BVRIT
HYDERABAD
College of Engineering for Women
Department of Information
Technology Project- Academic Year

TEAM 3

DETECTION OF MALICIOUS URL's

Abstract

The World Wide Web has become an application on the Internet which brought in an immense risk of cyber-attacks. The more we grow in technology over internet the more we get exposed to malicious data which leads to cyber-attacks. With the rise in the underground Internet economy, automated malicious programs popularly known as malwares have become a major threat to computers and information systems connected to the internet. Malicious URLs have been widely used to mount various cyber-attacks including spamming, phishing and malware. To avoid this we train a machine learning model that takes a dataset which contains 6,51,191 records of various URLs and then distinguish benign, phishing, malware and defacement URLs. The user gets an alert whether an entered URL is malicious or benign.

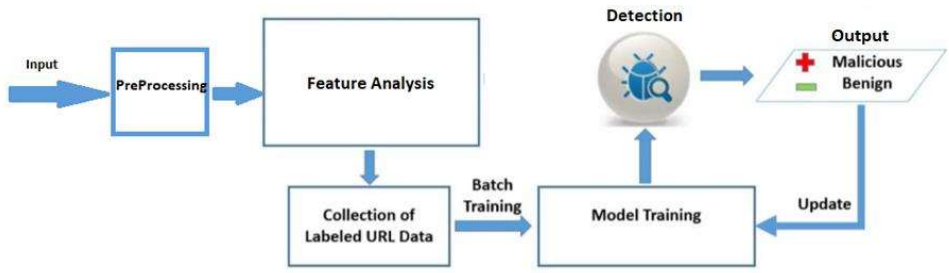
Modules

Feature Extraction

Detection

Interface

Control Flow



Tools and Libraries

- Anaconda3
- Python 3.6
- Flask
- Numpy
- Pandas
- Pickle
- Matplotlib
- Sklearn
- Beautifulsoap

Conclusion and Future Scope

We described a new approach for automatically classifying URLs as either malicious or benign based on machine learning techniques. By analyzing the features extracted from an URL our experiment shows a better approach than blacklist methods, which cannot predict the status of previously unseen malicious patterns. When compared with various classification algorithms we found that the Extreme Boosting algorithm has the best performance. In future work, we would also like to implement a chrome extension for easy use of the application.

Guide

Dr. S L Aruna Rao
Professor & Head, Dept. of IT
arunarao.sl@bvrithyderabad.edu.in

Team


A.Meher Gayatri Devi
18WH1A1206


T.G.S.N Sai Chandana
18WH1A1207


J. Anuhya
18WH1A1240

Github Links:

1. https://github.com/18wh1a1206/MajorProject
2. https://github.com/18wh1a1207/MajorProject
3. https://github.com/18wh1a1240/MajorProject