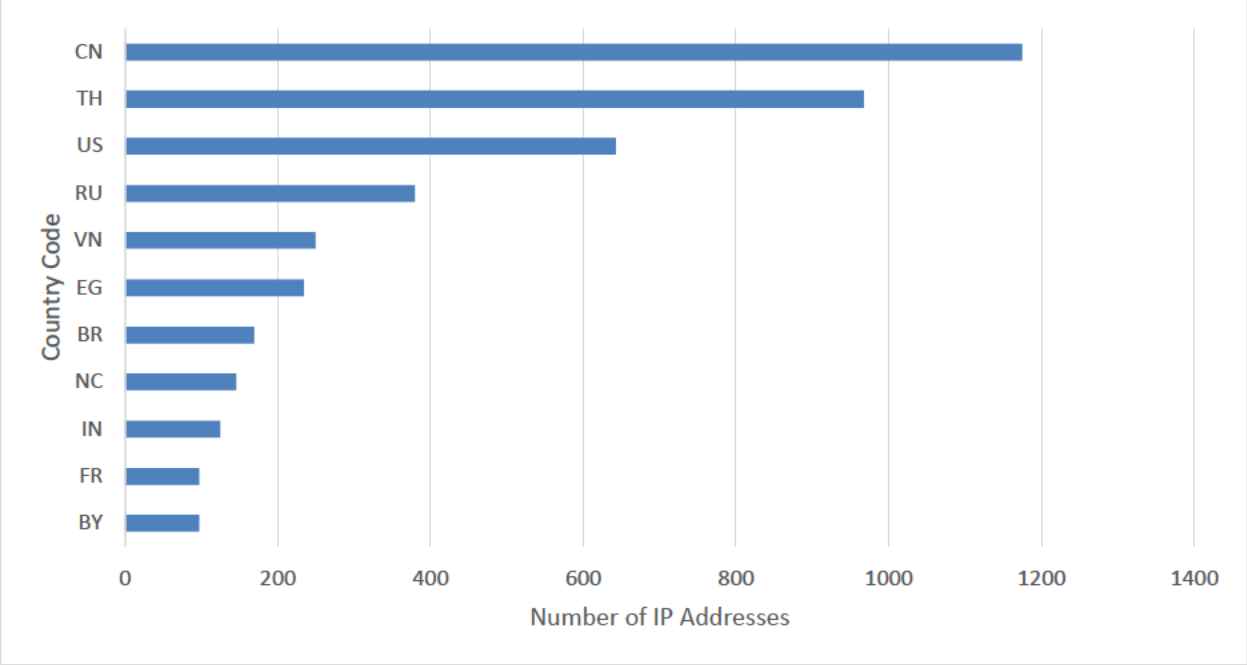


Figure 7

The top 10 network providers and their location by country, for the IP addresses that were involved in credential stuffing are shown in Figure 8 below. The spread across countries is large, making a whole country designation more difficult to use as an indicator of malicious access.

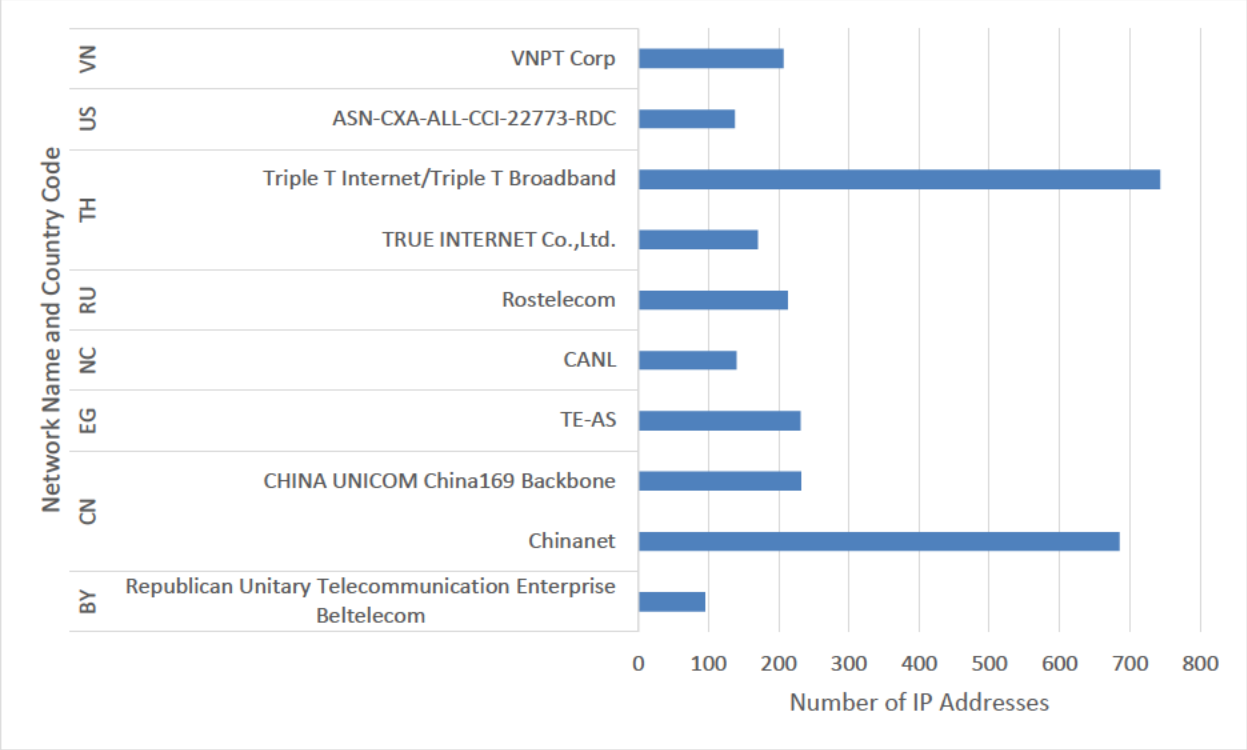


Figure 8

Table 2 shows the top 10 Autonomous System Number (ASN), it's related Organization Name and the network type. All of the network provider names are attributed to ISP type networks referred to in this paper as consumer-service IP addresses. Other types (non-Consumer IP addresses) are Content, Enterprise, Educational/Research, and Non-Profit services. By looking up network details it can be determined that the majority of IP addresses that are involved in credential stuffing are perpetrated from ISP type network connections that are likely to be consumer-style services. As it is unlikely that those orchestrating credential stuffing have dedicated infrastructure at each of these consumers-service IP address it's plausible that credential stuffing is perpetrated by compromised consumer machines, also known as botnets. With this being the case, it's possible for an organisation to maintain a list of known bad IP addresses and be alerted when a successful access attempt has been made. To extend this capability, it's also possible to keep a list of known network providers or IP address groups that have a high proportion of compromised computers, or botnets, and alert the IT support staff of accesses attempts from these IP addresses. This could lead to some false-positives, but it would depend on the amount of staff

that legitimately travel to those countries and use consumer-service IP addresses to login and interact with the organisations systems or Office 365.

Out of the 5535 IP addresses in the credential stuffing IP list only 13 IP addresses were consumer-service addresses located in Australia. Only 2 of these had tried more than 8 different accounts indicating that the threat posed by credential stuffing from Australian consumer-service IP addresses is low.

ASN	Organization Name	Type
45899	VNPT Corp	Consumer
22773	ASN-CXA-ALL-CCI-22773-RDC	Consumer
45758	Triple T Internet/Triple T Broadband	Consumer
7470	TRUE INTERNET Co.,Ltd.	Consumer
25490	Rostelecom	Consumer
17480	CANL	Consumer
8452	TE-AS	Consumer
4837	CHINA UNICOM China169 Backbone	Consumer
4134	Chinanet	Consumer
6697	Republican Unitary Telecommunication Enterprise Beltelecom	Consumer

Table 2

When the collection of credential stuffing IP addresses was initiated, there was an initial spike of 1024 IP addresses that were recognised as failing authentication attempts for more than 4 different accounts. The initial list consisted of 1024 IP addresses that did not represent a "per day" growth of the IP address list, so were removed for this analysis. Figure 9 and 10 below shows the count of IP addresses added every day since the list was created. Peaks of new IP addresses can be seen in July 2019, October to December 2019 and late March to early April 2020. On average 14.8 IP addresses were added per day, with 254 IP addresses added in the day of largest gain.

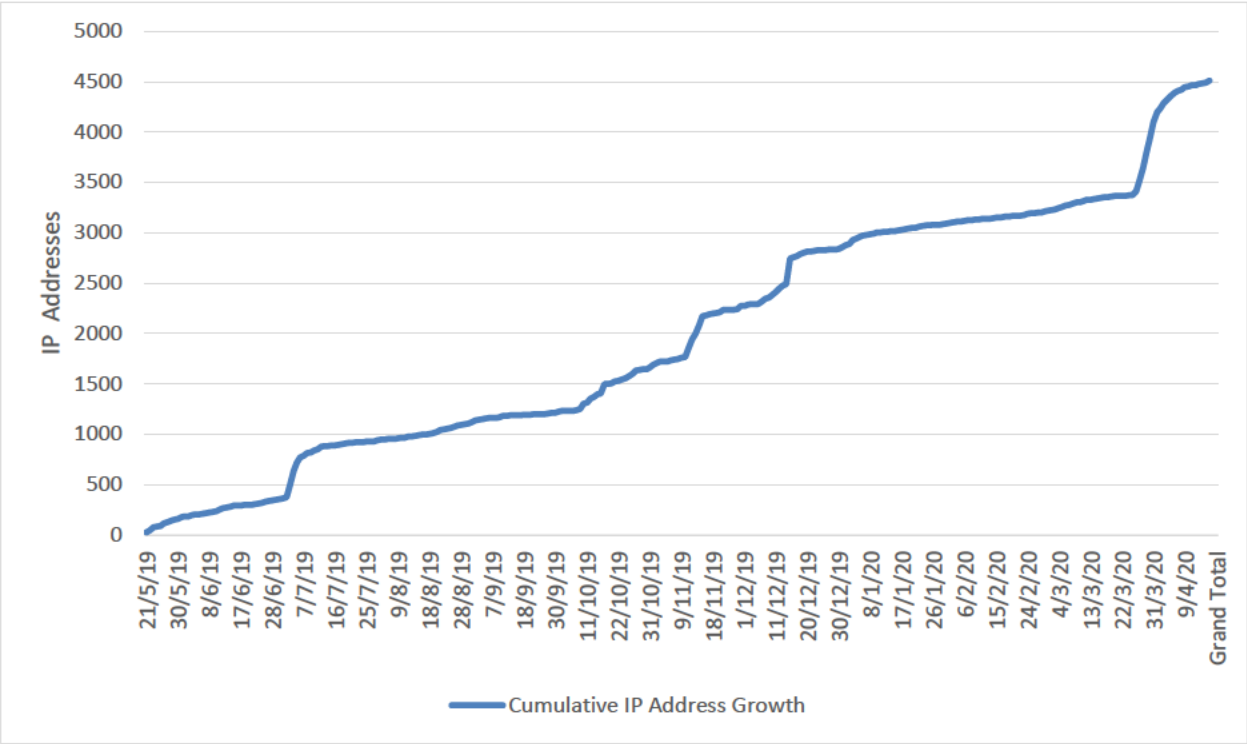


Figure 9

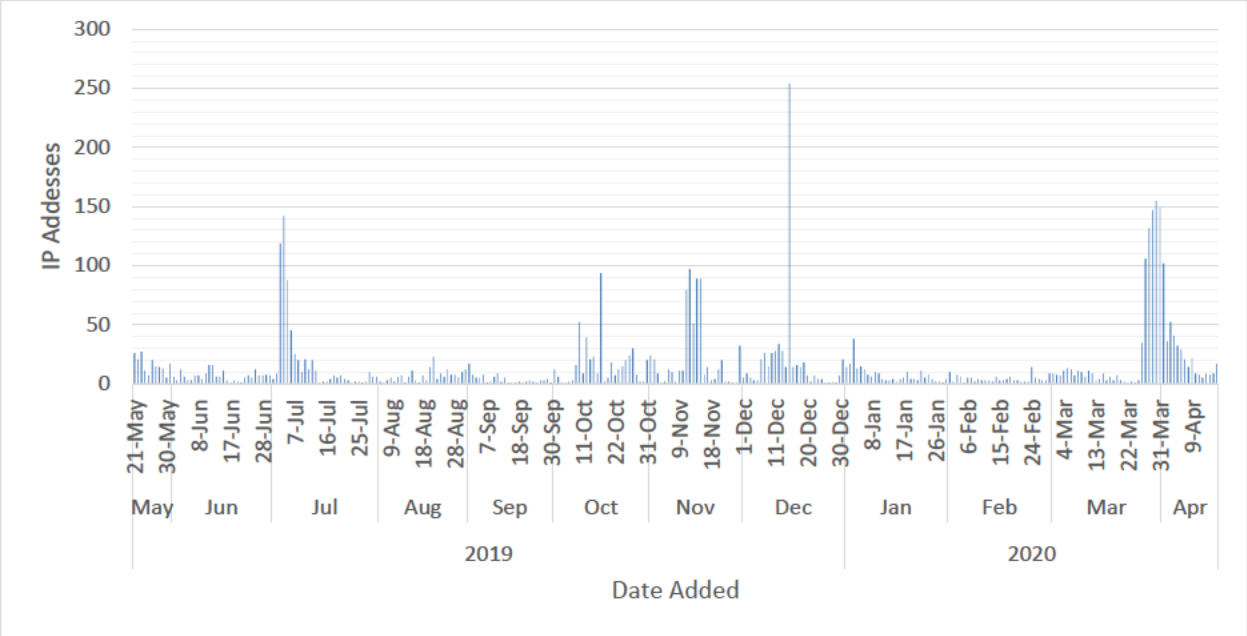


Figure 10

To determine if there is a correlation between credential stuffing IP address growth and credential loss, Figure 11 shows the dates in which credential stuffing IP addresses used

valid lost credentials. If credential stuffing IP address grew when there was a fresh list of lost credentials, then it might be possible to have forewarning of an increase credential stuffing success. Unfortunately, there does not seem to be a strong correlation between credential stuffing IP growth and credential loss. It's possible that the actors behind phishing and credential stuffing are the same even though they don't use the same IP addresses. One possible explanation for the difference in IP address sets is that there may be two different type of infrastructure used. Credential stuffing might be primarily utilising a credential dump from a data breach made public to facilitate the many authentication attempts involved in this activity. This would typically require a broad range of IP addresses, like that provided by a botnet, to avoid individual IP addresses being blocked due to too many failed attempts. However, the amount of credentials obtained by phishing would be much smaller in number and would not need the array of multiple IP addresses that successful credential stuffing requires. The use of five to 10 credentials from a single IP address is not as likely to trigger security control to block an IP address or alert staff. Credential stuffing may be perpetrated from botnet infrastructure whereas phishing campaign operators may be more likely to use a single IP address from an anonymous VPN provider.

However, Figure 11 does show that credential loss from credential stuffing occurs in batches with multiple lost credentials being used in short periods of time. This knowledge can allow IT Support staff to better pre-empt lost credentials when a single lost credential occurs. Once on high alert the IT support team can search identified IP addresses for further attempts and attempted credentials for further malicious IP addresses.

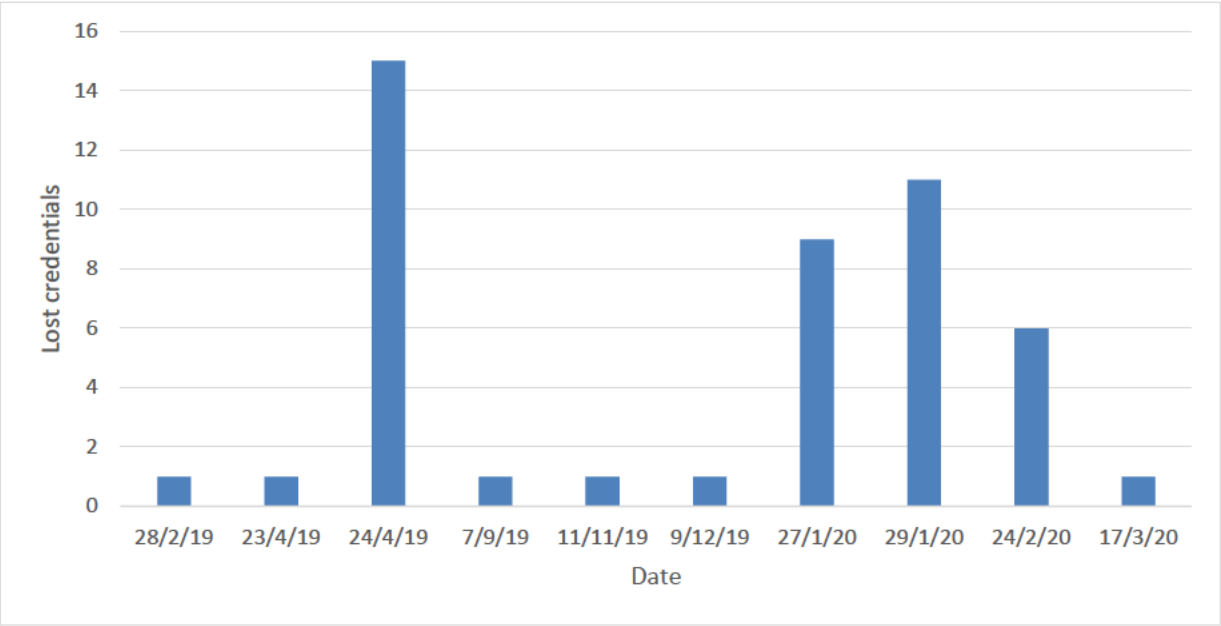


Figure 11

6.3.1. Peaks in credential stuffing IP address growth

On average credential stuffing IP addresses would try 8 different accounts before they were added to the IP address list. Figure 10 shows a number of peaks of botnet growth. Three peaks were analysed for similarity in geographical distribution. July 2019, November 2019 and March 2020.

July 2019 Peak

The peak of IP addresses added in July 2019 had an average of 3.8 attempts, from 44 different countries. The top 10 countries are shown in Figure 12 below. This peak of credential stuffing IP address growth had 30% of IP addresses located in Russia.

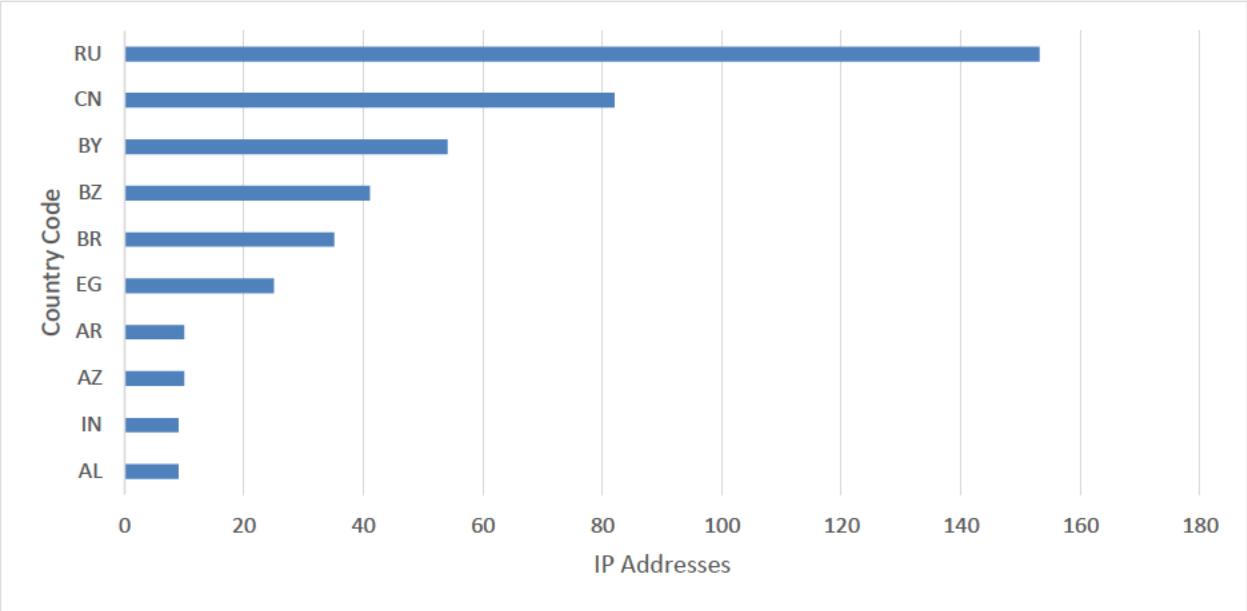


Figure 12

November 2019 Peak

The peak in November 2019 had an average of 4.9 attempts from 52 different countries with a bias towards Vietnam with 20 % of the IP addresses originating there. Graphs showing the number of IP addresses per country are shown in Figure 13 below.

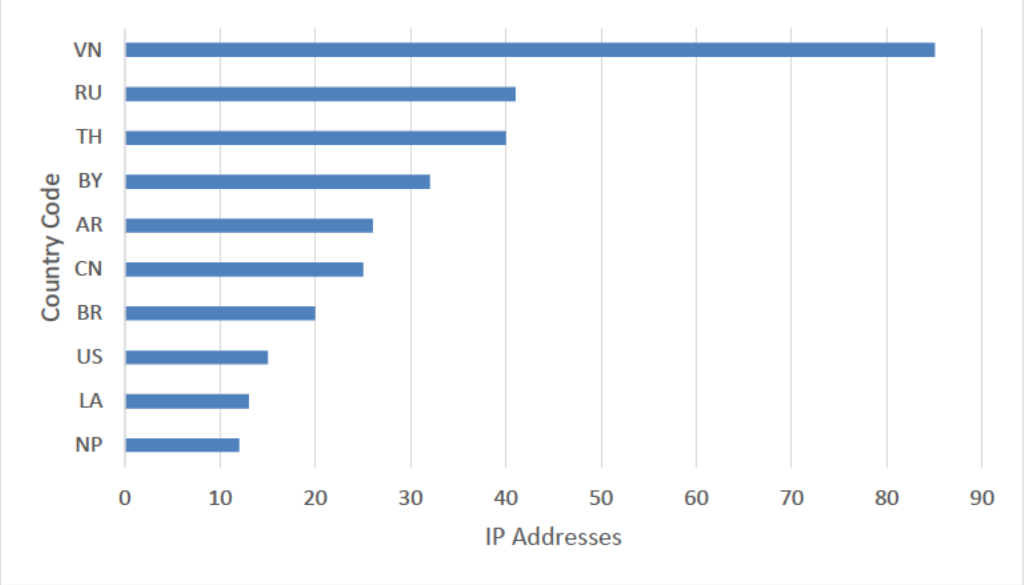


Figure 13

March 2020 Peak

The peak in March 2020 had an average of 6.1 attempts from 35 different countries with a heavy bias towards Thailand with 76% of the IP addresses originating there. Graphs showing the number of IP addresses per country are shown in Figure 14 below.

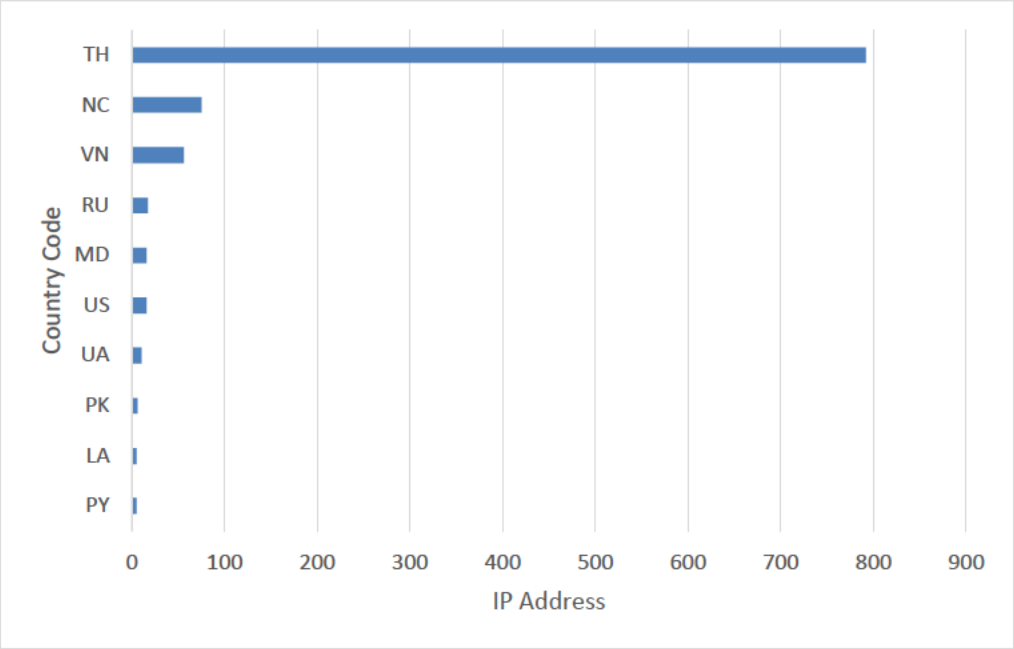


Figure 14

It's evident from the above set of graphs that some origin countries are more prevalent during a botnet growth than others. It's feasible that high growth rates in a specific country may indicate the recruitment of consumer machines by means of a malware campaign that targets that nationality or language in some way. This data might be able to be used by an IT Support team to pre-empt which region particular attacks might arise from during a botnet growth period.

The March 2020 peak also stands out because of the aggressive nature of the new botnets with an attempt average of 6.1 usernames per IP address. Constant monitoring of how aggressive credential stuffing is operating at, might be a way to identify different credential stuffing actors and understand when they become more active.

6.4. Credential stuffing success

Over the18-month period of lost credential stuffing tracking, only 14 IP addresses involved in credential stuffing used valid lost credentials. Figure 15 below shows the amount of lost credentials (obtained from the lost credentials dataset) used by individual credential stuffing IP addresses on individual days as well as network name and country code.

The bars shown in red for this graph are consumer-service IP addresses. Those in blue are non-consumer service IP addresses such as server hosting and data centre services. It is clear from this graph that the majority of risk related to credential stuffing comes from IP addresses from non-consumer services. The IP addresses that had the bulk of lost credentials originated from two known server hosting companies HOSTKEY-USA and SOFTLAYER.

For Organisation X, even though the majority of credential stuffing activity originates from consumer-service IP addresses, the actual risk is from a very small number (9) of IP addresses that are non-consumer service IP addresses. Even more specifically, 53% of lost credentials used in credential stuffing originated from the two non-consumer service IP addresses. While the HOSTKEY-USA IP address was added to the list in October 2019, it was in possession of 15 lost credentials in April 2019, before the credential stuffing IP addresses had been initiated. The SOFTLAYER IP address was added on the day it had possession of 11 lost credentials with a failed authentication count for 13 separate accounts. Adding these two figures the SOFTLAYER IP address had on this day 24 lost credentials, 11 of which were valid. Since that day, no more valid lost credential logins were attempted by this IP address.

This shows that the overwhelming majority of credential stuffing is unsuccessful for Organisation X. However, there are occasions where malicious parties have obtained a small set (11-15) of lost credentials. This behaviour is quite distinct from the mass of credential stuffing activity and can be focussed on to determine protection measures and causes.

It is also interesting to note that valid credentials were only attempted from non-consumer services. The non-consumer networks for SOFTLAYER IP and HOSTKEY-USA IP belong to hosting service providers that are often used by anonymous VPN service providers. It’s possible that once a malicious actor obtains leaked or stolen credentials that they have a high confidence in, the utilise an anonymous VPN service to hide their geolocation.

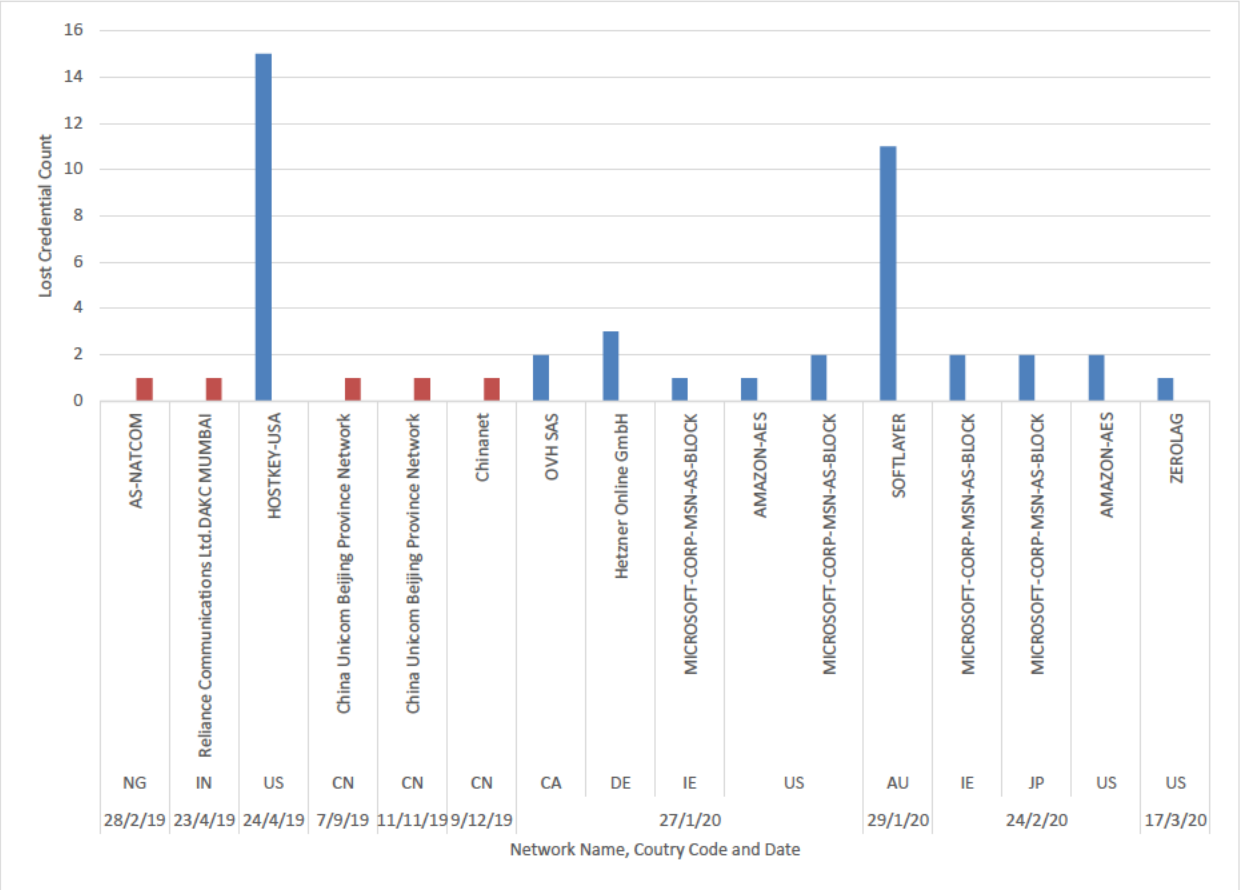


Figure 15

6.4.1. Credential Stuffing Lag

Of the 5535 IP address collected as credential stuffing IP addresses, 47 lost credentials were used during the data collection period. The graph in Figure 16 below shows how long (lag) it took for a credential stuffing IP address to make use of a lost credential once it had been recruited into the credential stuffing activity.

The negative entries before the May 2019 are indicative of credential stuffing IP addresses that were active before the credential stuffing data collection started.

Not including the lost credentials from before data collection occurred, on average, lost credentials were used by credential stuffing services 23 days after beginning the credential stuffing activity. This may indicate that there is little correlation between credential stuffing IP address recruitment and use of legitimate lost credentials, or that there is an inherent lag

in the process of feeding lost credentials into the credential stuffing system behind the credential stuffing IP address . This may be due to testing of the credential stuffing capability by the malicious actors, or an indication of multiple mechanisms in place between credential loss and credential use.

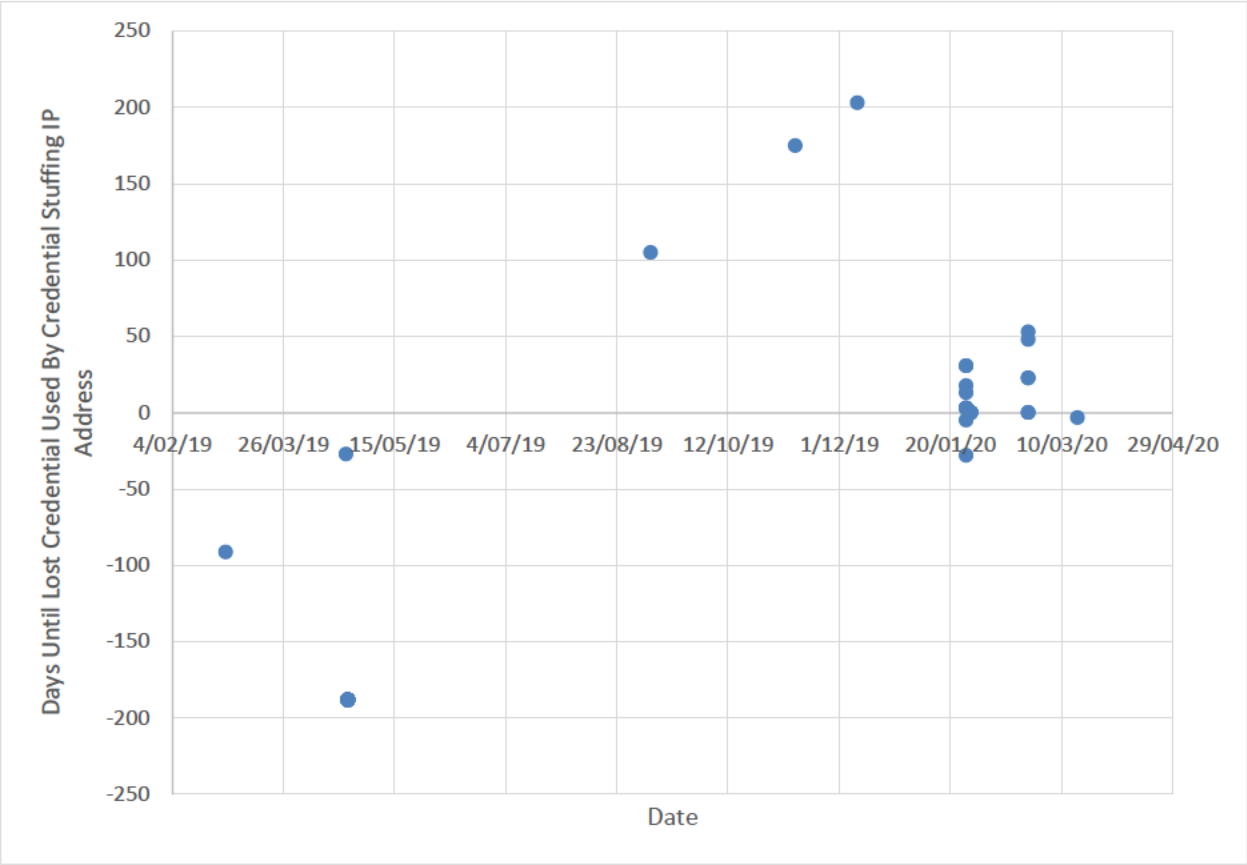


Figure 16

6.5. Credential loss

Figure 17 shows the level of credential loss per month over the capture period. It can be seen that the rate of credential loss has increased over the 18-month period.

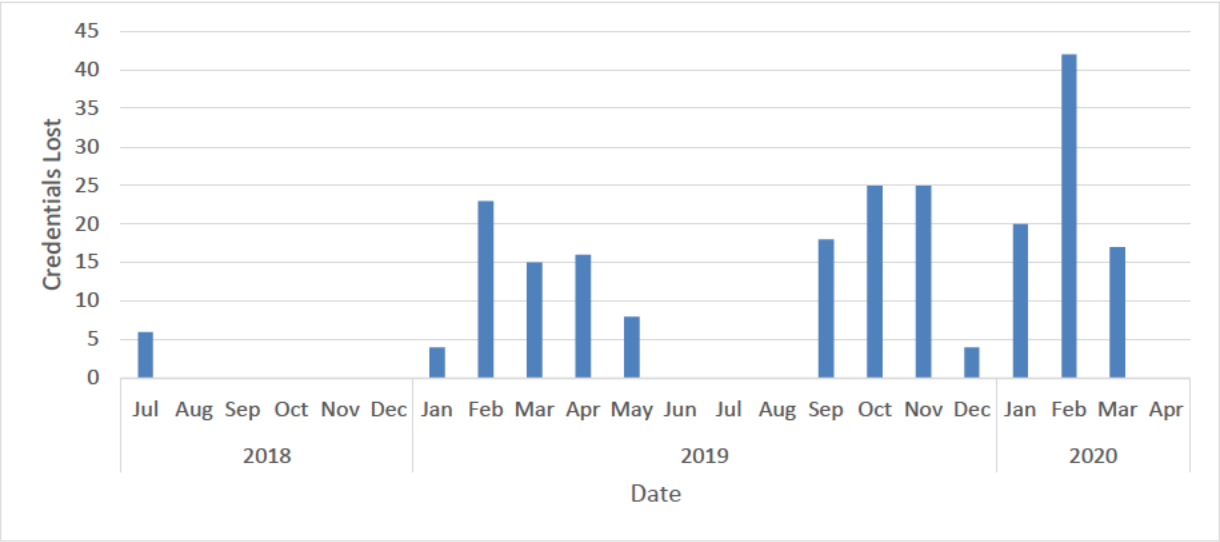


Figure 17

Credential loss was categorised into three separate reasons:

- Leaked weak passwords, undetected phishing, password breach or password guessing
- Phishing phishing detected by IT support team
- Unknown

Figure 18 shows that the majority of credential loss occurs through phishing campaigns (66%).

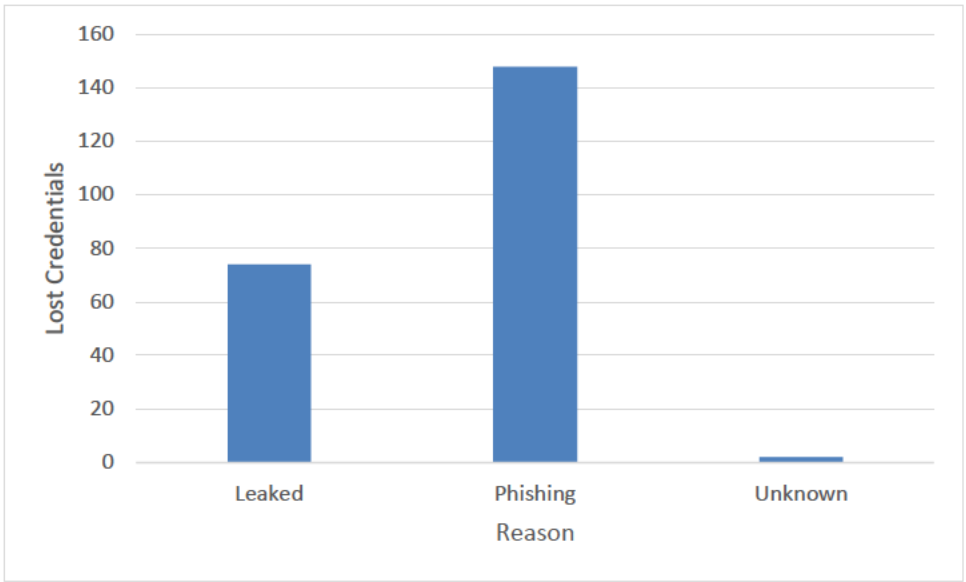


Figure 18

Of the three phishing campaign types observed the *Trusted Individual* type accounted for 89% of the lost credentials (Figure 19 below). This shows that phishing campaigns are much more likely to be successful if received from an individual that the sender knows.

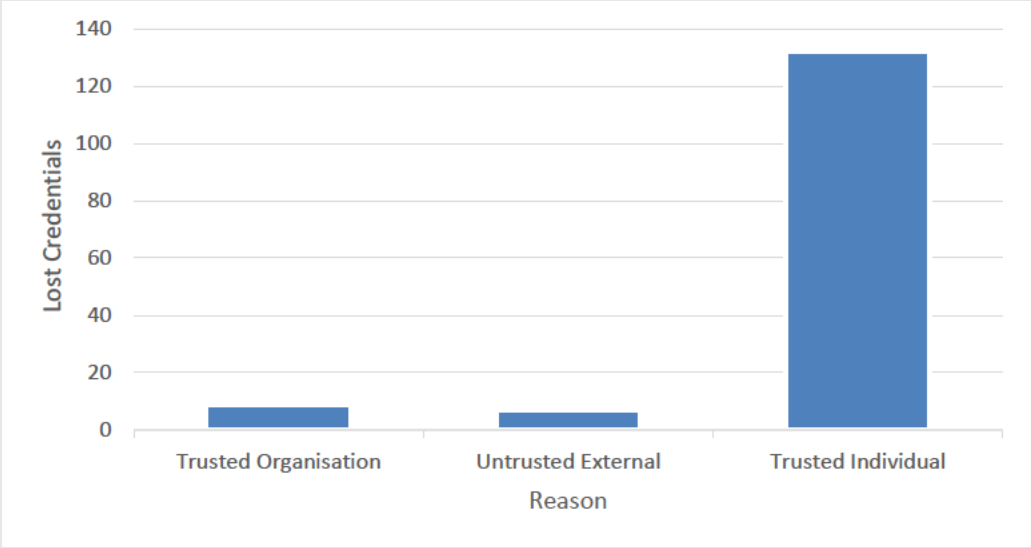


Figure 19

The graph below shows that the trusted individual type of phishing campaign has been well used over the collection period. The untrusted external type has accounted for less lost credentials

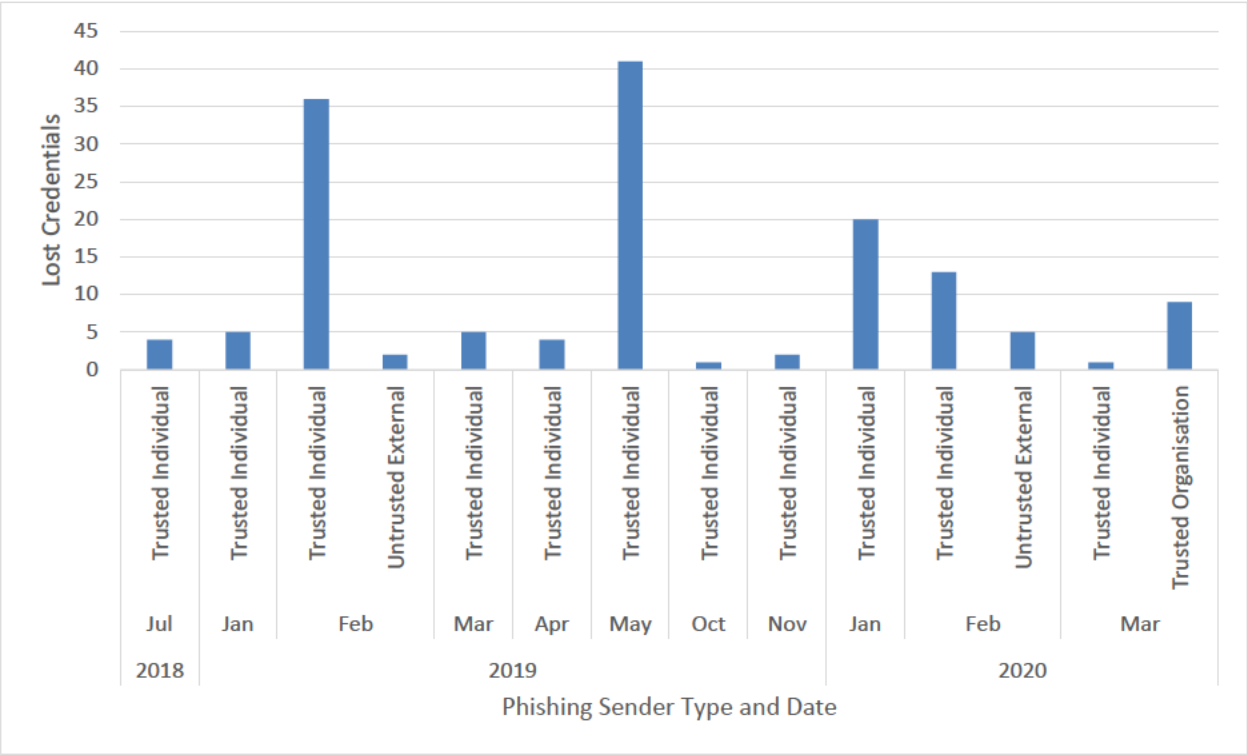


Figure 20

To answer the question if phishing techniques are changing, Figure 20 shows the distribution over time shows that the trusted individual has been prevalent through the data captures period. Trusted individual is shown to account for the majority of lost credentials in each month except for March 2020. This shows that the most effective phishing technique used is the trusted individual. Phishing campaigns that are received by an individual that the recipient knows or has communicated with previously account for more lost credentials. This trend could enable an organisation to focus phishing awareness and training not just on impersonation attempts, but also those received from trusted individuals who have themselves lost account credentials.

6.5.1. Phishing Technique Hook

The hook used by phishing campaigns can be described as the enticing element of a phishing email that compels the recipient to take action typically clicking on a link.

The two hooks used by phishing campaigns that accounted for credential loss during the data capture period were *Lockout threat* or *File shared.* 53% of lost credentials due to

phishing were attributed to phishing campaigns that used the *File shared* hook to entice users to click on links.

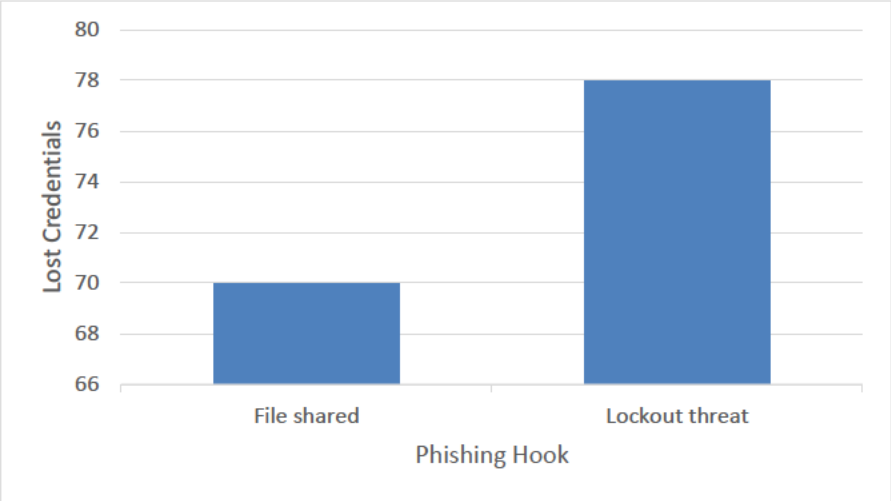


Figure 21

Figures 21and 22 show that the two main hooks used by phishing campaigns that led to lost credentials are relatively even in quantity and are well distributed over the collection period. This knowledge again allows an organisation to tailor education and awareness activities based on the successful hooks used by phishing campaigns.

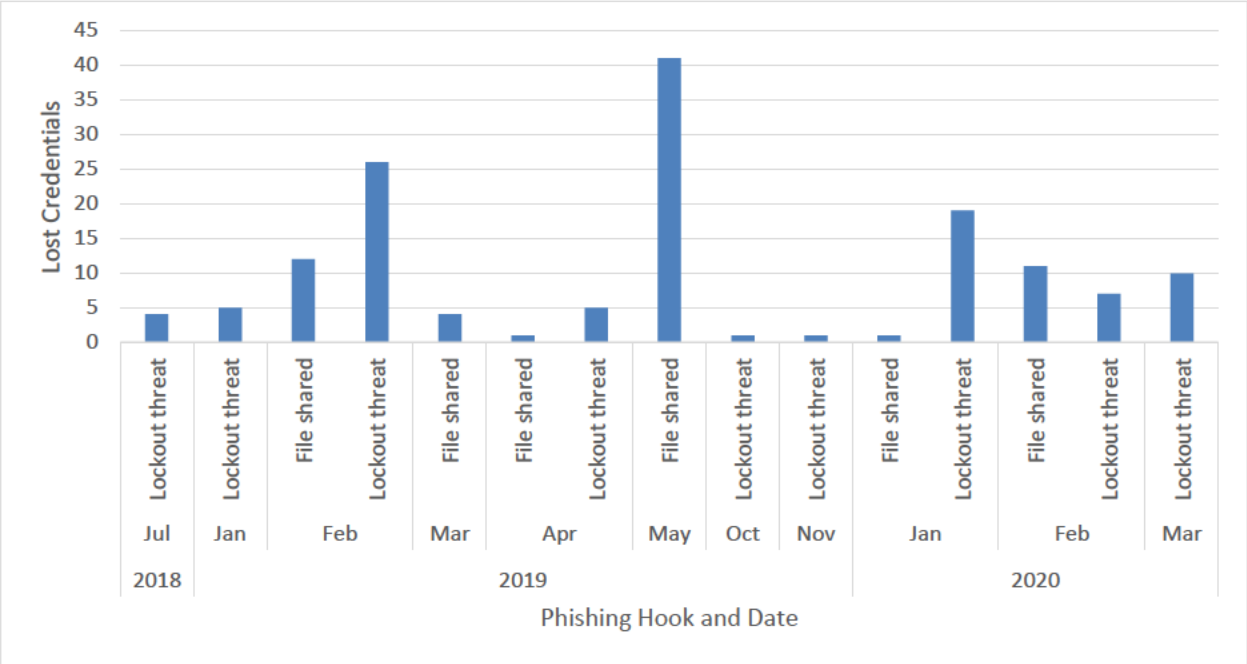


Figure 22

6.5.2. IP Address Analysis

The IP addresses used in attempts to use lost credentials may be used to profile malicious networks locations. Figure 23 shows that top locations that lost credentials were used from.

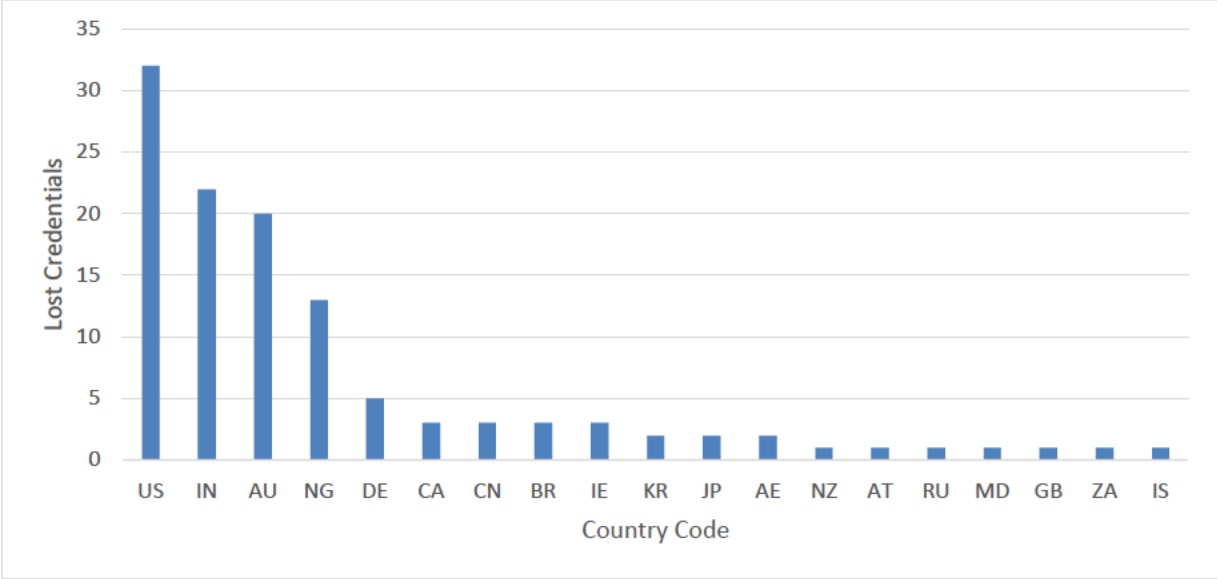


Figure 23

The distribution of IP addresses provides little in the way of patterns to determine if an access attempts is legitimate or not. However, with the inclusion of further data obtained by IP information website and network provider information, the network type can be determined. Figure 24 shows the distribution of IP addresses by country and if it is a consumer-service IP address or not.

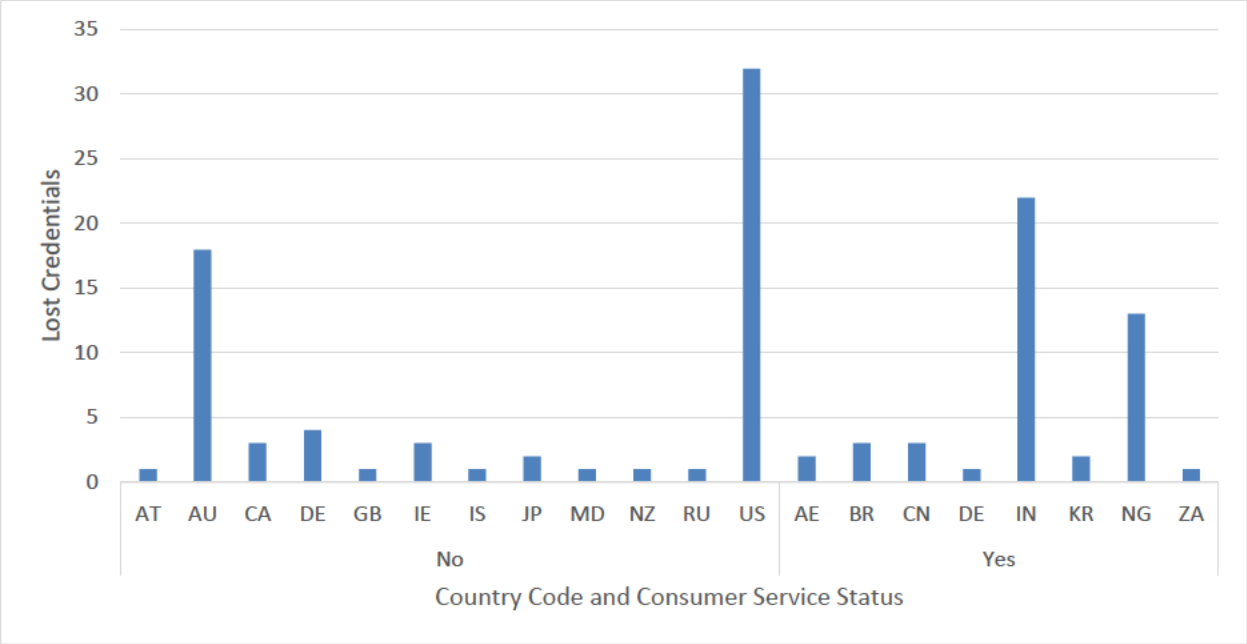


Figure 24

IP addresses that are not residential have only one overlap with countries that use non-residential services, which is Germany. A possible reason for the lack of overlap is that there are organisations that block IP addresses by location. If an authentication request comes from a country that is regularly the source of phishing, the access may be blocked or detected and repelled quickly. Malicious actors from these countries may have learned to hide their location behind anonymous proxies or VPNs to increase their success.

An organisation could use this information to identify possible suspicious behaviour. Monitoring for access attempts from consumer-service IP addresses located in a set of countries that are likely to be a threat and monitoring for access attempts from non-consumer IP address from countries that are also likely to be a threat. Being more specific about the types of IP addresses that are monitored for access attempts can reduce false

positives and allow an organisation to gain insight into where legitimate access is occurring from.

6.5.3. Repeat Offenders

The issue of repeat offenders is concerning for IT support staff in many organisations, as it shows that even if exposed to real life phishing attacks there is still a propensity for some users to fall for a phishing attack. For the data collected for credential loss, 11 of the 210 credentials were for repeat offenders. Only 1 of the repeat offenders lost their credentials 3 times, all other repeat offenders lost their credentials twice.

This has implications for the training and awareness activities conducted by IT Support staff. If there is a low occurrence of repeat offenders, then exposure to a phishing email appears to be a good method to reduce the likelihood of a phishing campaign success.

This may give support to friendly phishing activities that send simulated phishing emails to staff in an effort to make them more aware of phishing campaign techniques. However, to achieve an effective awareness level the phishing simulation must closely resemble or even replicate the actual phishing campaigns experienced by the organisation. Put more strongly, organisations should borrow most heavily from the phishing campaigns that successfully bypassed anti-spam measures and were able to obtain valid credentials from a staff member when designing friendly phishing activities.

6.6. Summary of findings

This section provides a summary of the findings arising from analysis described in the preceding parts of section six. Technical details of these findings are provided in their individual subsections. Findings in section six

- **Increased accuracy and effectiveness of monitoring for phishing** The early detection of phishing emails can greatly reduce the risk of damage from lost credentials. In section 6.1 a visible trend shows increased phishing activity during the early months of the year. If this pattern is sustained there is value in being able to

predict and cater for additional protection and detection in these times. Section 6.2 discusses the source networks where phishing campaigns were sent from. An organisation can more accurately detect malicious activity by understanding the use of non-consumer IP addresses within an organisation. Monitoring or blocking anonymous VPN usage from atypical locations can help to detect malicious access early. In 6.2.1 there is analysis of the phishing techniques that are successfully received by staff. Understanding the trends of techniques and the different content types will allow an organisation to configure its antispam system to identify these techniques and either block or notify IT support staff. Section 6.3 performs analysis of credential stuffing and its success. Valid credentials appear to be used in batches, making it possible to detect logins from multiple accounts from untrusted networks and placing IT support staff on alert that a successful batch of usernames is being utilised by a malicious party. Using this information as well as other specific location information derived from analysis in 6.3.1, 6.4 and 6.5, special attention can be paid to "hot" network locations being used by malicious parties.

- **Tailoring staff awareness programs –** The targeted training of an organisation's staff to identify and report malicious phishing emails is an effective measure to not only stop users falling for phishing emails, but also to respond quickly and protect an organisation from a phishing campaign. Findings from 6.2.2 show that some recipients are targeted more than others by phishing campaign operators. This can assist in the targeted training for those people, as well as those shown to be repeat offenders in section 6.5.3. It's also clear from analysis in sections 6.5 that there are very specific phishing hooks, and sender types that account for the majority of credential loss. Awareness campaigns such as all-staff notifications, training videos and even friendly phishing, can utilise this information to focus on the phishing campaign characteristics that are the greatest source of risk for the organisation. Finally, it is clear from section 6.5 that phishing from a trusted individual is the most effective phishing campaign sender type. Contrary to traditional education which warns staff of emails from unknown senders, more effective education can focus on reporting suspicious emails from known senders.

The findings of this analysis draw elicit thoughts of the ancient Sun Tzu saying: “know thy enemy and know yourself”. Organisations that have a good understanding of the attack characteristics specific to their organisation, and which of these their organisation is vulnerable to, can take measures both to identify these attacks and reduce vulnerabilities in a targeted way. This can bring efficiencies in the risk reduction activities as less effort is spent on measures that don’t directly address the identified threats.

7. Future Directions

There are a number of areas that are likely to lead to further benefits for an organisation protecting themselves against phishing activities. These are discussed below.

7.1. More comprehensive datasets

It’s clear from the data that phishing campaigns change fairly quickly over time as fraudsters change tactics to continue to remain effective. A dataset that extended for a longer period of time would likely be less valuable than a dataset with a shorter timeframe, but with a greater breadth.

Capturing the URLs, mail client details, time zone, sending domain and other header information within phishing emails that successfully bypassed the antispam filters might allow for the profiling of particular senders or phishing techniques that a malicious actor might be using. While some of these items can be used in mainstream anti-spam platforms for blocking emails, many might be able to be used for alerting, through monitoring and notification systems.

Capturing the usernames, protocols and client devices details of all login attempts by credential stuffing IP address would allow for an understanding of those accounts being targeted, overlap of credentials amongst IP addresses, characteristics of group sets of lost credentials and an understanding of how quickly new email addresses are included in credential stuffing account lists. All of this additional information could help a targeted organisation to gain a deep understanding of those parties that are attempting to obtain access to their systems and information.

7.2. Greater IP address and domain name enrichment

The datasets used in this study are from freely available sources. There exists a number of suppliers of commercial IP information databases, often available through APIs or downloadable databases. These databases include fields such as TOR nodes, public proxies, anonymising VPNs and known threat related IP addresses. This additional information is likely to help further profile phishing emails, and unauthorised access attempts.

There is also a significant amount of information that is related to domain names that are involved in sending phishing emails or hosting phishing sites. Further details about campaigns can be derived from collection details of sending domains, phishing URLs, and even reverse DNS lookups. This can then be further enriched with domain age, last seen date, WHOIS information and abuse reporting contacts.

7.3. Compare threat information with commercial feeds

Continuing on from the additional IP and domain name enrichment, the reputation of IP addresses and domain names derived from local systems can be compared with that of commercial threat intelligence. It would be an interesting study to discover how many locally identified IP addresses appear in commercial threat intelligence services, and how many threat intelligence services identify malicious activity that the locally identified intelligence miss.

7.4. Collaboration with other organisations

Even though this study has been primarily focussed on data obtained from attacks on a single organisation, once the data has been distilled to a number of profiles of attack methods, hook types and networks, there may be benefit in understanding the similarity of these profiles with other organisations in the same industry. It would not be surprising if threat actors targeted particular industries and there could be interesting collaboration possibilities if they did. The MISP project is one such avenue for performing this collaboration (MISP 2019).

7.5. Positive profiling

This study has focussed on the building of phishing campaign profiles from data obtained by malicious activity, which could be called, negative profiling. A valuable dataset would also be the profiling of legitimate access that an organisation sees in email transport data as well as Office 365 access, which could be called positive profiling.

These two profiles might be quite distinct for an organisation that has strict access policies (no access from anonymous services or cloud clients) and has relatively few travelling staff members. For organisations that have a wide range of staff access and travel habits there might be greater overlap.

7.6. Automated data collection and profiling

Machine learning and artificial intelligence hold promise for removing the manual tasks of profiling datasets. The automated gathering of datasets and the use of algorithms that identify patterns and anomalies may be a way for future systems to protect themselves and notify their owners. This would involve the creation of large datasets and applying machine learning models across these datasets to identify unauthorised access attempts or phishing campaigns. Understanding first how to extract data automatically and reliably from key systems within the organisation is the first step.

8. Conclusion

Phishing continues to be an issue for computer users across the internet. Many individuals and organisations are harmed by the criminal activities of phishing campaign operators. While there has been significant research and work identifying phishing emails and websites, phishing is still a very popular and successful method of obtaining account login credentials.

Spending time analysing phishing campaigns that successfully make it through anti-spam measures may allow an organisation to better target phishing campaigns specific to their organisation for blocking and training purposes. For Organisation X there were only a few techniques used by the majority of successful phishing campaigns that made it through

anti-spam defences. Simply relying on the anti-spam vendor or threat intelligence vendor to take data from the broader installation base of their product to build anti-phishing measures may leave the individual organisations exposed to phishing attacks that are specific to their organisation or industry.

Using similar techniques described in this paper to track data from phishing campaigns, an organisation can spend time and effort in creating very specific threat intelligence and training material. This can be valuable in raising awareness and reducing the risk from the activities of phishing campaign operators and those perpetrating credential stuffing.

Organisation can increase the accuracy and effectiveness of monitoring by:

- Heightening alertness during periodic cycles of phishing.
- Preparing for further successful credential stuffing from a particular IP addresses when one valid account is detected.
- Focussing monitoring on IP addresses from certain countries and service types.

Organisations can tailor their awareness programs to:

- Target regular recipients of phishing emails.
- Model actual phishing campaign techniques and hooks that have been successful in the organisation already.
- Educate staff on how to correctly identify phishing campaigns that are received from known individuals.

Office 365 is currently a battleground for many organisations, with phishers using sophisticated techniques to obtain credentials and profit from Office 365 account access. With the growth of Office 365, this trend is set to continue. While many protection measures are available, there is still much work to ensure that the valuable information held in the cloud accounts of organisations globally stays secure.

Utilising the technology resources and data at an organisation's disposal, techniques can be used to gain valuable information. The data obtained through the day-to-day work of repelling phishing attacks and unauthorised access can be used to gain insights into successful phishing campaigns, credential stuffing and credential loss. By the addition of IP

address metadata from public services, location, network providers and service types can also be used to profile sources of phishing and high-risk access attempts.

It is also desirable to make phishing more difficult for the fraudsters. Continually creating hurdles for fraudsters and repelling phishing attacks reduces the gain for the phishing campaign operator. This paper hopes to increase the effort and resources that those perpetrating phishing are required to spend to benefit from phishing. By developing creative techniques, strong protections and swift responses, individual organisations can decrease the risk of phishing attacks and credential loss while capitalising on the systems within direct reach.

References

APNIC (2020). Autonomous System numbers - FAQs - APNIC. https://www.apnic.net/get-ip/faqs/asn/. Accesses 14 May 2020.

Aonzo, S., Merlo, A., Tavella G., and Fratantonio, Y. (2018*). Phishing Attacks on Modern Android.* 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18).

Cui, Q., Jourdan, G., Bochmann, G., Couturier, R., Onut., I. (2017). *Tracking Phishing Attacks Over Time.* 2017 International World Wide Web Conference Committee (IW3C2), WWW 2017, April 3 7, 2017, Perth, Australia.

Greenberg, A (2019). *Hackers are passing around a megaleak of 2.2 billion records | WIRED*, https://www.wired.com/ story/collection-leak-usernames-passwords-billions/ Accessed 14 May 2020.

Han, X., Kheir, N., Balzarotti. D. (2016). PhishEye: *Live Monitoring of Sandboxed Phishing Kits.* CCS'16, October 24 - 28, 2016, Vienna, Austria.

Harrison B., Svetieva E., Vishwanath A. (2016). *Individual processing of phishing emails.* How attention and elaboration protect against phishing.

Ho, G., Sharma, A., Javed, M., Paxson, V., Wagner, D. (2017). *Detecting Credential Spearphishing Attacks in Enterprise Settings*. 26th USENIX Security Symposium August 16 18, 2017 • Vancouver, BC, Canada.

Laperdrix, P., Rudametkin, W., Baudry, B. (2016). *Beauty and the Beast: Diverting modern web browsers to build unique browser fingerprints.* 2016 IEEE Symposium on Security and Privacy.

MaxMind (2020). *GeoLite2 Free Downloadable Databases; MaxMind Developer Site.* https://dev.maxmind.com/geoip/geoip2/geolite2/ Accessed 14 May 2020.

Miniwatts Marketing Group (2019), *Internet World Stats, June 2019.* https://www.internetworldstats.com/stats.htm Accessed 29th July 2019.

The MISP Project (2019), MISP, MISP - a threat information sharing platform - *The Open Source Threat Intelligence Platform.* https://www.misp-project.org/ Accessed 20, December 2019.

Neupane, A., Rahman, L., Saxena, N., Hirshfield, L. (2015). *A Multi-Modal Neuro-Physiological Study of Phishing Detection and Malware Warnings*. CCS'15, October 12 16, 2015, Denver, Colorado, USA.

Michael Spencer, Satya Nadella, Amy Hood, (2019). *Microsoft FY20 First Quarter Earnings Conference Call. https://view.officeapps.live.com/op/view.aspx?src=https://c.s-microsoft.com/en-us/CMSFiles/TranscriptFY20Q1.docx?version=be2962be-e9f2-6dad-c9da-99da2c3077fa* - Accessed 2/3/2019.

OAIC - Office of the Australian Information Commissioner (2019) *Notifiable Data Breaches Report July-December 2019.* https://www.oaic.gov.au/assets/privacy/notifiable-data-breaches-scheme/statistics/Notifiable-Data-Breaches-Report-July-December-2019.pdf Accessed 14 May 2020.

PasswordRandom.com (2020), *PasswordRandom.com - Top 10000 most common passwords list Page 1*, https://www.passwordrandom.com/most-popular-passwords - Accessed 20th August, 2020.

PeeringDB (2020). *PeeringDB.* https://www.peeringdb.com/. Accessed 14 May 2020.

Shape Security, Inc (2017), 2017 Credential Spill Report. http://info.shapesecurity.com/rs/935-ZAM-778/images/Shape-2017-Credential-Spill-Report.pdf Accessed 20th August 2020.

The Anti Phishing Working Group (2019). *Phishing Activity Trends Report 4th Quarter 2018*. https://docs.apwg.org/reports/apwg_trends_report_q4_2018.pdf. Accessed 29th July 2019.

Thomas, D., Pastrana, S., Hutchings, A., Clayton, R., Beresford, A. (2017b). *Ethical issues in research using datasets of illicit origin*. In Proceedings of IMC '17, London, UK, November 1 3, 2017.

Thomas, K., Li, F., Zand, A., Barrett, J., Ranieri, J., Invernizzi, L., Markov, Y., Comanescu, O., Eranti, V., Moscicki, A., Margolis, D., Paxson, V., Bursztein, E. (2017a). *Data Breaches, Phishing, or Malware? Understanding the Risks of Stolen Credentials.* CCS'17, Oct. 30 Nov. 3, 2017, Dallas, TX, USA.

Thomas, K., Pullman J., Yeo, K., Raghunathan, A,. Kelley P.G., Invernizzi L., Benko, B., Pietraszek, T., and Patel, S., Boneh, D., Bursztein, E. (2019). *Protecting accounts from credential stuffing with password breach alerting*. In Proceedings of the 28th USENIX Security Symposium. Santa Clara, CA, USA. August 14 16, 2019 .

Vadrevu, P., Liu, J., Li, B., Rahbarinia†, B. Lee, K., Perdisci, R. (2017). *Enabling Reconstruction of Attacks on Users via Efficient Browsing Snapshots.* NDSS '17, 26 February - 1 March 2017, San Diego, CA, USA.

Vastel, A., Laperdrix, P., Rudametkin, W., Rouvoy., R. (2018). *FP-STALKER: Tracking Browser Fingerprint Evolutions.* 2018 - 39th IEEE Symposium on Security and Privacy, May 2018, San Francisco, United States.

Verizon Enterprise Solutions (2019), *2019 Data Breach Investigations Report.* https://enterprise.verizon.com/en-au/resources/reports/dbir/. Accessed 11th May 2020.

WiFiPhisher (2020), *GitHub - wifiphisher/wifiphisher: The Rogue Access Point Framework*, https://github.com/wifiphisher/wifiphisher Accessed 20th August, 2020.

Ethics Approval

Office of the Deputy Vice-Chancellor (Research)

Research Services
Research Hub, 27 Wally's Walk
Macquarie University
NSW 2109 Australia
T: +61 (2) 9850 7987
http://www.research.mq.edu.au/
ABN 90 952 801 237
CRICOS Provider No 00002.

MACQUARIE University
SYDNEY·AUSTRALIA

13/03/2020

Dear Professor Mohamed Ali Kaafar,

Reference No:52020515014436

Title: 5150 Phish Phlooding

Thank you for submitting the above application for ethical and scientific review. Macquarie University Human Research Ethics Committee HREC Humanities & Social Sciences Committee considered your application.

I am pleased to advise that ethical and scientific approval has been granted for this project to be conducted by Professor Mohamed Ali Kaafar and other personnel: Mr Jeremy Koster.

Approval Date: 13/03/2020

This research meets the requirements set out in the *National Statement on Ethical Conduct in Human Research* (2007, updated July 2018) (the *National Statement*).

Standard Conditions of Approval:

1. Continuing compliance with the requirements of the *National Statement*, which is available at the following website: http://www.nhmrc.gov.au/book/national-statement-ethical-conduct-human-research
2. This approval is valid for five (5) years, subject to the submission of annual reports. Please submit your reports on the anniversary of the approval for this protocol.
3. All significant safety issues, that adversely affect the safety of participants or materially impact on the continued ethical and scientific acceptability of the project, must be reported to the HREC within 72 hours.
4. Proposed changes to the protocol and associated documents must be submitted to the Committee for approval before implementation.

It is the responsibility of the Chief investigator to retain a copy of all documentation related to this project and to forward a copy of this approval letter to all personnel listed on the project.

Should you have any queries regarding your project, please contact the Ethics Secretariat on 9850 4194 or by email ethics.secretariat@mq.edu.au

The HREC Humanities & Social Sciences Committee Terms of Reference and Standard Operating Procedures are available from the Research Office website at: https://www.mq.edu.au/research/ethics-integrity-and-policies/ethics/human-ethics

The HREC Humanities & Social Sciences Committee wishes you every success in your research.

Yours sincerely,

Dr Karolyn White
Chair, HREC Humanities & Social Sciences Committee

This HREC is constituted and operates in accordance with the National Health and *Medical Research Council's (NHMRC)National Statement on Ethical Conduct in Human Research (2007, updated July 2018)* and the *CPMP/ICH Note for Guidance on Good Clinical Practice*