YILDIZ TECHNICAL UNIVERSITY

FACULTY OF ELECTRICAL AND ELECTRONICS

SECURITY OF COMPUTER SYSTEMS (BLM4011)

TERM PROJECT REPORT

Umutcan Sevdi - 19011091

İsmet Güngör - 19011100

Semih Yazıcı - 19011087

Oğuzhan Ercan - 18011054

umutcan.sevdi@std.yildiz.edu.tr

ismet.gungor@std.yildiz.edu.tr

semih.yazici@std.yildiz.edu.tr

oguzhan.ercan@std.yildiz.edu.tr

DEPARTMENT OF COMPUTER ENGINEERING

Data Security: Machine Learning and Regex Matching Based Phishing Detection System Development for a System That Performs SMTP Phishing as an Attack

1. INTRODUCTION

Phishing is a type of cyber attack involving carefully crafted emails or websites to trick individuals into revealing sensitive information such as login credentials or financial information. These attacks often take the form of fake login pages or emails purporting to be from legitimate organizations, and they can have serious consequences for both individuals and organizations.

Phishing attacks have become increasingly common as attackers have become more sophisticated in their techniques. These attacks can be highly effective, as they often prey on individuals' trust and lack of awareness of the risks involved. In addition to stealing sensitive information, phishing attacks can install malware on individuals' computers, giving attackers access to even more sensitive information.
There are several common ways to handle the problem of phishing attacks. One of the most effective approaches is to implement strong password policies and use two-factor authentication whenever possible. This helps to prevent attackers from being able to gain access to sensitive information even if they can obtain an individual's login credentials.

In our project, we developed a phishing scenario and a program to protect from it. For this scenario, we hosted an SMTP server and a phishing server. Phishing server tricks users into thinking that the website is legit.
When the victim clicks on the link, a login page imitates edevlet.gov.tr is returned. However, when the user logs in, all credentials are sent to the attacker. Phishing site responds with a fake dashboard to be unnoticed.
To be protected from attacks like those, organizations can also take steps to protect themselves from phishing attacks. This can include implementing email filtering systems to block suspicious emails and implementing security measures such as firewalls and intrusion detection systems to prevent attackers from gaining access to sensitive data.
In this project, we aimed to develop a machine learning and regex matching-based phishing detection system to identify and prevent phishing attacks on a system that performs SMTP phishing as an attack. The use of machine learning algorithms and regex matching allows the system to analyze and classify email content and identify patterns and keywords commonly used in phishing attacks. This approach has the potential to be highly effective in detecting and preventing phishing attacks, as it can quickly and accurately identify suspicious emails and take action to block them.

2. METHOD

We have designed a system including an attacker and a victim.

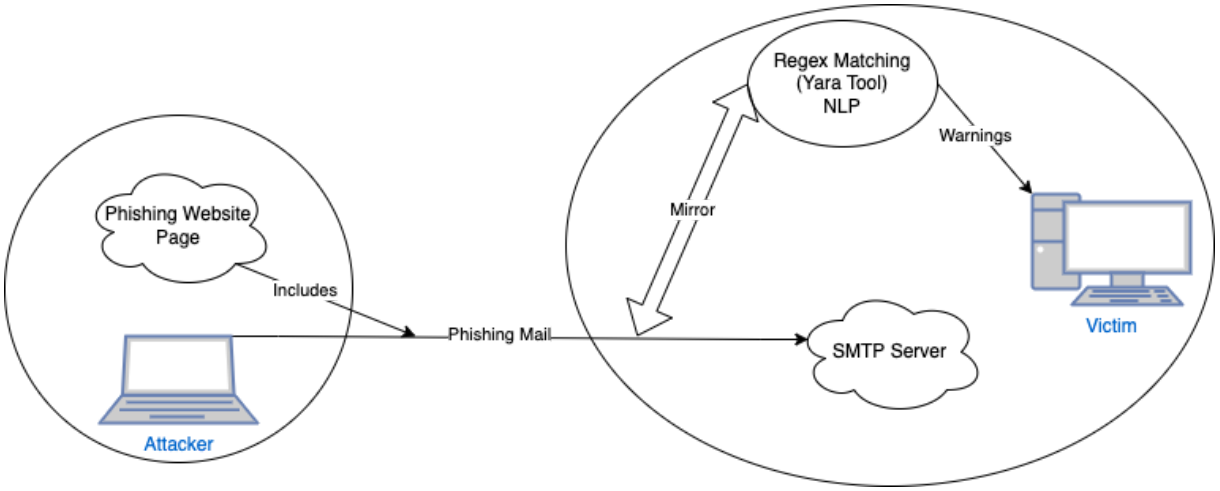

Figure 2.1 Diagram of the System

In the system, the attacker's purpose is to attack the victim via SMTP. For that, the attacker hosts a web server that imitates edevlet.gov.tr for a phishing attack. This server sends an HTML page that looks exactly like the original site. However, unlike the original, this version does not encrypt the username and password. It sends the form data to the attacker's server.
In this mail, the attacker sends a mail containing the link demanding a login to gain the victim's TCKN and password for logging into e-devlet. When the user logs in, it redirects to a website that looks exactly like the homepage of edevlet.gov.tr.
We developed the phishing web server in Go. This server hosts two websites depending on the incoming request. If a GET request is received, it returns the login page of Edevlet. If a POST request is received, it scans the form data to find the username and password and produces a fake dashboard that contains the TCKN of the user. So the victim doesn't realize whether the website is compromised or not.
We used a MailHog server running from a docker-compose file as a container for testing purposes. We assumed a scenario where the victim hosts a web server and connects to it. The attacker hosts a phishing website and sends an email to the victim.
To protect the victim against phishing attacks, we have implemented a system that listens to the ongoing traffic and parses SMTP to examine the mail body. After obtaining the mail body, firstly process with Yara using rules specifically generated for detecting phishing mail attacks. After checking possible malicious keywords with the Yara tool, transferring the plaintext body to a Python program, a machine learning method that determines whether the incoming mail is a phishing attack or innocent.
We have called Long Short Term memory, a type of recurrent neural network (RNN) well-suited for modeling long-term dependencies in time series or sequential data. It can effectively retain information over long periods and handle variable-length input sequences. The attention layer weighs the input sequences, and the classifier predicts based on the weighted input. The model also has methods for generating initial hidden states for the LSTM layer, encoding input text using the embedding layer and LSTM layer, and applying attention

to the output of the LSTM layer. In addition, we detect which words cause phishing thanks to the attention layer placed between LSTM and linear classifiers in the model.

The text that came over TCP and converted to the string was not in a format that could be fed into our LSTM model. For this reason, we performed the text preprocessing steps frequently used in natural language processing tasks. The utils_preprocess_text function is used for cleaning and preprocessing text by removing punctuation and lowercasing, removing stop words, and optionally applying stemming or lemmatization. The textCleaner function applies the utils_preprocess_text function to a column of a pandas DataFrame and stores the processed text in a new column.

3. RESULTS

In our project, we successfully implemented four different programs while hosting another program from a docker container. We developed a phishing server that imitates a legit website. We developed an SMTP traffic analyzer and regex-based detection program based on YARA. We developed a machine learning-based detection program to support the previous program. Finally, we developed a program that sends phishing emails.
All developed programs run successfully and can protect the victim.

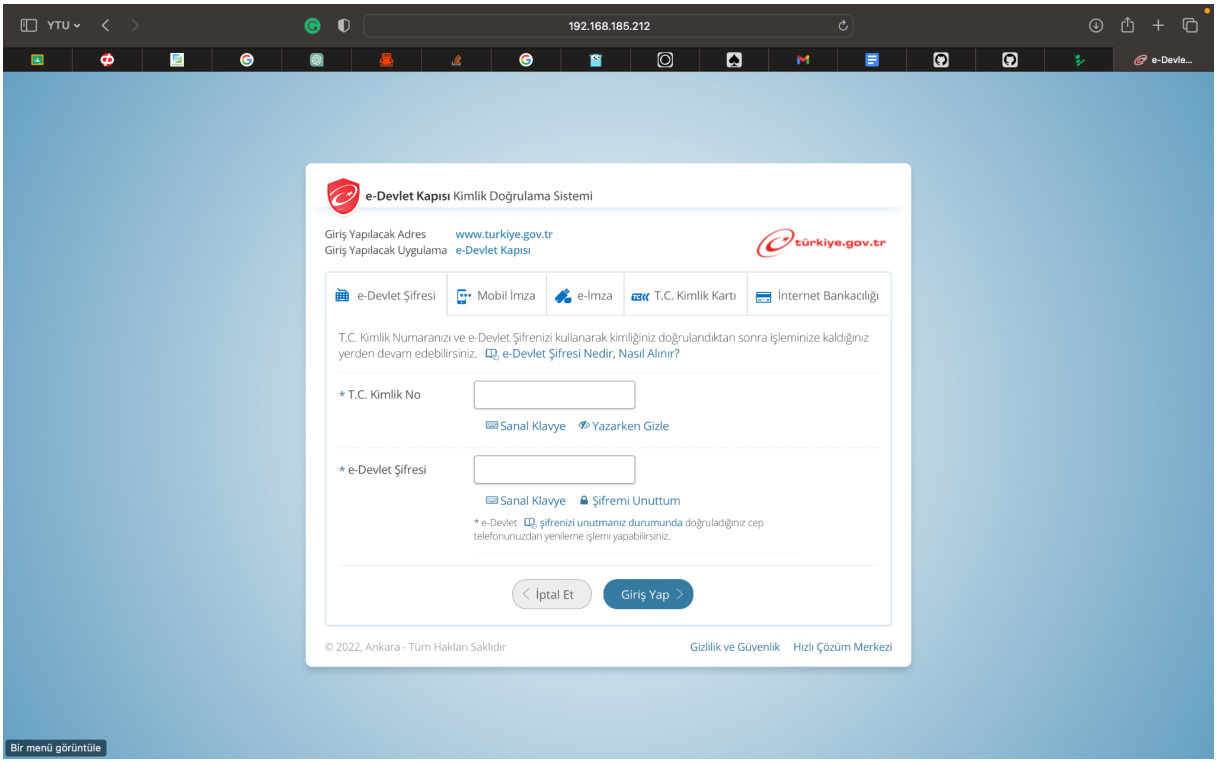

Figure 3.1 Phishing Website Login Page

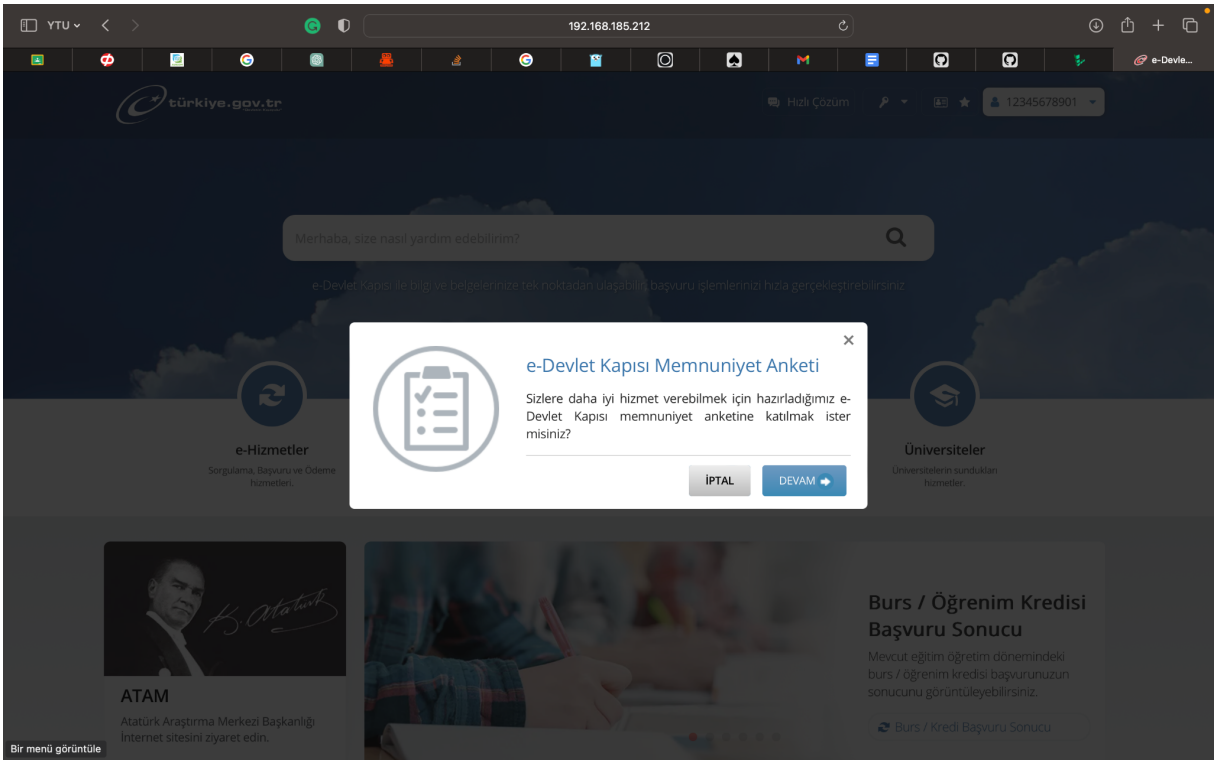

Figure 3.2 Phishing Website After Login

```
Collected username and password
username:  12312312312
password: 1231231
2022/12/25 22:29:23 request index: /
2022/12/25 22:29:23 Phishing started
Form data:
jsonData : []
submitButton : [Giriş Yap]
encTridField : [12345678901]
tridField : []
egpField : []
encEgpField : [testtesttest]
currentPageToken : [2396833e-ddb5-4f59-ba9e-825f7f190acc]
actionName : [giris]
Collected username and password
username:  12345678901
password: testtesttest
```

Figure 3.3 Obtained TCKN and Password via Phishing Website Login

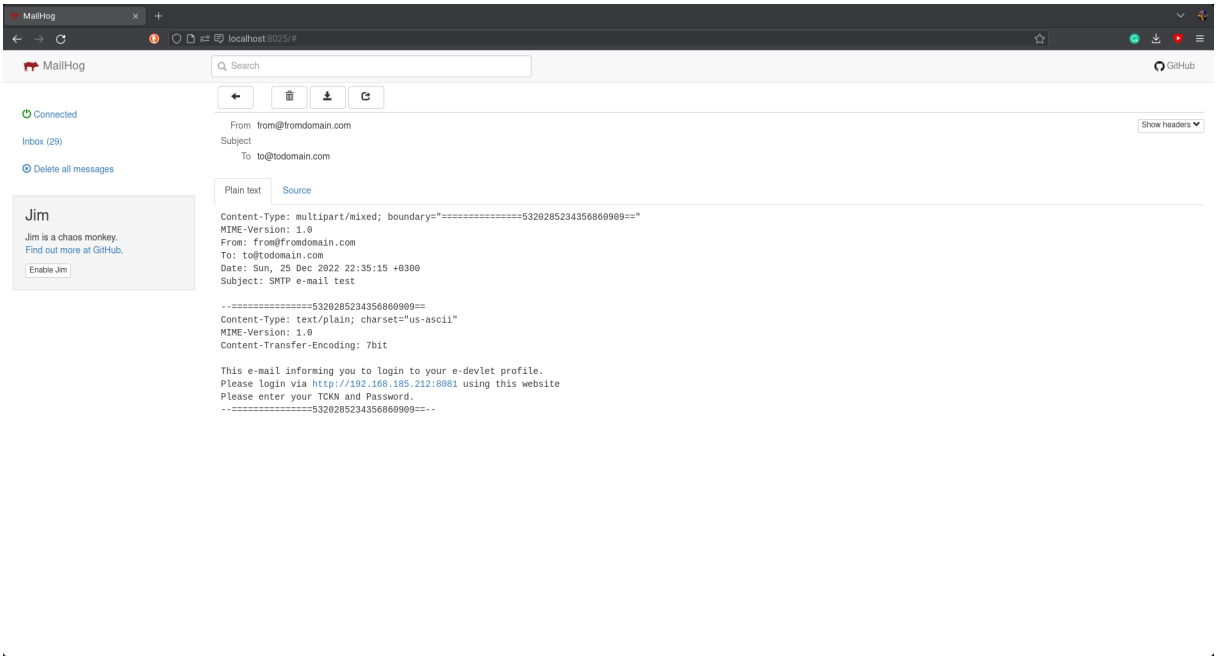

Figure 3.4 Received Mail From SMTP Server

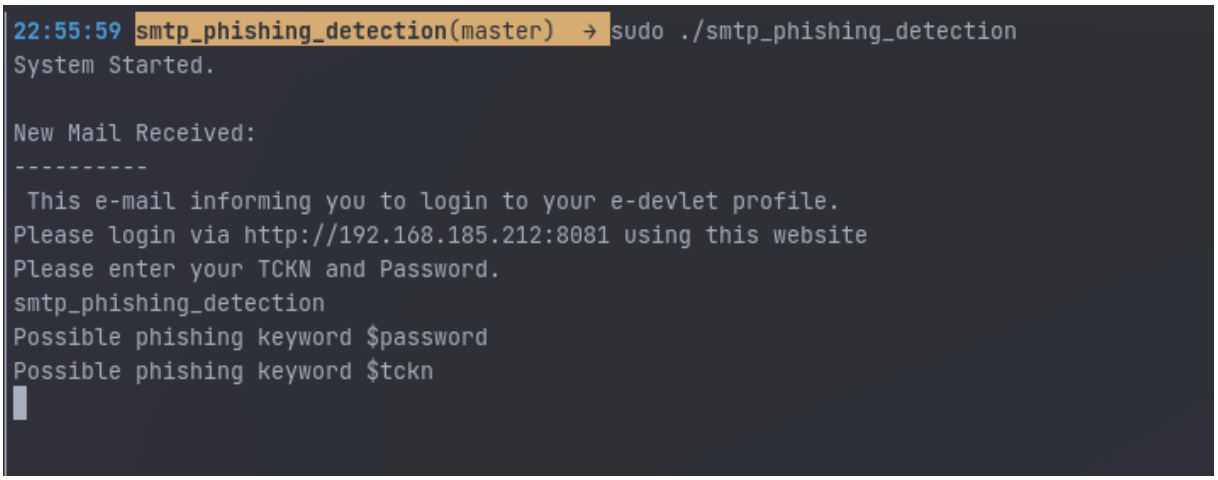

Figure 3.5 Detected Keywords With Regex Matching Using Yara Tool

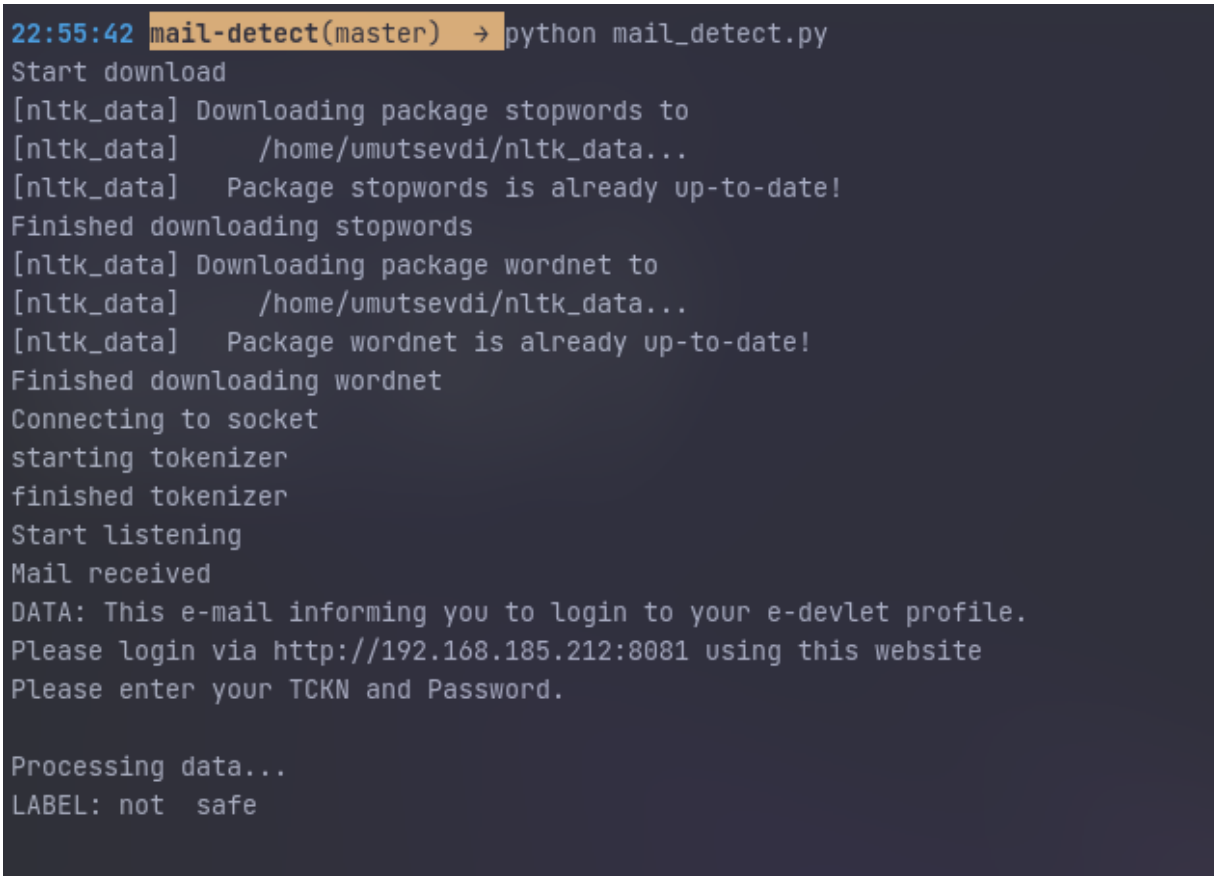

Figure 3.6 NLP Based Detected Result