**SYNOPSIS REPORT**

**on**

**Malware Detection using ML and Python**

A Project Report submitted in partial fulfilment of the requirements for the award of the degree of

**B.Tech CSE with LLB(H) in Cyber Laws**

**Submitted by**

Isha Mittal (Enroll No: R120218007)

Ramya Mihir (Enroll No: R120218011)

**Under the guidance of**

Ms. Dhiviya J Rose

Assistant Professor(SS), SoCS

Department of Applications



**SCHOOL OF COMPUTER SCIENCE**

UNIVERSITY OF PETROLEUM & ENERGY STUDIES

Bidholi Campus, Energy Acres, Dehradun – 248007.

**August 2021**

**ABSTRACT**

There has been an exponential growth in the number of malwares in the cyber world in the last few years. Modern malwares use sophisticated techniques such as polymorphism and metamorphism to thwart the malware detection and analysis. Detecting malware on the basis of their features and behavior is critical for the computer security community. Most anti- virus depends on the signature-based detection which is relatively easy to evade and is ineffective for zero-day exploit based malwares. With this project, we propose a new approach to identify malwares using static analysis, i.e. without executing. With the help of different machine learning models, we will identify malwares and analyses the performance of different models for the same.

**Keywords:** Malwares, Static Analysis, Machine Learning

**PROBLEM STATEMENT**

In today’s world people have no idea about that their devices are infected by malwares that make their computers bots or zombie machines that can-do task directed by the attacker without users’ permission or knowledge and many more tasks like that. Prevention Mechanism based on Hash computation to check if the file is malware have become old and attackers have discovered new ways to avoid them.

**NEED FOR THE PROBLEM**

Over the past few years, The Internet has become a vital part in our lives with increased usage of services like online banking, online reservation etc., and our dependence on the Internet is expected to grow. With the rise of the Internet, there has been huge growth in the amount of malware samples in the wild.

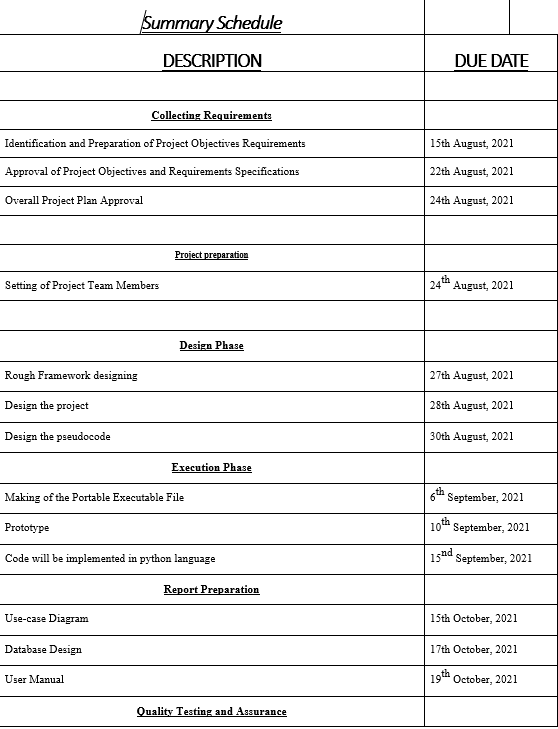
The term Malware is used when referring to a malicious software, which is designed to disrupt or gain unauthorized access to a system. There are various types of malwares, brief description of each is given below:

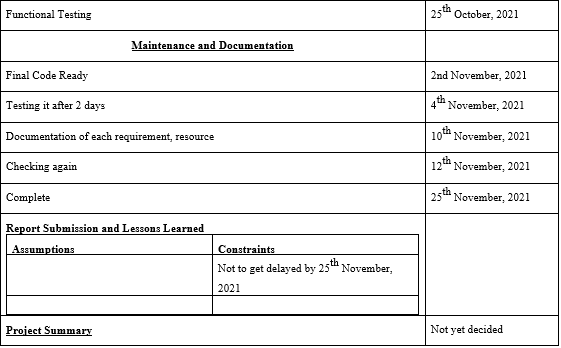
* Viruses and Worms: These are famous for the way they spread to other systems. Virus can infect any normal file or program on the system by attaching itself to it and will infect the whole system as well when executed. On the other hand, worms spread automatically using the computer network and performs various malicious actions
* Spy-Wares: These programs are crafted to intrude on the privacy of users by leaking confidential information like screen-shots, on-line banking credentials or Internet surfing history etc. of users to the third party without user’s knowledge. They get installed on the system very secretly and could be difficult to detect.
* Trojans: Often referred as Trojan Horses, are programs designed to have an appearance of serving a useful utility but actually contains some hidden malicious payload which may be to install spy-wares
* Root-kits: These are collection of tools which are used to gain administrator/root privi- legs on a system by exploiting a known vulnerability in the kernel or through passwords (Brute-forcing commonly used passwords or using social engineering). Once installed, they are very difficult to detect.

It is therefore required to differentiate between a malware and a clean ware once an unknown file has entered into the system in order to ensure it will not perform any malicious activity. Several ways by which a malware could enter in your system include:

* Downloading a file from Internet.
* Opening attachments in email or social networking websites such as face-book.
* Plugging external hard drives, USB and CDs to the system.
* This project aims to perform malware detection, where models will be trained on large corpus of executable using good set of discriminative predictors extracted through static analysis

**PERT CHART**

****

****

**REFERENCES**

* https://media.kaspersky.com/en/enterprise-security/Kaspersky-Lab-Whitepaper-Machine-Learning.pdf
* https://www.ccsinet.com/blog/machine-learning-malware-analysis/
* https://www.imperva.com/learn/application-security/malware-detection-and-removal/