Social Engineering

A Project Work Synopsis

Submitted in the partial fulfilment for the award of the degree of

BACHELOR OF ENGINEERING
IN
COMPUTER SCIENCE WITH SPECIALIZATION IN INFORMATION SECURITY

Submitted by:

21BCS8403--Kartik Kumar

21BCS8404--Pratik Mukherjee
21BCS8405--Krishma

Under the Supervision of:

Ms. Priyanka Jamawal

CHANDIGARH UNIVERSITY,GHARUAN,MOHALI-140413,PUNJAB

May2023

Abstract

Social engineering has emerged as a significant threat vector in the realm of cybersecurity, exploiting human psychology to manipulate individuals into divulging sensitive information or performing actions that compromise security. This research delves into one of the most prevalent and insidious forms of social engineering: phishing. The project abstract provides an overview of the study's objectives, methodologies, and anticipated contributions.

Phishing is a deceptive practice wherein cyber attackers masquerade as legitimate entities through emails, messages, or websites, aiming to trick recipients into revealing confidential information such as passwords, financial details, or personal data. The goal of this research is to comprehensively analyze the concept of phishing within the context of social engineering, shedding light on its underlying mechanisms, tactics, and potential countermeasures.

The study employs a multi-faceted approach, combining literature review, case studies, and empirical analysis. By examining real-world phishing attacks and their outcomes, the project seeks to uncover the psychological triggers that make individuals susceptible to such manipulation. Additionally, technical aspects of phishing campaigns, such as email spoofing and website replication, will be explored to understand the intricacies of attack execution.

Keywords: —Phishing; URL features; Social Engineering, Phishing Detection, Confidential Information.

Table of Contents

1. INTRODUCTION

Phishing is a type of social engineering attack that uses email or malicious websites to solicit personal information by posing as a trustworthy organization. The goal of a phishing attack is to trick the victim into clicking on a malicious link, opening an infected attachment, or providing sensitive information such as their username, password, or credit card number.

Social engineering is a broad term that refers to any type of attack that relies on human interaction to succeed. Phishing is a specific type of social engineering attack, but there are many other types, such as:

- Baiting: This involves sending the victim a gift or other incentive to get them to click on a malicious link.
- Quid pro quo: This involves offering the victim something in exchange for their personal information, such as a free trial or a discount.
- Tailgating: This involves following the victim into a secure area without authorization.
- Pretexting: This involves creating a false scenario in order to trick the victim into giving up their personal information.

1.1 Problem Definition

In today's digitally driven society, the convergence of technology and communication has brought about numerous benefits, but it has also given rise to increasingly sophisticated cyber threats. Among these threats, the practice of phishing within the domain of social engineering has emerged as a major

concern. Phishing attacks capitalize on human vulnerabilities, exploiting trust and deception to steal sensitive information, compromise accounts, and gain unauthorized access to systems.

The problem at hand is the alarming rise in successful phishing attacks and the resulting detrimental consequences for individuals, businesses, and institutions. Despite advancements in cybersecurity, attackers continue to exploit psychological manipulation to deceive victims into divulging confidential information or performing actions that jeopardize security.

The urgency of this issue lies in the financial, reputational, and emotional damage that phishing attacks inflict on victims. Organizations face significant financial losses, legal liabilities, and tarnished reputations due to data breaches and compromised customer information. Individuals experience personal data theft, identity fraud, and loss of privacy. These attacks undermine trust in digital interactions and hinder the potential of technology for positive change.

1.2 Problem Overview

Project phishing is a type of social engineering attack that is specifically designed to target organizations. In a project phishing attack, the attacker will typically impersonate a legitimate project manager or other employee of the organization in order to gain access to sensitive information or resources.

The overview of project phishing under social engineering can be summarized as follows:

- Phishing is a type of social engineering attack that uses email or malicious websites to solicit personal information by posing as a trustworthy organization.

- Project phishing is a type of phishing attack that is specifically designed to target organizations.
- The goal of project phishing is to gain access to sensitive information or resources.
- Project phishing attacks can be very successful, as they often exploit the trust that employees have in their organization.

1.3 Hardware Specification

- Laptop / PC.
- 1 GB RAM

1.4 Software Specification

- Visual Studio Code

2. LITERATURE SURVEY

A literature survey is an insightful article that presents the existing information including considerable discoveries just as theoretical and methodological commitments to a specific topic. A very effective detection of phishing website model which is focused on optimal feature selection technique and also based on neural network (OFS-NN) is proposed. In this proposed model, an index called feature validity value (FVV) has been generated to check the effects of all those features on the detection of such websites. Now, based on this newly generated index, an algorithm is developed to find from the phishing websites, the optimal features. This selected algorithm will be able to overcome the problem of over- fitting of the neural network to a great extent. These optimal features are then used to build an optimal classifier that detects phishing URLs by training the neural network.

Feature engineering plays a vital role in finding solutions for detection of phishing websites, although the accuracy of the model greatly will be based on knowledge of the features. though the features taken from all these various dimensions are understandable, the limitation lies in the time taken to collect these features. To fix this drawback, the authors have proposed a multidimensional phishing detection feature approach that concentrates on a rapid detection technique by making use of deep learning (MFPD) To detect phishing occurrence accurately, a three-phase detection called Web Crawler based Phishing Attack Detector (WC- PAD) has been proposed. This takes the web's content, traffic and URL as input features. Now considering these features, classification is done.

A method called parse tree validation has been proposed to find if a webpage is phishing or legitimate. This is an innovative approach to find such web sites by

intercepting every hyperlink of a present page through API of Google, and developing a parse tree from all those hyperlinks that were intercepted. In this, parsing begins from the root node. It goes by the Depth-FirstSearch (DFS) algorithm to determine if any child node has the same value as the root node.

1.5 Existing System

- Malicious Web sites are the basis of most of the criminal activities over theinternet.
- The dangers that arise due to the malicious sites are enormous and the end-users must be prohibited from visiting such sites.
- The users should prohibit themselves from clicking on such Uniform ResourceLocator (URL).
- In order to prevent such attacks, the paper proposes the use of machine learningalgorithms to detect
- Phishing Websites. The Existing PWD (Phishing Website Detection) model is trained using an existing dataset which contains URLs, each with unique features, and is applied to three different
- machine learning classifiers—support vector machine, logistic regression and Naïve Bayes. After training and testing the algorithms, it is observed that NaïveBayes classifier recorded the highest accuracy.

Disadvantages

- Low Accuracy Due to Training Loss.
- Many Website features not included for the consideration.

2.2 Proposed System

- Collect dataset containing phishing and legitimate websites from the open-source platforms.
- Write a code to extract the required features from the URL database.
- Analyse and pre-process the dataset by using EDA techniques.
- Divide the dataset into training and testing sets.
- Run selected machine learning and deep neural network algorithm (DNN) onthe dataset.
- Write a code for displaying the evaluation result considering accuracy metrics.
- Compare the obtained results for trained models and specify which is better.
- **DNN-** This is also one of the classification algorithms which is supervised and is easy to use. It can used for both classification and regression applications, but it is more famous to be used in classification applications.

Advantages

- Provide clear idea about the effective level of each classifier on phishing emaildetection.
- High level of accuracy by take the advantages of classifiers many aspect.
- High level of accuracy.
- Fast in classification process fast, less consuming memory, high accuracy, Evolving with time, online working.

2.3 Literature Review Summary

YEAR	TITLE	TECHNIQUE	LIMITATION
2019	OFS-NN: ”An Effective Phishing Websites Detection Model Based on Optimal Feature Selection and Neural Network”	Proposedmethodhas 3stages:1.Definesan ewindex -FVV. 2. Designs an optimal feature selection algorithm.3. Produce theOFSNN model	The continuous growing of features that are sensitive of phishing attacks need collection of more features for the OFS
2019	“Phishing Website Detection based on Multidimensional Features driven by Deep Learning”	1.character succession features of the URL are extricated as well as utilized for fast characterization. 2.the LSTM (longshort-term memory)network is utilizedto catch settingsemantic and dependency features of URL character groupings.	It requires more computation and therefore an expensive method.
2019	WC-PAD: Web Crawling based Phishing Attack Detection.	It is a 3-phase detection of phishing attack approach. The 3 phases of WC-PAD	Time consuming as it involves three phases and each website has

		are1) black list of DNS2) Approach based on Heuristics and3) Approach based on WebCrawler. Feature extraction as well as phishing attack Detection both makes use of WebCrawler.	to go through the three phases.
2018	Phishing Detection in Websites using Parse Tree Validation	If the number of recurrences of root node is: 1. More than half the number of nodes, then probability of authenticity is more. 2. Quarter the number of nodes, the probability of authenticity is moderate. 3.less than the quarter number of nodes, then probability of authenticity is low which means phishing probability is high.	The false Negative and false positive rates are high.

3. PROBLEM FORMULATION

Problem: Phishing is a type of social engineering attack that uses email or malicious websites to solicit personal information by posing as a trustworthy organization. The goal of a phishing attack is to trick the victim into clicking on a malicious link, opening an infected attachment, or providing sensitive information such as their username, password, or credit card number.

- Formulation: How can we formulate a project to improve the ability of organizations to identify and avoid phishing attacks?
- Specific objectives:
- Identify the different types of phishing attacks and the techniques that attackers use.
- Develop a training program to educate employees on how to identify and avoid phishing attacks.
- Implement a security policy that prohibits employees from clicking on links or opening attachments from unknown senders.
- Evaluate the effectiveness of the training program and security policy.

This is just a sample problem formulation, and the specific objectives of the project will vary depending on the organization's needs. However, this formulation provides a good starting point for developing a project to improve the ability of organizations to identify and avoid phishing attacks.

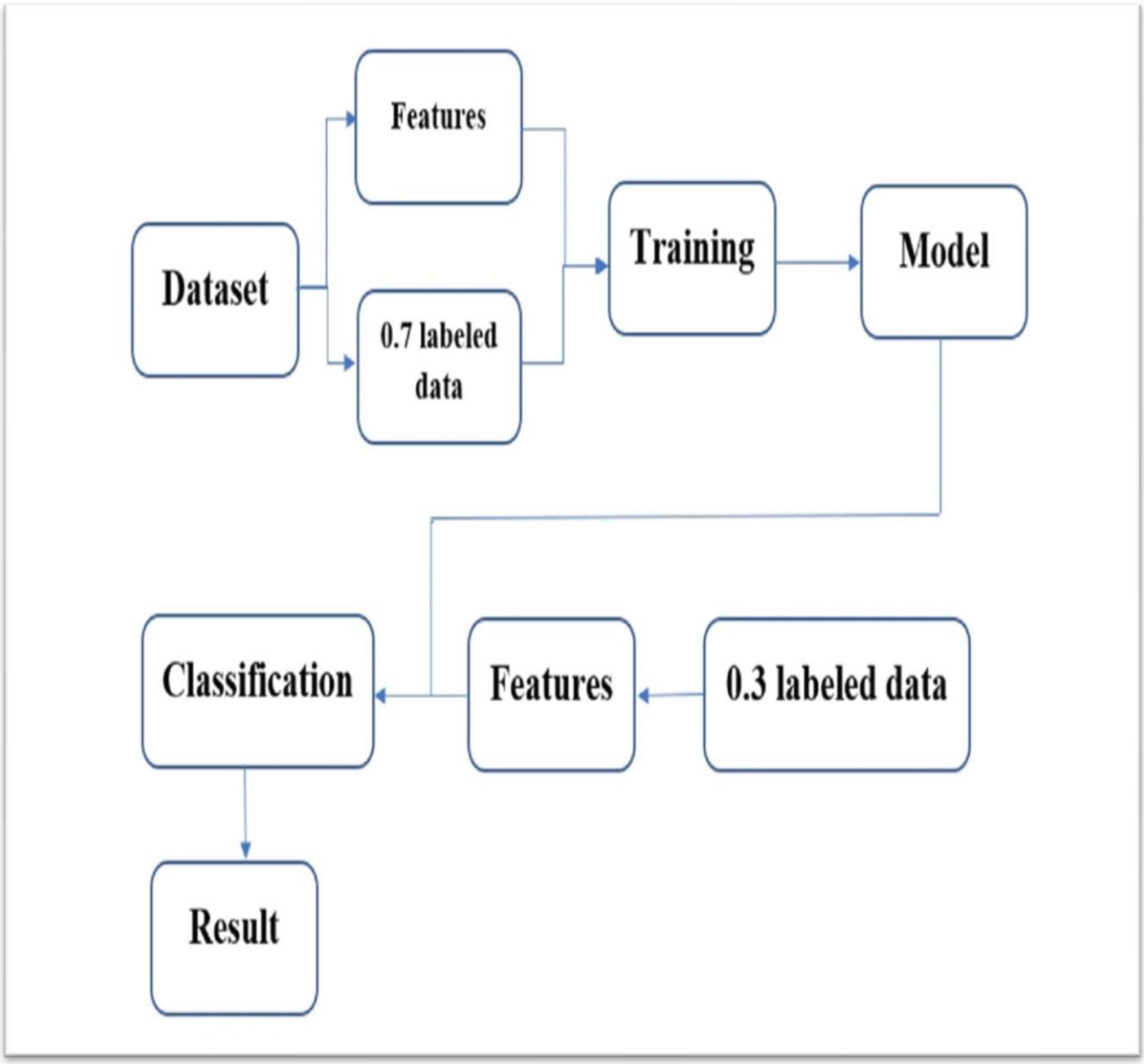
4. OBJECTIVES

The objectives of phishing under social engineering project can vary depending on the specific goals of the project. However, some common objectives include:

- To identify the different types of phishing attacks and the techniques that attackers use.
- To develop a training program to educate employees on how to identify and avoid phishing attacks.
- To implement a security policy that prohibits employees from clicking on links or opening attachments from unknown senders.
- To evaluate the effectiveness of the training program and security policy.
- To raise awareness of phishing attacks and their risks.
- To improve the security posture of the organization by reducing the number of phishing attacks that are successful.

The specific objectives of the project will depend on the organization's needs and the resources that are available. However, by achieving these objectives, organizations can help to protect themselves from phishing attacks and the risks that they pose.

5. METHODOLOGY



6. CONCLUSION

In conclusion, this project has illuminated the intricate landscape of phishing attacks within the context of social engineering, shedding light on the multifaceted tactics that exploit human psychology and trust. Through comprehensive analysis and investigation, we have gained insights into the methodologies employed by malicious actors, ranging from deceptive emails to convincing phone calls and fraudulent websites. Our exploration of the psychological triggers and cognitive biases that render individuals susceptible to these attacks has underscored the importance of awareness and vigilance in the digital realm.

The classification of different phishing attack types has provided a nuanced understanding of their unique characteristics, enabling us to tailor effective prevention strategies. The evaluation of existing defenses has revealed both strengths and limitations, reinforcing the need for a holistic approach that combines technological solutions with informed user behavior. By developing proactive mitigation strategies and empowering individuals with knowledge, we aim to bolster cybersecurity measures and foster a culture of skepticism and cautious engagement.

As we navigate the evolving threat landscape, ethical considerations have remained paramount, ensuring that our research respects user privacy, consent, and responsible conduct. By disseminating our findings, recommendations, and insights, we aspire to contribute to a more resilient digital ecosystem. Ultimately, the endeavor to combat phishing attacks through social engineering hinges on collective efforts—uniting researchers, organizations, and users—in an ongoing quest for a safer and more secure digital world.

7. REFERENCES

1. What is phishing ? : **https://www.phishing.org/what-is-phishing**.

2. Sarker, I.H. Deep Cybersecurity: A Comprehensive Overview from Neural Network and Deep Learning Perspective. *SN Comput. Sci.* **2021**, *2*, 154.

3. Kim, D., Kim, Y.-H., Shin, D., Shin, D.: Fast attack detection system using log analysis and attack tree generation. Clust. Comput. **22**(1), 1827–1835 (2019).

4. Phishing Detection using Machine Learning based URL Analysis: A Survey- https://www.ijert.org/research/phishing-detection-using-machine-learning- based-url-analysis-a-survey-IJERTCONV9IS13033.pdf.

5. An intelligent cyber security phishing detection system using deep learning techniques https://link.springer.com/article/10.1007/s10586-022-03604-4.

6. https://github.com/chamanthmvs/Phishing-Website-Detection.

7. https://www.researchgate.net/publication/328541785_Phishing_Website_Detection_using_Machine_Learning_Algorithms.

8. https://bmsit.ac.in/public/assets/PG_NBA_Data/Program_Level_Documents/5%20PROJECT%20REPORTS/2018-20/Bhagya_report_final.pdf.

9. https://www.academia.edu/43230944/A_REPORT_on_DETECTION_OF_PHISHING_WEBSITE_USING_MACHINE_LEARNING.

10. Phishing detection system using machine learning https://youtu.be/pjewBSmWTlU.