**Phishing detection extension**

***מגישים : סאלח אבו צאלח 315810101 , מוחמד מחאמיד 314993130***

1. תוכנית עבודה: מצורף מסמך "תוכנית עבודה" מעודכן(final work plan).
2. הסבר על התוכנית לעומק ותמונות של דוגמאות ריצה (עמוד 2 עד 6).
3. מצורף קובץ הקוד שמכיל צד השרת וצד הלקוח(miniProject)

server ,client side - the extension)) שאפשר לריץ אותו לפי ההורואות בקובץ readme.md.

1. ליניקם ורפרנסים שנעזרנו בהם (עמוד 7).

**Phishing detection extension**

Deep in the program

**Server side:** we took the backend from pre-trained machine learning model, which for each URL it receives, it builds a vector of values {-1,1}. each value, represent data about the URL and website itself, and using machine learning to determine how much these values reflect on the decision of whether this URL is phishing or not. the decision is binary, “phishing URL” or” legitimate URL”.

These values indicate the legitimacy of an URL attributes such that:

1) length of the URL (too long return -1 ..).

2) protocol of URL (http ….).

3) domain of URL (.gov .com …).

4) host name of URL (google ,…).

5) path of URL (/login ….).

And by using “Whois” servers that give us information about websites for example: expiration date , first online date , reputation , …

Collecting this information, building the vector and returning binary answer using the model.

Server receives URL through “Get” property, then processing the URL and returning an answer. (the server runs on a local host in our project). the backend is written with python.

**Client side:** the front end, is Google Chrome extension which is written with JavaScript and Html. the interface is pop up style. you have the app icon on upper side which opens small html page and there you have the option to check the current URL “check current tab”. clicking on the button activates the logic side of the app (Java script). it takes the current URL using chrome API and send request to the server to check the URL then answer will be printed on the html pag. its that simple! .

**-we will demonstrate how the server works, how the extension works and see some examples and one example where the app fails.**

- first, we run the server - after installing all the requirements, we run the command line in the Project directory: “python3 server.py”

A close up of text on a white background

Description generated with very high confidence

-second, we run the extension, after loading it to chrome we press the green button on top right.

A close up of a piece of paper

Description generated with high confidence

Examples: google.com A screenshot of a computer

Description generated with very high confidence

Phishing site 1

A screenshot of a cell phone

Description generated with very high confidence

Phishing site 2

A screenshot of a social media post

Description generated with very high confidence

Fail example:

A picture containing indoor

Description generated with very high confidence

-here we can see that URL is phishing PayPal info but the app says its legitimate URL, the app fails this one but that doesn’t happen often .

**overview**

Issues we faced:

We have tried more than 3 modules to help us to get the best results we can. And that enabled us dig deep into the modules to get the best out of them. (we read about machine learning in general in able to find the best module).

Strong sides of the program:

1. It uses pre trained machine learning module to decide for each URL, so no need for database to store phishing URLs, and the program will detect the new phishing sites.

2. The google chrome extension is so user friendly so it just require one click from the user and extension do the rest of the work .

Weak sides of the program:

Phishing websites may develop new techniques to not get caught and the module may not cover all these techniques and will not be able to detect some of these sites.

**RESOURSES**

1. <https://openphish.com/> ……………………………….phishing examples.(live sites thus they may change URLs)
2. <https://github.com/swagster420/Phishing-Url-Detection-Using-Machine-Learning>..................Machine learning model code and server code but we modified it and fixed some issues.
3. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6731669> …………paper about phishing website and machine learning , we read some of it to understand the logic behind the ML.