Phishing Attack Prevention: How to Identify and Detect Phishing Attacks

Thesis Proposal

by

**Suraj Pratap Samirlal Poddar**

Harrisburg University of Science and Technology

Spring 2021

Table of Contents

[Abstract 4](#_Toc70616195)

[Introduction 4](#_Toc70616196)

[Research Questions 7](#_Toc70616197)

[Significance and Justification 9](#_Toc70616198)

[Background 9](#_Toc70616199)

[Justification 11](#_Toc70616200)

[Literature Review 12](#_Toc70616201)

[Problems and Challenges 13](#_Toc70616202)

[Prevent Phishing 13](#_Toc70616203)

[Detect Phishing 14](#_Toc70616204)

[Providing Stake holder training 14](#_Toc70616205)

[Impact of Anti-Phishing Training 14](#_Toc70616206)

[Introduction 14](#_Toc70616207)

[Recent Survey 15](#_Toc70616208)

[Threat 16](#_Toc70616209)

[Summary 17](#_Toc70616210)

[Learning to Detect Phishing Emails 17](#_Toc70616211)

[Hybrid Learning based Phishing Detection 19](#_Toc70616212)

[CRI approach to address phishing 19](#_Toc70616213)

[Taxonomy of phishing Attack 19](#_Toc70616214)

[Anti-phishing Technique Modus Operandi 20](#_Toc70616215)

[Lifecycle of Phishing Attack 20](#_Toc70616216)

[Visual Similarity Based Phishing Detection and Filtering Approach 20](#_Toc70616217)

[Taxonomy of Phishing Detection and Filtering Based on Visual Similarity 21](#_Toc70616218)

[Visual Similarity Assessment 21](#_Toc70616219)

[Research Methodology 22](#_Toc70616220)

[Justification 23](#_Toc70616221)

[Target 24](#_Toc70616222)

[Interviews & Questionnaire 24](#_Toc70616223)

[Summary 26](#_Toc70616224)

[Development Methodology 27](#_Toc70616225)

[Requirements Gathering and Analysis 28](#_Toc70616226)

[Data Collection 28](#_Toc70616227)

[Data Analysis 28](#_Toc70616228)

[Proposed Plan 29](#_Toc70616229)

[Time Line 30](#_Toc70616230)

[Conclusion 30](#_Toc70616231)

[References 32](#_Toc70616232)

# Abstract

Internet has become the means to live a life which has been deeply integrated and embedded in our day-to-day life. It can be anything ranging from an online app which you visit or any social networking sites, financial sites, booking a Doctor’s appointment, ordering food, navigating to a destination, advertisements, movies and many more which makes humans life easier by accessing them in a short amount of time and mostly on the fly. Because of extensive use of Internet in daily life the technology has started to become vulnerable and therefore intense threats have been discovered following the technology. It has given birth to a new way for hackers to lure a victim into giving away their personal identity or to steal information digitally. Phishing is a form of cybercrime where a hacker/attacker for that matter pretends to be the real person or an entity by elevating them as the official owner through any communication medium, the famous of all and most popular is through e-mails, where an attacker spoofs an email or through instant messaging. The attack mainly focusses on people with no knowledge on Internet Security or to those who don’t take care heir security or privacy seriously. It can be a Bank Account or any social engineering account such as Facebook, Yahoo, Instagram, Gmail or any Online account. This paper will provide information about types of phishing attacks and steps that can be taken to shield user’s confidential information

# Introduction

Internet has changed the way we live, it has dominated many fields by increasing the comfort level for users/humans, however on contrary it has also brought few disadvantages with it, one of the most important things that comes to everybody’s mind is security, when surfing on the Internet. Everyone online takes utmost caution to protect their business by patching their server and by fixing the vulnerabilities to safeguard their business through Internet. Phishing is one of the online forms of identity theft that can be used to steal private information of an individual. Phishing scams has been receiving the utmost attention as this clearly have been on the rise, there is no single remedy that can protect a user from phishing scams (A. Alswailem, 2019)

The word ‘Phishing’ was introduced in 90s, in the earlier day’s hackers used to use ‘ph’ in replacement to letter ‘f’ and the hackers in that generation were also known as *Phreaks*. Phreaking is the term used for study, exploring, experimenting of telecom system. Since phreaks and hackers were closely linked, ‘ph’ was linked to phishing instead of fishing and since this attack lures a victim to give away their personal information just as the analogy with ‘angling’ where the attacker sets email lures by positioning the hooks to ‘fish’ for sensitive information which can include their identity information or banking passwords or any type of financial data from the ‘sea’ of users on the Internet (Gürel, Detection of phishing attacks, 2018)

One of the main goals of phishing is to conduct a fraudulent transaction by using a fake email by luring the user to click on the URL and redirect them to an illegitimate website which looks exactly the same as their banking website and steal their username and password that way or to entice a user to give away their social security number or any private information such as name, address , phone of an user which can be used to either apply for a loan or for a credit-card on behalf of the user

There are various types of phishing scams and these tend to continue as the techniques has been successful enough for cyber criminals in making hefty profits. Phishing has been around since the origination of Internet and they are not going away any time soon. There is always a chance for you to fall for these traps and therefore it becomes imperative to adapt basic guidelines in keeping yourself safe. Always keep up with the information on Phishing, think twice before clicking on any URLs, Install an Anti-phishing toolbar, Verify the certificate and the site security of the website, check your Online A/c’s and passwords regularly, rotate the passwords frequently, enable Multi Factor Authentication, never give away your private/personal information

While many methods have been developed to tackle phishing attacks none of them are efficient enough to mitigate new phishing attacks that are constantly being developed and cannot filter all the malicious email content. This paper describes about the methodology and approach along the various phases and the differences between a normal approach to a website versus a phishing attack redirect. While Phishing websites are designed to looks like the original page there is always a slight difference which one needs to understand. The Sony hack which was pulled in 2015 is a great example of a successful phishing campaign and illustrates the damage that can do to an organization, the damage done to Sony was at a bigger scale in millions of USD

Therefore, the key is to take a proactive approach by educating your users which is the most important part in anti-phishing strategy. To Design an effective anti-phishing strategy includes various factors, and is intended to improve person’s knowledge on phishing, this can be in the form of videos or simlets or printed materials. Phish tank is a website which has database of all the phishing websites, it updates its records by monitoring the entire Internet and I mainly created for user’s awareness. By submitting the URL, you can verify whether the url is a potential phishing site. Deep packet inspection is one of the Network Techniques which is used to dig into a packet level, many countries use this mechanism to surveillance on the Internet and to monitor any malicious activity over the Internet (M. Adil, 2020)

# Research Questions

Question 1: How do you classify & identify Phishing Attacks

In the following Research question, we will classify the types of phishing attacks such as Email Phishing, Smishing & Vishing, Spear Phishing, Angler Phishing, Deceptive phishing, fight Phishing, pharming etc. by explaining their techniques and the methods used to exploit or lure a victim

As lot of the users become Internet savvy the Internet has become an integral part of our life. On the contrary it has also brought lot of disadvantages and threat opportunities for Users with ill intent to perform malicious activities like Phishing. While many methods have been developed to identify and detect Phishing, one of the most successful among them is through Machine Learning. ML provides simple and effective methods for analyzing the data and has brought some promising in eliminating some real time problems. ML can be a powerful tool in the current situation where it can quickly detect any fraudulent transaction and help develop learning-based solution (Vahid Shahivari, 2020). The risk of disclosing private sensitive identifiable information is very high and therefore become applicable that a thorough research is conducted to identify and discover the outcomes associated with phishing attacks

Question 2: What steps can you take to avoid phishing scams

In this, we are going to discover how phishing scams are being developed all the time and in order to stay on the top of this, the first step towards is to educate and spread awareness which can be carried out through security awareness training. By providing awareness we can have our users take the responsibility to think before clicking on any random links.

Phishing has been defined as social engineering which tends to steal information through guise of a trusted source (Jakobsson, 2007) With over a billions of Users currently using Internet all over the world, there are not a less number of potential victims to achieve identifying information from, Social Engineering is further defined as an ‘old way of conning’ with efforts put in place to gather personal identifiable details to launch a successful attack (Easttorn, 2011) The following research uses a real time approach to gather data which leads to phishing scams. Knowledge/Educating users is one of the important factors to prevent a user from falling to phishing or social engineering attacks. There are lot of training videos out there which provides a user with ammo and provides them with an opportunity to explore and understand the response to a phishing attack

Question 3: Anti phishing techniques and best practices

This question will help in determining what software’s/network devices can be installed on the perimeter and inside the organization to protect from these attacks like installing Firewall, enabling IDS/IPS, by keeping the browser updated, by keeping the virus signatures updated on the Antivirus client, by checking Sites Security and its certificate, by using reverse proxies etc (Chhikara, 2013)

Providing with the right tools and the necessary knowledge to a User to not fall for scams becomes very important. Phishing is a cybercrime however crimes related to it can be difficult to investigate to get to a verdict due to number of reasons such as Delay in crime reporting, temporary phishing sites which are brought up and down on a regular basis in just a short amount of time deleting any trace related to the crime etc. Most of the Users often do not report these scams if their financial loss is low or negligible (White, 2008) While Phishing scam websites are taken down in a very swift manner research (Moore, 2007) show it may have number of potential victims before taking down the website (Chhikara, 2013) The research will provide valuable insights to inner and knowledgeable response that appears when a User experiences a Phishing attack. There might be many reasons as to how and why a user chooses to fall for a phishing scam, even after knowing the consequences, knowing this will provide insight related to the response to phishing attacks and will help develop training sessions to further assist in preventing an individual from falling these attacks

# Significance and Justification

## Background

In the past few years, the amount of unique phishing emails has exponentially increased, Phishers can steal sensitive data from users to access an individual’s financial information and initiate a illegitimate wire transfer or withdraw money without the users consent, this has put the organization and the individuals both at risk. Due to these consequences, most of the companies have now started to educate their employees about different types of phishing through security awareness and various training program and have implemented a way to inform their IT department of any phishing email s they receive. The training provided by the organization can be limited for a short time and therefore the long-term benefit of Phishing attack prevention remains a question. Fascinated by this challenge, this study/research targets to determine the stretch one can experience by using various techniques to keep the user engaged in preventing Phishing attacks (APWG, 2003)

Chart, pie chart

Description automatically generated

Chart, bar chart

Description automatically generated

## Justification

The main objective of this paper is to propose a study/research to assess the concept model of how to encourage individuals to participate which will help a user in long run to identify and prevent phishing attacks at the same time by improving phishing reporting. The research will take a mixed approach and will include various measures or methodologies developed to tackle Phishing scams. Furthermore, a research is going to be performed on various phishing software platforms and browser addons to gauge the tools capabilities and use them in the real-world shielding user’s sensitive data. The outcome of this research will have potential to be published in a security research journal and can potentially be a funding initiative by NIST as this is an important area of research for Cyber security. This research will help explore the emotions of users falling for these phishing scams as well as provide training to make them realize the negative side of providing their private information and the consequences related to it

# Literature Review

This research is fairly new and is subject to change as the new threats continue to appear. The research focusses on different types of Phishing scams that have currently gained traction, understand how a user falls for these scams and how to identify and respond to the scams. By providing a theory related to phishing scams with existing research will not help in expanding the research in general for Phishing but also provides an opportunity to explore more about the phishing communications and what can be done to bring awareness to the Internet users

Over the years, given the fast adoption of latest technological advancements, security exploitation/attacks have significantly on a rise topping the human vulnerability. Phishing has become one of the threats to cyber world creating a massive damage of worth billion dollars every year. Phishing can be performed in different methods involving web, email, messages etc. Over time there has many articles describing the techniques and new attacks based on phishing however, they have failed to outline all the risk factors involved and provide and efficient/effective solution to mitigate this form of a threat. In this section we will be comparing different data from various sources to perform a literature review. Emails has been always one of the main online application which is used extensively by lots of users, business, Govt and different organization for communicating with one another and sharing data and therefore phishing emails has always shown us a serious threat to digital commerce as they are used to scam both individual and financial institutions (A. N. Shaikh, 2016)

# Problems and Challenges

Hacker/phishers have become advanced and more talented/skilled at recreating/forging sites to appear identical including recreating logos and graphical presentation. No doubt Phishing is evolving to tackle the defenses and bypass the detection. In these circumstances I think a sequential approach is needed which will help mitigate these kind of attacks

1. Prevent Phishing
2. Detect Phishing
3. Provide Stake holder training

Diagram

Description automatically generated

Figure 1 Step by Step process

### Prevent Phishing

Phishing can be prevented by taking extra precautions and blocking it before it reaches the users, this includes blocking or blacklisting the phishing site by filtering emails. The first procedure is to identify the URL either manually or by ML, although this can help catch few sites, it cannot catch them all

### Detect Phishing

This method is deemed to be one of the effective methods as this can identify the link to phishing sites through the spam filters which are majorly used by the email servers. Phishing filters are constructed by Machine Learning techniques. Most of the browsers already have an extension installed on their browser which has a passive or an active indicator warning the users and alerting them at the same time

### Providing Stake holder training

By providing training to all the users can help avoiding the users clicking random sites and falling into Phishing scams. Most of the organizations create internal training program to combat against advanced phishing attacks which is often referred as anti-phishing techniques/methods. One of the famous anti-phishing games is Anti-Phishing Phil which is a micro game that helps a user identify suspicious URLs and spread awareness also provide a detail view of the consequences (Jampen, 2020)

# Impact of Anti-Phishing Training

### Introduction

There are many publications that provide insight on how a user will become a victim to phishing scams even after providing them with various anti-phishing techniques or through an integrated workflow training. The training is basically where an automated system sends a set of phishing emails to users, the users have to identify and report them to respective team during a typical workday. These types of training are often referred as embedded training, they are designed to be more realistic and take a practical approach. The accuracy of Anti-phishing techniques is not always effective against all type of threats. While the studies show the accuracy of Anti Phishing is a critical factor in reduction of successful works of phishing attacks there are other things like enhancing user’s knowledge, improving tool etc. which has a strong implication for an individual as well as to the organization

### Recent Survey

In (Neupane A, 2015) one of the studies conducted in identifying a real website vs a phishing website at neuro physiological level governing the human processing shows how we take a 3-dimensional approval to detect phishing and how a user thinks about these tasks. As per the multi modal neuro physiological measures, users do not like to spend additional time analyzing the important indicators and therefore fail identifying these attacks. They may still have or think about the task however can subconsciously proceed to phishing sites when compared to a real site

One of the most recent work by one of the authors in a retrospective carried out in 6 health care institutions (Kumaraguru P, 2007) show there were approx. about 95 campaigns that were run from 2001 to 2018 with a whopping 3 million phishing emails by sending them to users of their respective organization. (Carella A, 2017) The universal clicking of email were relatively high, however varied per institution, on an average there were about 14% of those simulated emails which were clicked by their employees during a workday. By Conducing numerous phishing campaigns have reduced the odds of potential clicking on possible phishing emails, the odds of phishing emails went down by 0.5 lower for about every 4 to 7 campaigns. They have also found difference in click rates among the institutions (J, 2017)

There are various ways how phishing scams can be designed, they are most commonly designed in one of two ways, the first one is by targeting or triggering an individual greed and the other type is to manipulate a user by sending threatening communications. Sentences/phrases such as “act now, its urgent” failure to comply will result in permanent suspension etc are used to make it look like you need to take an immediate action or face consequences. One of the examples is sending a communication related to your email storage, which typically says, you have exceeded your 50Gb of storage and threatens the recipient that they will not be able to send/receive emails or that their Account will be revoked. Users don’t think twice about providing log-in information.

### Threat

A threat is inclined to cause harm or loss to another person, in this scenario phishing threat is to fraudulently steal personal identifiable information from an individual user or an organization or to inject a malicious software by acting as a trusted entity. Threats are commonly delivered by visiting unknown websites or delivered via emails which can include any type of advertisements on genuine sites that have been potentially exploited. One of the most common form of phishing related deception is Stealing Identity (Bakhshi, 2008) where the attacker targets a victim by gathering private information to socially engineer or to tailor fit the attack

### Summary

The use of various training programs helps enhancing both user and enterprise awareness towards phishing scams providing new insights to prevent phishing and eventually stopping users to give away their personal identifiable information or sensitive information such as usernames password or their financial details. Phishers are adopting new tricks to lure and confuse victims, it can be in the form of a hyperlink, website URL encoded or having the victim click on malicious links and by redirecting them to a phishing website to install malicious code not the host machine or to append victims machine host file to change their local mappings

# Learning to Detect Phishing Emails

Even though there are few advantages to filter a phishing attack delivered in the form of an email, there are not that many methods tailored to concentrate on phishing emails as opposed to spams in general. The closely related option is to create policies on the smtp server based on the classification of what we call it as a good email vs a phishing email. As the attacks improve, we expect there will be new features that will be added which will help in identifying and combining the information form internal as well as external resources. This approach is used to detect phishing websites, or the content sent via email to lure a user. While most of the email gets filtered it is very hard to accurately differentiate, although 90% of the phishing emails are blocked there is still a 10% chance where the user can always fall into the trap

The prosaic approaches for phishing detections have always been at a lowest accuracy, machine learning provides us a better outcome for phishing detection, however, are time consuming and not a match at scale. There are various methods that can help us in Phishing detection (Khonji, 2013)

Machine Learning Based Phishing Detection

Diagram

Description automatically generated

Figure 2 Machine Learning

To tame ML and make it effective in detecting phishing attacks is to collect data at hand and they must have features related to legit websites and phishing websites. Figure 2 Machine Learning shows the working model of ML for phishing attack detection. A stream of data is provided as input to train ML model to help predict the phishing attacks and legit traffic (Gupta, 2016)

Author in El Asssal et al. introduced a conventional structure which is named after phish Bench, which enables us to not only analyze but also assess the current detection techniques by understanding the framework and performance measurements. The tests show us that the classification dropped when the portion dropped among the authentic and phishing decreasing from 1 to 10

### Hybrid Learning based Phishing Detection

The technique compares real and phishing with counter strategies like tokenization, word/text parsing and by stopping word evacuation, this type of approach has been found to be accurate against existing techniques (Abawajy, 2013)

### CRI approach to address phishing

This type of approach focuses on a solution resulting in a holistic anti phishing review. All the approaches primarily focus on how to prevent the attack without actually thinking about the crime and the respective problem associated with it, the following literature shows that providing solution without knowing/exploring the actual issue is not the right way to tackle the threat. As a reason to which a CRI approach was implemented to explore the crime factor, review mitigation techniques and investigate the missing pieces. The major goal is to provide a future researcher with the information at the same time strengthen the literature review. Crime has increased with the expansion of IT, IoT and any form of digital devices that can talk over the Internet where cyber criminals have got their hands on to exploit in the real world (Purkait, 2012)

### Taxonomy of phishing Attack

Attackers use different techniques for phishing. In social way they try to attract by sending emails to provide bank information, credit card and convenience them to respond. They might also include links to open which can directly hack your system and attackers can also remote to victim’s computer to collect all the data.

### Anti-phishing Technique Modus Operandi

After receiving phishing emails anti-phishing techniques are also implemented either by redirecting them to SPAM folder or by showing a warning when user tries to open the link (Vayansky, 2018)

### Lifecycle of Phishing Attack

* Attackers try to create a fake copy of organization and sends URL to all Internet users using email id & social networking sites.
* In case of fake emails there are different solutions to block them based on structure features of mails so that fake email ids are not blocked.
* If fake email id works and user tries to open link, there will be some browser techniques based on client end.
* If it by passes through all the blockages, then attacker can steal all the data from user.

### Visual Similarity Based Phishing Detection and Filtering Approach

A user will not understand phishing attack as it will look similar to legitimate sites by seeing high standard URL’s layout, Images, fonts. Fake and original sites look exactly same except the URL’s are not same, but user may not check URL’s before clicking the link. If attacker doesn’t match with the original website, then there are chances of users using it are very less. (Basit, n.d.)

An attacker can fool user by following

1. Visual Appearance: Phishing website will exactly look like original website as attackers will copy the HTML source code to look same.
2. Address Bar: They can also cover the address bar by using any image so that users will not understand.
3. Embedded Objects: Use of different images, text to hide the content from phishing detection techniques.
4. Favicon: Every website will have a unique favicon image and attacker will try to copy same image if not user can easily identify fake website.

Based on different surveys 90% of users couldn’t identify fake websites by seeing the visual appearance and content. Even experienced users cannot understand and most of them doesn’t even look at the address bar because of this most of the users are trapped by attackers. (El Aassal, 2020)

### Taxonomy of Phishing Detection and Filtering Based on Visual Similarity

Different approaches have been proposed to detect phishing attack based on user’s education and software. An educated user should always check address URL before clicking any link and software-based approach like black-list will have all phishing URL’s but fails as it’s difficult to update new websites that are created every day.

### Visual Similarity Assessment

Based on approach there are two modules in which first module will detect the suspicious URL’s and other module will detect the layout, content. In Block level each block is compared with original website and then it shows matching blocks. Layout level is dividing the total number of matching blocks to all blocks in original site. Overall style is calculated by style feature and similarity if threshold is more than legitimate website then it is considered as phishing webpage

# Research Methodology

As mentioned earlier, Phishing attacks have become more dangerous in the recent past. Hackers are privy to the latest technology and use the same in an unconventional manner to trick the users into revealing their personal information. Although there are several training programs to help users understand and prevent phishing attacks, most provide temporary advantage as users gradually forget the same and fall prey to phishing attacks in a latter period. There is also the fact that phishing attacks are also developing, which warrants the need for frequent training. In this research paper, we are looking at how users can be better equipped to identify and prevent phishing attacks. If more individuals participate in training and use the knowledge efficiently, more users would be able to prevent phishing attacks. This can also help them report phishing attacks immediately, which is one of the main challenges that hinder legal actions during a phishing attack. The research methodology is focused on conducting user interactions via a qualitative approach to ensure better understanding of the topic.

The main goal of the research methodology is to understand the psychological aspect of phishing attacks. The qualitative research methodology that is used will be able to support the surveys mentioned earlier in the research paper. The qualitative research will be focused on understanding user reactions during and after a phishing attack to learn more about the psychological aspects of the phishing attack. Gaining the understanding of the psychological behavior of users during a phishing attack will help us look at the topic from a unique perspective. This will also help the organization create better training programs that can help the users in the long run (Asfoor et al., 2018).

# Justification

The qualitative research methodology was chosen because it provides opportunities for direct user engagement, which is necessary to gain a different perspective on the topic of phishing attacks. The research methodology will also add to the neurological survey about phishing attacks, as it will provide an in-depth research from a new angle of user psychology. This will improve the usability of the research for the organization. Another important factor for using the qualitative research methodology is that it will make it easier to collect information from users directly. By engaging with users, the organization can gather more data to better target the users who are more likely to be vulnerable in case of phishing attacks. The qualitative research paper is intended for the corporate users as well, because there are many organizations that already use the qualitative research methodology. This is the main reason for conducting a qualitative research on phishing attack identification and prevention (Conway et al., 2017).

The purpose of the qualitative research is to add value to the research paper through its findings. The research will be done to analyze the user behaviors and their interaction with phishers through the research paper. The research will be able to compare those behavioral patterns on the different phases of the phishing attack, such as the communication phase, in that they will be able to determine which stages are the most vulnerable in the process. To ensure the same, the qualitative research will be conducted via in-depth interviews and questionnaires. This will allow for an in-depth analysis on phishing attacks, especially the types of phishing attacks that are common and which are not. By doing this, the qualitative research paper will have a clear and detailed view on how users were tricked by phishing attacks. This, in turn, can be used to create training programs that can have major impact.

# Target

Primarily, the qualitative research methodology will be focused on individuals who had previously fallen victims for phishing attacks. The purpose of the same is to get in-depth information about the situation and the psychological status of the individual during and after the phishing attack. The interview will be focused on finding more information that can aide the organization in their phishing prevention and training methods. By focusing on phishing victims, the research paper will be able to learn lessons on phishing as an action to make phishing prevention even more effective. This will be of vital help to both the organizations and researchers working in this field. As mentioned earlier, the purpose of the paper is to increase user participation in phishing training and enhancing phishing training to help users retain the information longer and report phishing activities immediately. To ensure the same, the qualitative research will purposely look for individuals who have previously resulted phishing attacks and conduct interviews from them. The purpose here is the reason for reporting and the reporting timeframe. Both these factors are important for the usability of the phishing training and the reduction in phishing activities in the short-term (Md Ilyas, 2017).

# Interviews & Questionnaire

The qualitative research methodology will be comprised of both interviews and questionnaires. The main purpose for including both the methodologies is to segregate the target base and make the research more focused on the topic at hand. The questionnaires will help identify individuals who have fell victim to phishing attacks. The focus of the questionnaire is to identify the relevant targets for in-depth interviews. The questionnaire will also help in defining the specific phishing behavior to be investigated on that specific target. For the first part, the questionnaire will contain generic questions based on phishing and have a defining question that will be focused on whether an individual suffered from phishing attacks. By including generic questions, the research will be able to attract a broader range of individuals, which is necessary for the success of the research. On the other hand, this will also help individuals understand the nature of phishing attacks and help them realize whether they have fallen victim for the same. This is important as phishing attacks are often reported late or not reported at all because individuals failed to identify the phishing attack. Since the questionnaire broadly identifies the nature of phishing attacks and the methodologies used, individuals will be able to better understand the same and identify whether they have fallen victims to phishing attacks (Ghazi-Tehrani & Pontell, 2021).

The second part of the qualitative research will be focused on in-depth interviews. The interviews will be conducted to individuals who have passed the initial screening questionnaire to ensure that the research stays focused. The purpose of the interview is to understand the victims and identify their reactions during and after the phishing attack. Hence, the interview questions will be related to the victim's actions during the phishing attack. This will help to identify the victim's reactions when it comes to phishing attacks. By investigating the victim’s experiences and analyzing the same, the research paper will be able to gain much-needed knowledge about the victim’s psyche during the attack. The interview will also be focused on further investigating the response of the victim and the result of the same. A clear understanding of the victim’s response and the time it took for the victim to identify and report the phishing attack will enable the organization to better assess the success or failure of phishing attacks. This will further strengthen the phishing training program immensely (Ferreira & Marques, 2018).

# Summary

The focus of the qualitative research methodology is to look at phishing attacks from the victim’s perspective and help establish the victim’s psychological activities during the same. This is important for seeking solutions to the core research purpose. The purpose of the research paper is to help individuals prepare themselves better, so that they can efficiently identify and report phishing activities. Although there are several training programs that is focused on the same, most of them lose effect in a short period as individuals fall victim to phishing attacks later. Therefore, the main goal here is to create an enhanced training program that could help individuals gain sustainable knowledge and efficient understanding of phishing attacks and helping them report the same instantly. This can only be done when the research paper understands the psyche of the victims and the time it took for them to identify and report the phishing attacks. With the help of the questionnaire and the subsequent interviews, I believe that the qualitative research can help create this better training program (Conway et al., 2017).

Another important aspect is that the qualitative research will provide significant discoveries that can help researchers understand the roadblocks for phishing identification and reporting. This, coupled with the neurological surveys, can form the base for future research into phishing and the phishing attack methodology itself. It will facilitate their efforts to gather additional information about the victim during their attacks and it can assist research on the victims' psychological reactions during the phishing attacks, which will enable them to identify and report the same in a more reliable manner. Therefore, this quantitative research will be a huge help for phishing researchers to understand the attack methods and help them improve their response and defense to the attack. To conclude, I believe that the qualitative research will be critical in helping identify the solutions for the research questions and help move the research forward. This is the main purpose of the qualitative research methodology (Asfoor et al., 2018).

# Development Methodology

The following paper will explain the research methodology and justification to why they are preferable, the methodology has six phases which consists of Requirement Gathering, System Design, Implementation, System Testing, User Evaluation and Delivery

Diagram

Description automatically generated

## Requirements Gathering and Analysis

The following was covered in the above two sections while illustrating outlined objective and literature review. This has helped in clear definition of analysis and problem to achieve these objectives

## Data Collection

The Data Collection involved collection of phishing email s for testing, the phishing email were taken from online organizations that have reported phishing attacks, for example (Berkeley Information Security and Policy , 2017) Gathering all the information and their respective key process while identifying emails. This helped design an algorithm from various findings of wide range of phishing emails by ensuring the algorithm met the criteria required to prevent phishing emails

At the end of the training program if the source data is found to be not heuristically trustworthy, the browser will have built in feature to notify/alert the user on clicking on a potentially dangerous URL, if the source data is found to be heuristically trustworthy either because this was already recorded from user’s interaction history or either because it is reputed then the system operation coded through this message is going to be monitored as the data can be altered.

## Data Analysis

The anti-phishing techniques will be evaluated in the following context, can the anti-phishing technique efficiently and effectively prevent existing phishing attacks or the new phishing techniques which are on a constant rise? Does the training distract users from their routine work? To find out we can use the APWG archives to recreate new phishing attacks with novel tricks to bypass the heuristic prevention and detection rules to protect. In terms of usability, we will see how users/an individual react towards our solution and what is their attitude towards it, by integrated visuals the training will look like it is being carried out in real time and finally detecting how many legit sites are flagged dangerous when the users first interact. The implementation of detecting and visualizing will be to analyze an iterative design loop.

Based on the findings, the goal will be to identify methods through which anti-phishing techniques can be made more credible based on perceived quality. When an user visit a website in the training program the tool bar will provide with recommendations against clicking a malicious Url versus a legit url based on specific cues available on the site. While the system is providing recommendation or advise, the users assessment will play an important role to be a factor in determining whether the user follows the advice

It has been found that users often try to dismiss recommendations or advice provided by a system, since they consider themselves as tech savvy and pretend to know everything on the Internet.

# Proposed Plan

The above Research study will be conducted over a period of 4 months. Once the required information has been collected the results will be posted in accordance to experiment and analysis. The proposed paper will discuss issues and problems related to phishing while we take a deeper dive in the analysis as mentioned above. The above data will be backed up with practical analysis and an effective education tool to spread awareness

# Time Line

Apr 29th 2021

Submit Thesis Proposal

Aug, 2021

Start the design implementation

Interviews and Questionnaire

User Study for efficient and effectiveness of different integrated visualization

Sep, 2021

Evaluate interviews and Questionnaire collected from the users

Design and implement effective training for users based on user study

Evaluate and choose improvised version of phishing detection framework

Nov, 2021

Final Evaluation of Phishing attack Prevention

Submission GRAD699

# Conclusion

Phishing attacks have clearly evolved over time and pose a tremendous threat to individual or to an organization. Phishers tend to exploit human vulnerabilities by leveraging technological vulnerabilities in the system. Various phishing techniques are on rise, besides use of social media-based phishing has increased in use parallel to advancement/growth of social media. Concurrently Phishing has developed above and beyond gathering just sensitive information, it has come to a point where it is used for cyberterrorism, damaging reputations, nation-state attacks etc. The following thesis has dived into the problem deeper and have come up with a proposed solution which can be taken by an individual or an organization to avoid phishing

Although, educating humans is the most efficient and effective way to tackle phishing, it is very difficult to eradicate the threat permanently due to the constant change in the attacks and different techniques. An ongoing training program can help contribute in security awareness training to reduce its impact, by developing and taking the right measures/by following anti-phishing techniques which acts as a defense to users for phishing scams is one of the prime steps in mitigating phishing attacks.

Further research will be needed to identify and investigate susceptibility among the users, this will help in designing a stronger defense or anti-phishing security systems. Law and Legislation made for phishing are still in the early stages, there is no law for phishing in few of the countries, most of them are considered under criminal laws like an identity theft or a computer crime. Making a specific law for phishing is imperative and will help in mitigating these attacks i.e, by determining the source and the stage of phishing cycle and by enforcing specific law legislation on the phisher can help in restricting phishing scams drastically

# References

(2017, April 24). Retrieved from Berkeley Information Security and Policy : https:?/security.berkeley.edu/resources/phishing/phishing-examples-archive

A. Alswailem, B. A. (2019). Detecting Phishing Websites Using Machine Learning. *2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia*, pp. 1-6, doi: 10.1109.

A. N. Shaikh, A. M. (2016). A literature review on phishing crime, prevention review and investigation of gaps. *10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA), Chengdu*.

Abawajy, R. I. (2013). A multi-tier phishing detection and filtering approach. *Journal of Network and Computer Applications, vol. 36, no. 1*, pp. 324–335, 2013.

*APWG*. (n.d.). Retrieved from APWG: https://www.antiphishing.org

*APWG*. (2003). Retrieved from APWG: https://www.antiphishing.org

Asfoor, A. R. (2018, June). *Factors influencing information security awareness of phishing attacks from bank customers’ perspective: A preliminary investigation. In International Conference of Reliable Information and Communication.*

Asfoor, A. R. (2018, June). *Factors influencing information security awareness of phishing attacks from bank customers’ perspective: A preliminary investigation. In International Conference of Reliable Information and Communicatio.*

Bakhshi, T. P. (2008). *A Practical Assessment of Social Engineering Vulnerabilities”, Proceedings of the 2nd International Conference on Human Aspects of Information Security and Assurance.*

Basit, A. Z. (n.d.). *A comprehensive survey of AI-enabled phishing attacks detection techniques*. Retrieved from Telecommun Syst 76, 139–154 (2021). : https://doi.org/10.1007/s11235-020-00733-2

Carella A, K. M. (2017). *Impact of security awareness training on phishing click-through rates. In: 2017 IEEE international conference on Big Data (Big Data).*

Chhikara, J. (2013). Phishing & Anti-Phishing Techniques. *Case Study International Journal of Advanced Research in Computer Science and Software Engineering*.

Conway, D. T. (2017 pp. 115-129). *A qualitative investigation of bank employee experiences of information security and phishing. .* In Thirteenth Symposium on Usable Privacy and Security .

Easttorn, C. &. (2011). *Computer Crime, Investigation, and the Law. Boston, MA: Course Technology.*

El Aassal, A. B. (2020). An in-depth benchmarking and evaluation of phishing detection research for security needs. *IEEE Access, 8, 22170–22192*.

Ferreira, A. &.-2. (n.d.).

Gürel, M. B. (2013). Detection of phishing attacks. *2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, 2018, pp. 1-5, doi: 10.1109/ISDFS.2018.8355389 Chhikara, Jyoti. (2013). Phishing & Anti-Phishing Techniques*.

Gürel, M. B. (2018). Detection of phishing attacks. *6th International Symposium on Digital Forensic and Security (ISDFS), Antalya*, pp. 1-5, doi: 10.1109/ISDFS.2018.8355389.

Ghazi-Tehrani, A. K.-3. (n.d.).

Gupta, A. K. (2016). Comparative analysis of features-based machine learning approaches for phishing detection. *10th INDIA-COM, New Delhi, India*.

J, S. (2017). *Persistent training. In: Advanced persistent training, Apress, Berkele.*

Jakobsson, M. &. (2007). *Phishing and Countermeasures:Understanding the Increasing Problem of Electronic Identiry Theft.* Hoboken, New Jersey.

Jampen, D. G. (2020). *A comparative literature review. Hum. Cent. Comput. Inf. Sci. 10, 33* . Retrieved from https://doi.org/10.1186/s13673-020-00237-7

Khonji, M. &. (2013). Phishing Detection: A Literature Survey. *IEEE Communications Surveys &amp Tutorials*, PP. 1-31. 10.1109/SURV.2013.032213.00009.

Kumaraguru P, R. Y. (2007). *Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. In: Proceedings of the anti-phishing working groups 2nd annual eCrime researchers.* Pittsburgh, PA.

M. Adil, R. K. (2020). Preventive Techniques of Phishing Attacks in Networks. *3rd International Conference on Advancements in Computational Sciences (ICACS), Lahore, Pakistan, 2020*, pp. 1-8, doi: 10.1109/ICACS47775.2020.9055943.

Md Ilyas, S. S. (n.d.).

Moore, T. &. (2007). *Examining the Impact of Website Take-down on Phishing. Proceedings of the Anti-Phishing Working Group 2nd Annual e-Crime Researchers Summit.* Pittsburgh.

Neupane A, R. M. (2015). *A multi-modal neuro-physiological study of phishing detection and malware warnings. In: Proceedings of the 22nd ACM SIGSAC conference on computer and communications security—CCS ’15. ACM Press,.* Denver, Colrado.

Orunsolu AA, S. A. (2017). *An empirical evaluation of security tips in phishing prevention: a case study of Nigerian banks. Int J Electron Inf Eng.*

Purkait, S. (2012). *Phishing counter measures and their effectiveness – literature review*. Retrieved from Information Management & Computer Security, Vol. 20 No. 5 pp. 382-420.: https://doi.org/10.1108/09685221211286548

Vahid Shahivari, M. M. (2020). *Phishing Detection Using Machine Learning Techniques.*

Vayansky, I. &. (2018). Phishing – challenges and solutions. *Computer Fraud & Security* , 15-20. 10.1016/S1361-3723(18)30007-1.

White, M. &. (2008). *Assessing Our Knowledge of Identity Theft: The Challenges to Effective Prevention and Control Efforts. Criminal Justice Policy Review.*