CMR COLLEGE OF ENGINEERING & TECHNOLOGY

Kandlakoya, Medchal, Hyderabad – 501401

Department of CSE - Cyber Security

(BATCH – 18)

PHISHING DETECTION

Under esteemed guidance of

Dr.Punyaban Patel

(Professor, CSC)

Name of the student	Roll No.
1. G.Charan Sai	20H51A6279
2. A.Rishika	20H51A6297
3. G.Dheeraj Reddy	20H51A62A0

Outline

1.Abstract

2.Introduction

3.Literature Survey

4.Existing Systems

   a. Problems in Existing System

5.Problem Definition

6.Proposed System

   a. Architecture

   b. Methodology

   c. System Requirements

   d. Proposed system vs Existing System.

7.Conclusion

8.References

ABSTRACT

Abstract

- Phishing is a form of **Social Engineering** and **Scam**.
- Purpose : **Login credentials , Financial details**
- Phishing detection is the process of **identifying** ,**alerting** and **mitigating** the phishing attacks.
- It is a **semantics-based attack**, that mainly focuses on human vulnerabilities.
- Our project is designed to **enhance online security .**
- This comprehensive approach involves :

a) URL analysis

b) Database comparison

c) Greek Alphabet Analysis

d) Homographic Attack URLs

e) Port forwarding detection

INTRODUCTION

Introduction

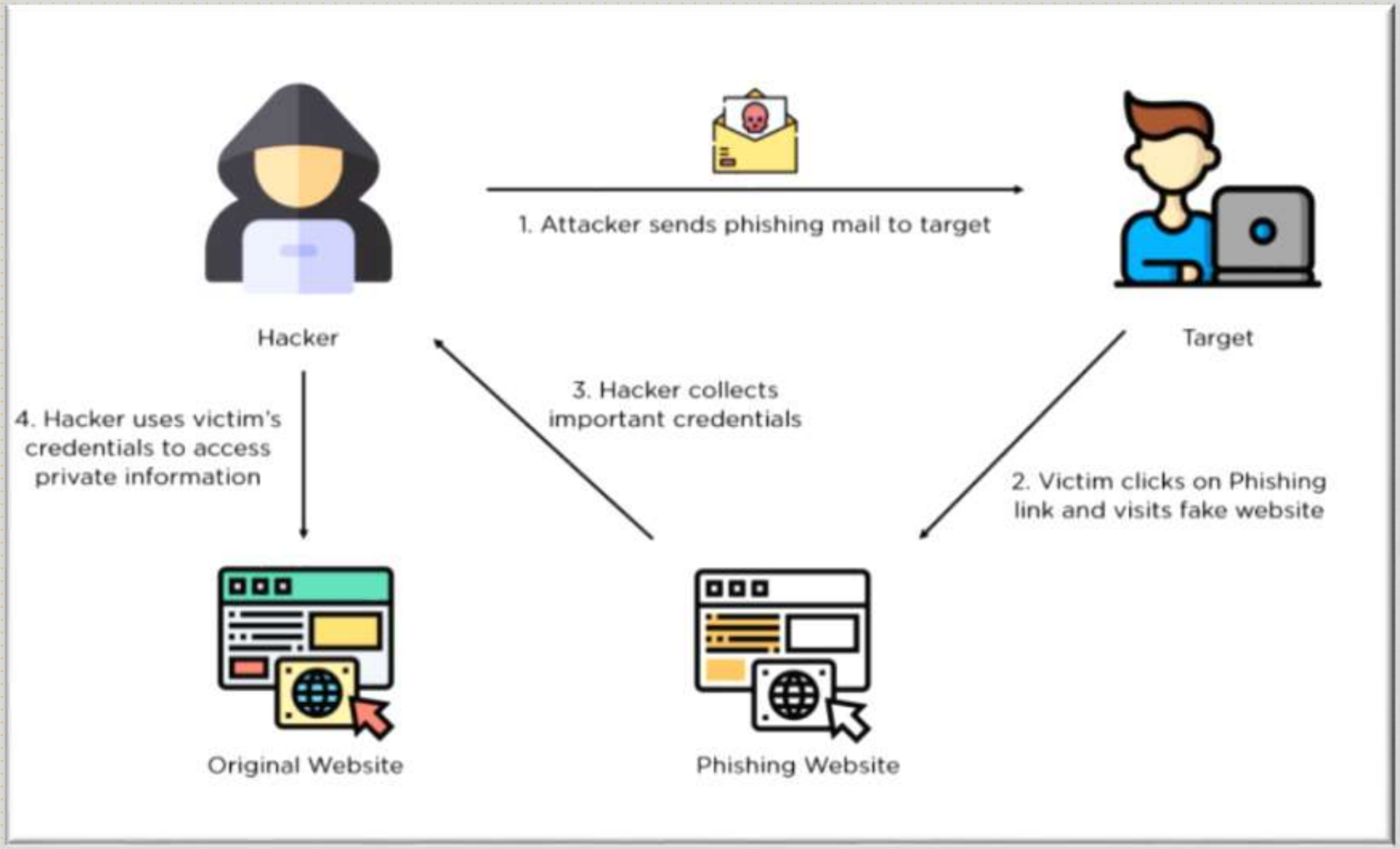
- Phishing is an attack in which the **threat actor** poses as a **trusted person** or organization .

- These attacks consists of three main entities.They are:

<u>THE SENDER</u>: The sender imitates someone trustworthy and the phishing messages mimic emails from large companies like PayPal, Amazon, or Microsoft, and also banks or government offices.

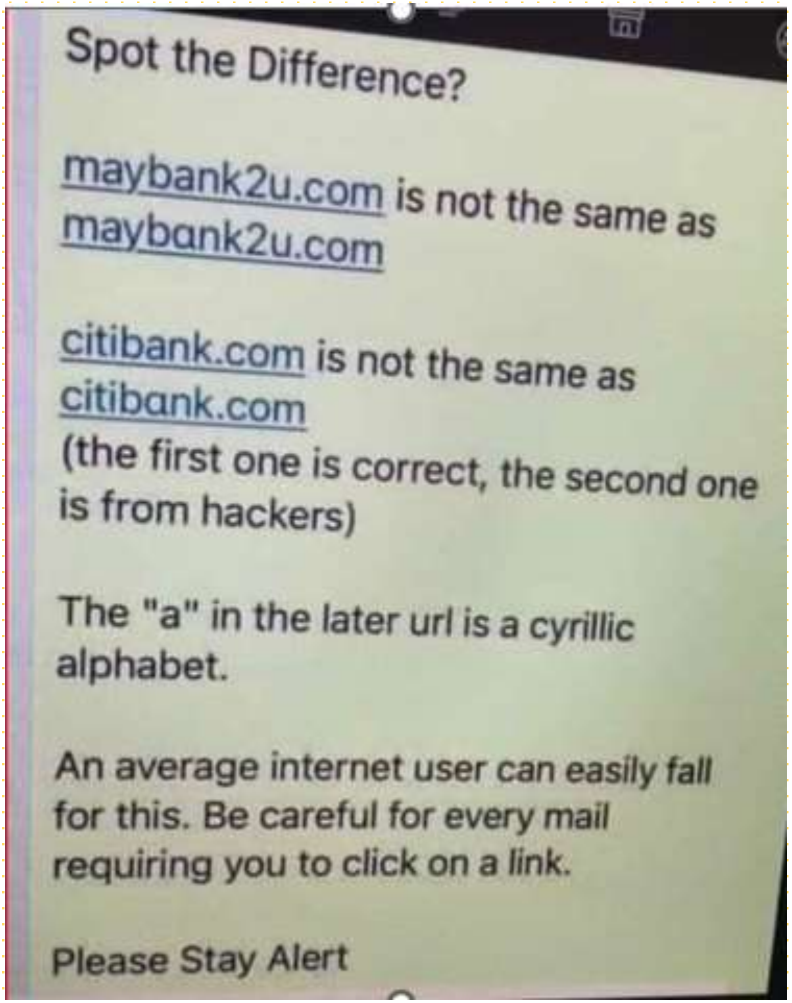
<u>THE MESSAGE</u>: The attacker will ask the recipient to **click a link**, **download an attachment**, or to **send money** or they create an emergency situation.

<u>THE DESTINATION</u>: If users click the link, they're sent to an imitation of a legitimate website. From here, they're asked to log in with their username and **password credentials**.


1. Attacker sends phishing mail to target
Hacker
Target
4. Hacker uses victim's credentials to access private information
3. Hacker collects important credentials
2. Victim clicks on Phishing link and visits fake website
Original Website
Phishing Website

OVERVIEW OF A PHISHING ATTACK

How does a Phishing Attack look like?



Spot the Difference?

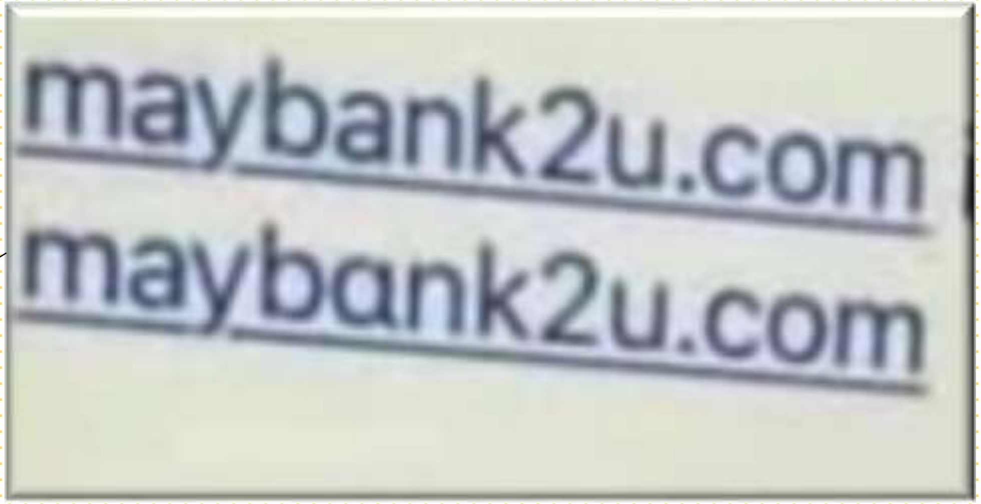
maybank2u.com is not the same as maybɑnk2u.com

citibank.com is not the same as citibɑnk.com
(the first one is correct, the second one is from hackers)

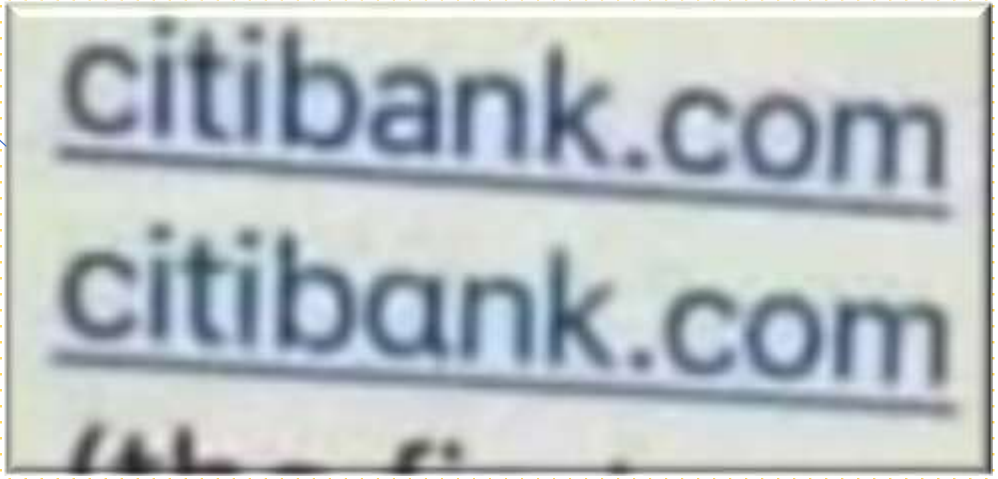
The "a" in the later url is a cyrillic alphabet.

An average internet user can easily fall for this. Be careful for every mail requiring you to click on a link.

Please Stay Alert

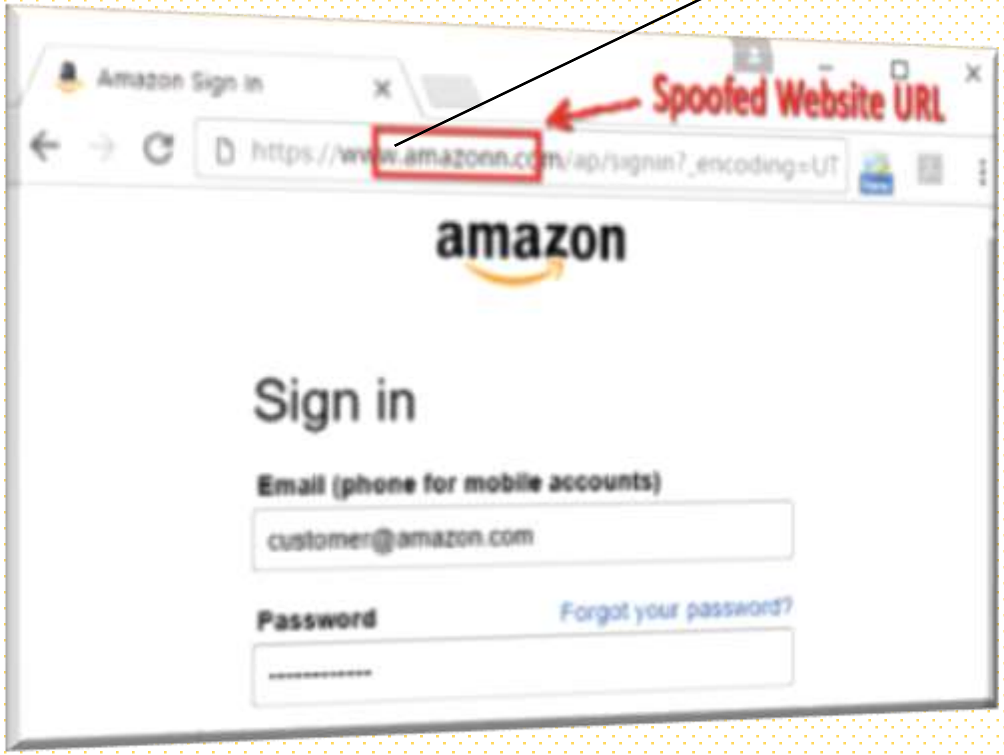


maybank2u.com
maybɑnk2u.com



citibank.com
citibɑnk.com

www.amazonn.com



Amazon Sign in

https://www.amazonn.com/ap/signin?_encoding=UT

Spoofed Website URL

amazon

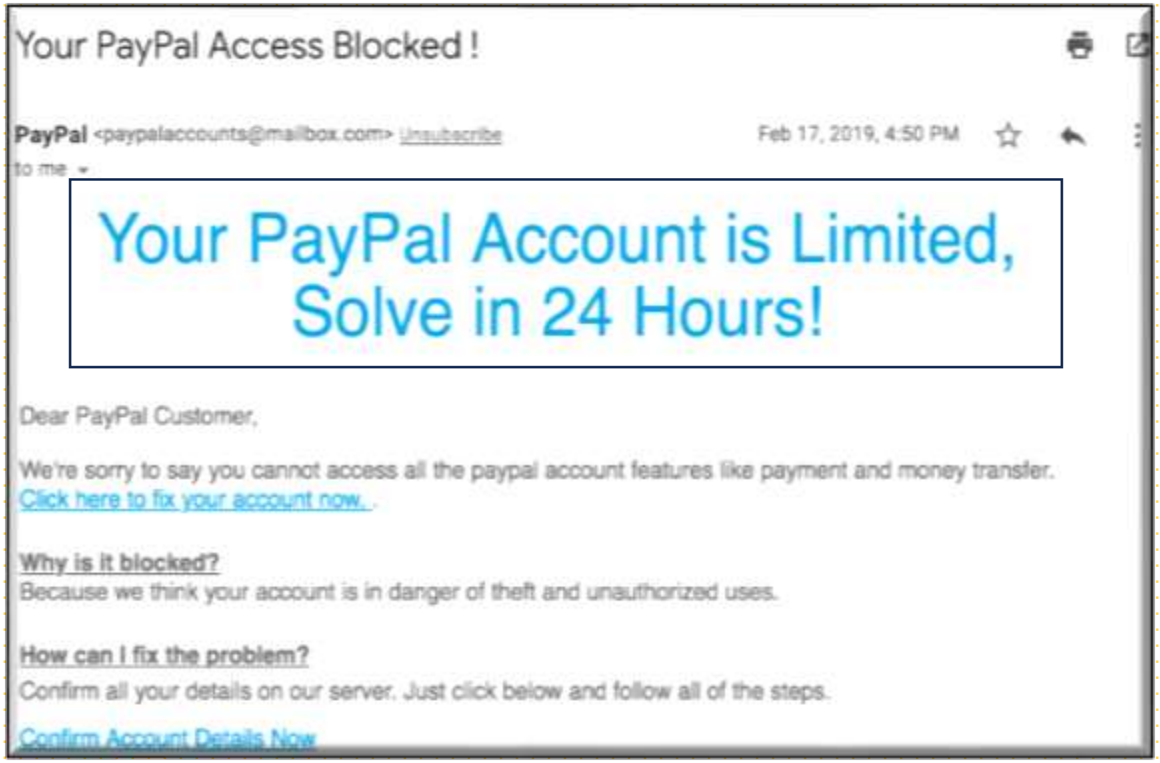
Sign in

Email (phone for mobile accounts)

customer@amazon.com

Password

Forgot your password?



Your PayPal Access Blocked !

PayPal <paypalaccounts@mailbox.com> Unsubscribe

Feb 17, 2019, 4:50 PM

to me

Your PayPal Account is Limited, Solve in 24 Hours!

Dear PayPal Customer,

We're sorry to say you cannot access all the paypal account features like payment and money transfer.
Click here to fix your account now. .

Why is it blocked?
Because we think your account is in danger of theft and unauthorized uses.

How can I fix the problem?
Confirm all your details on our server. Just click below and follow all of the steps.

Confirm Account Details Now

LITERATURE SURVEY

Literature Survey

1) Email based Spam Detection: June-2020

- **AUTHOR :**Thashina Sultana, K A Sapnaz, Fathima Sana, Mrs. Jamedar Najath
- Published by International Journal of Engineering Research & Technology (IJERT)

https://www.researchgate.net/publication/342113653_Email_based_Spam_Detection

2) SemanticPhish : Nov- 2020

- **AUTHOR :**Guy-Vincent Jourdan, Gregor v. Bochmann
- Published by Institute of Electrical and Electronics Engineers (IEEE)
- https://ieeexplore.ieee.org/document/9493252/authors#authors

EXISTING SYSTEMS

Existing Systems

1.PhishTank: PhishTank is a community-driven website that tracks known phishing websites. You can search their database or report suspicious URLs.

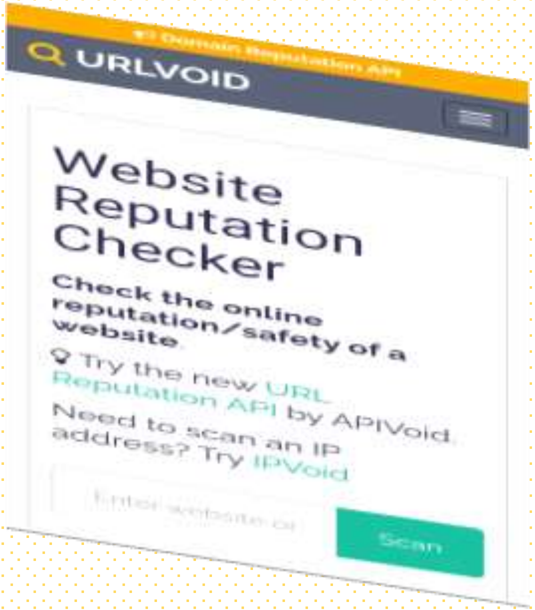
PhishTank

Join the fight against phishing

What is phishing?

What is PhishTank?

2.URLVoid: URLVoid provides a website reputation checker that scans websites against multiple blacklists and security services. It can help you determine the trustworthiness of a website.



Domain Reputation API

URLVOID

Website Reputation Checker

Check the online reputation/safety of a website

Try the new URL Reputation API by APIVoid

Need to scan an IP address? Try IPVoid

Enter website or

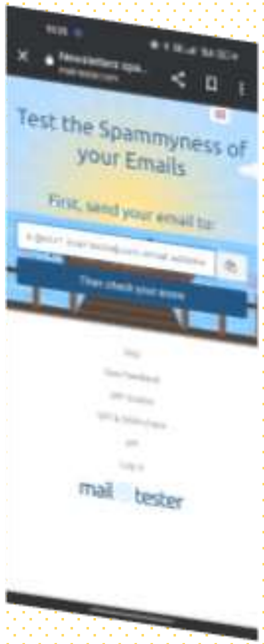
Scan

3.CheckPhish: CheckPhish is an online tool that allows you to analyze and check the legitimacy of a website.

Are you keeping tabs on domain typosquats?

Monitor one domain for free to see all typosquat and lookalike variants

4.Mail Tester: Mail Tester is an email validation service that can help you verify if an email address is legitimate. While it doesn't specifically detect phishing, it can be useful in verifying the authenticity of email senders.



Test the Spammyness of your Emails

First, send your email to:

- ## Problems in Existing Solutions:

1) Less Accuracy
2) Inability to Detect unknown attacks
3) High False Acceptance Rate
4) Not using multi layered approach

PROBLEM DEFINITION

Problem Definition

Phishing websites are duplicate web pages created to mimic real websites in-order to deceive people to get their personal information. Identifying these phishing websites is typically a challenging task because phishing is mainly a semantics-based attack, that mainly focuses on human vulnerabilities, not the network or software vulnerabilities. Most of the phishing attacks are sent via email.

PROPOSED SYSTEM

Proposed Solution

- Our proposed system is a website with comprehensive and **multi-layered solution** .
- The system begins by

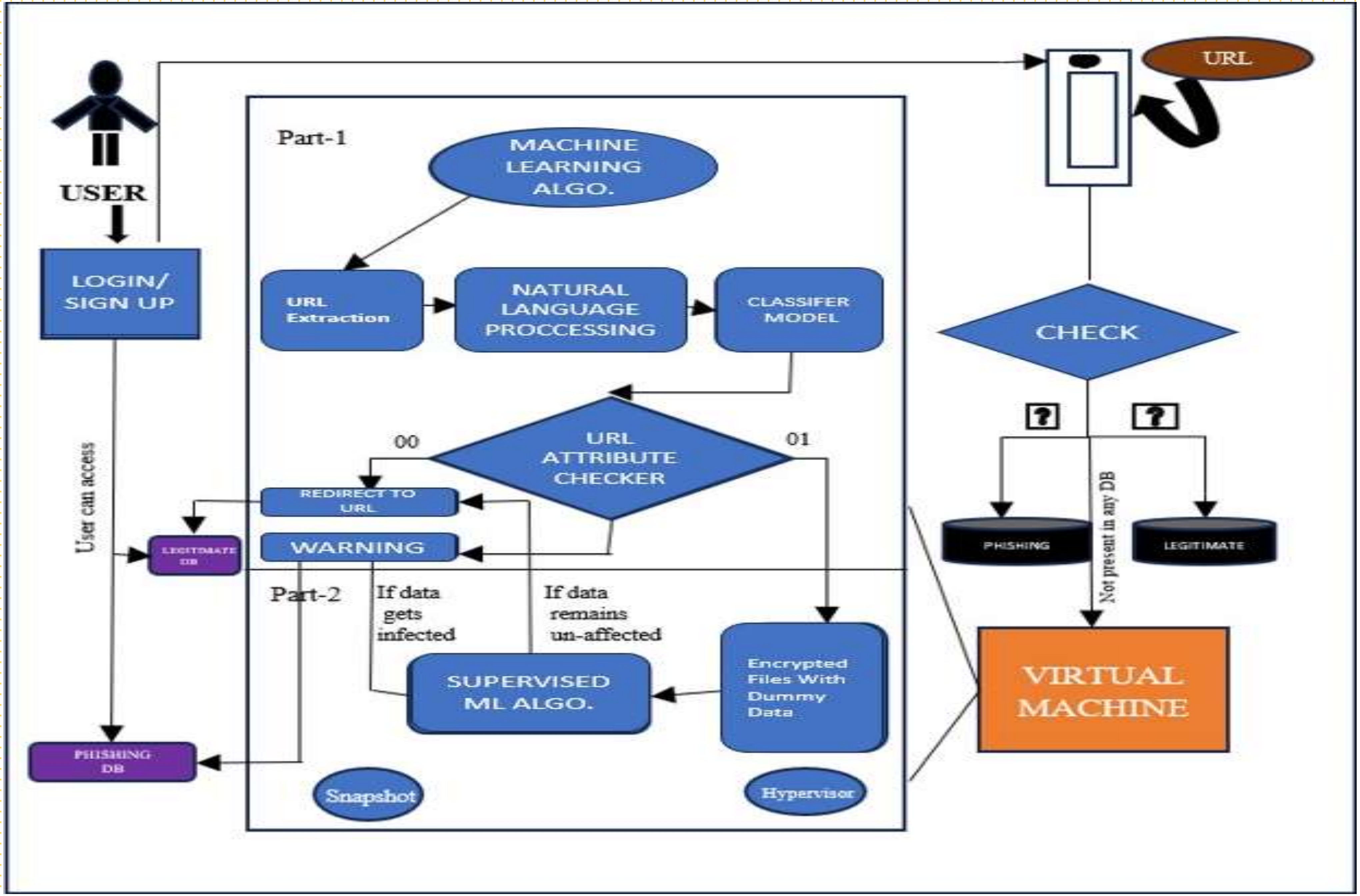
➢ **Meticulously scrutinizing** the **structure** of the URL , alerting users to potential threats based on linguistic anomalies.

➢ It then compares provided **URLs** to an extensive **database** of known phishing and malware-infested websites, offering instant alerts in case of a match.

➢ To further fortify security, we investigate the URLs for any **port forwarding** techniques, a common method used by cybercriminals.

➢ Finally, our system conducts **Greek alphabet analysis**, **Homographic attack URLs** specifically targeting the presence of Greek characters within URLs and alerting the users or administrators.

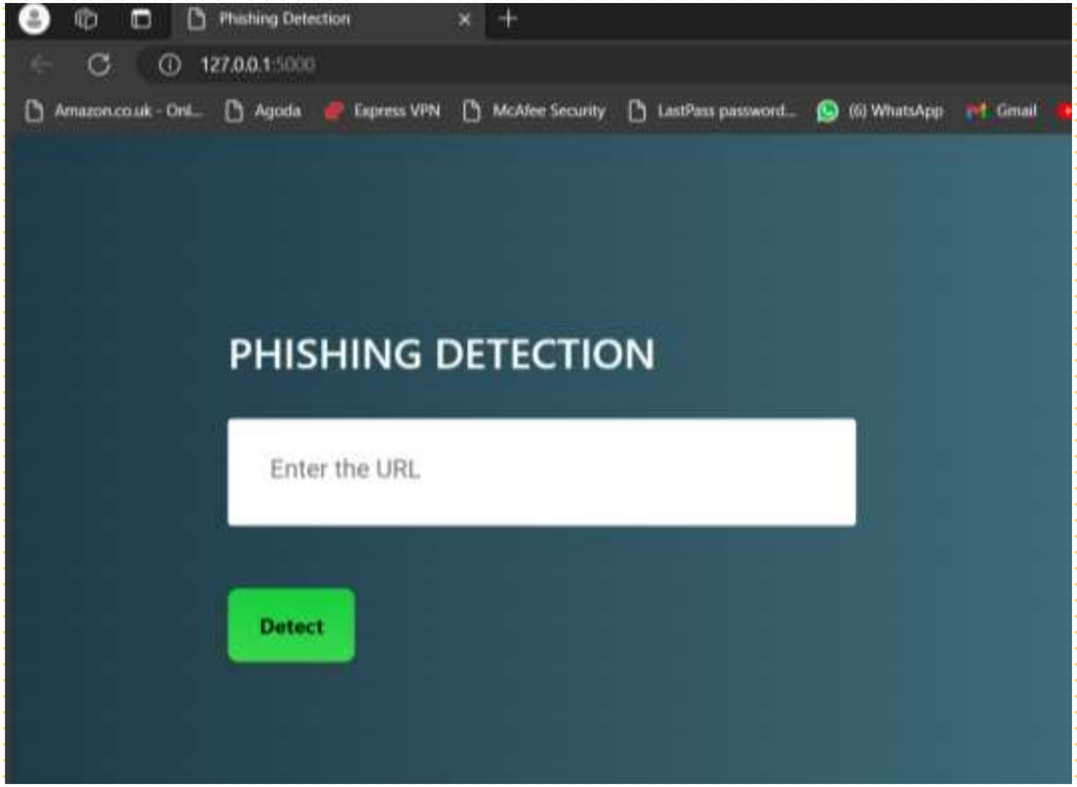
ARCHITECTURE / BLOCK DIAGRAM


USER
LOGIN/ SIGN UP
User can access
LEGITMATE DB
PHISHING DB
Part-1
MACHINE LEARNING ALGO.
URL Extraction
NATURAL LANGUAGE PROCCESSING
CLASSIFER MODEL
URL ATTRIBUTE CHECKER
00
01
REDIRECT TO URL
WARNING
Part-2
If data gets infected
If data remains un-affected
SUPERVISED ML ALGO.
Encrypted Files With Dummy Data
Snapshot
Hypervisor
URL
CHECK
?
?
PHISHING
LEGITIMATE
Not present in any DB
VIRTUAL MACHINE

METHODOLOGY

Methodology

- This comprehensive approach involves several layers of analysis to assess the legitimacy of URLs. The URL which we want to check has to be entered in the text box and click on the submit button for the result.



PHISHING DETECTION

Enter the URL

Detect

a) URL analysis
b) Database comparison
c) Greek Alphabet Analysis
d) Homographic Attack URLs
e) Port forwarding detection

System Requirements

Hardware Requirements:

• RAM	: 16 GB or above
• Processor	: intel i5 or above

Software Requirements:

• Operating System (used)	: Windows
• Programming Language	: Python
• Technologies used	: Python Libraries , Flask framework
• Editor	: Visual Studio code

PROPOSED vs EXISTING SYSTEM

PROPOSED SYSTEM	EXISTING SYSTEM
More Accurate than the existing systems	Less Accurate
Can detect unknown attacks and alert the users	Inability to Detect unknown attacks
Multi- Layered Approach	Single – Layered Approach

Conclusion

Our project is a **major advancement** in online security, countering phishing and malware threats with a multi-layered approach.

Our system detects **semantic anomalies** and **evolving cybercriminal techniques**, promptly alerting users.

We aim to significantly **reduce** phishing and malware success rates, creating a safer digital environment.

➢ Future Scope :

- **Extension for e-mail and message application:**

Sending the phishing mails and messages directly to the spam folder of email and messaging application respectively.

- **User Education**:

Develop features or modules that provide educational content to users, helping them recognize and avoid phishing attempts. This could include interactive tutorials, quizzes or simulated phishing exercises.

REFERENCES

1)Author: Emalderson ,Published in the year 2022

https://github.com/emalderson/ThePhish

2) Author : Ashit kumar Datta, Published on 2021 Oct

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8504731/

3) Author : Fatima Salahdine, Zakaria El Mrabet, Naima Kaabouch, Published on 2022 January.

https://www.researchgate.net/publication/358142755_Phishing_Attacks_Detection_--

4) https://www.sciencedirect.com/science/article/pii/S1319157823000034

THANKYOU