RAMAIAH
Institute of Technology

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CSP: PROJECT WORK TERM: Jan-May 2019

PROJECT SYNOPSIS

Submitted to
Dr. S. Rajarajeswari

PROJECT TEAM MEMBERS

Sl. No	USN	Name
1	1MS15CS002	Aatish Kayyath
2	1MS15CS005	Abishek Padaki
3	1MS15CS024	Aravind P Anil
4	1MS15CS040	Devika Anil

M.S. RAMAIAH INSTITUTE OF TECHNOLOGY
(Autonomous Institute, Affiliated to VTU)

Malware Detection System using Deep Learning Techniques

Problem Statement

The project aims at developing a model which accurately predicts the probability that an operating system will be hit by a malware. Its primary goal is to malware detection among many operating systems and building a model which can accurately do the task at hand.

Objective

There are two main types of defence against malware - malware detection and malware prevention. We focus on the former and remove the problem before it even arises. There are over a billion potential systems in the world that potentially be affected by malware. Designing a technique to prevent attacks from all different types of malware can be extremely difficult as compared to detecting the causes of malware and predicting if a machine will be hit with malware.

We work on an operating systems dataset that has over 7 million recorded operating systems and their various features. We create a model that tries to predict which of those operating systems will be hit with any type of malware. Essentially, the model will predict the vulnerability of the system and the probability that the system will be affected soon. The dataset will have actual results of the systems after receiving status of if they're affected which is when we can decide and calculate the accuracy of our model.

Rather than inspecting malware, we can get much better results by analyzing the system and its own vulnerabilities against targeted attacks from malware. The outcome is a predictive model that outlines the probability of a system with many parameters being affected in the near future or not. A generic solution that fits all the systems will help plug holes in many security systems - primarily windows machines as they are the most used operating system on the planet currently.

Scope

Malware is software that is aimed at intentionally causing a system to behave in a way that it should not. Its main objective is to disrupt the system's normal functioning or cause damage to the system itself and its components or both. Malware comes in many different forms, but irrespective of the type, their objective is one - damage to a system.

Malware attack is a very large domain of cybersecurity attacks. Cryptanalysts across the world has been in a long drawn out battle which has intensified in the past decade with malware. With increase in the technological capabilities of computers, there is also a distinct sharp increase in the capabilities of what malware can do and more importantly, how a malware can prevent from being detected. This is what our project is aimed at - Malware detection.

Methodology

Prevention is better than cure but we take it a step further to even avoid prevention and stop the chance of a malware infection by assessing the vulnerability of the system.

If we check and analyze different types of malware that exist at this point in time, we may come up with a very efficient system to guard against attacks. However, this is very difficult as it requires documentation of every malware that has ever existed not to mention the very complicated and intricate system which requires different techniques against different types of malware. This makes malware detection essentially a time-series problem, but it is made complicated by the introduction of new machines, machines that come online and offline, machines that receive patches, machines that receive new operating systems, etc.

This leaves us in a predicament - stopping malware attacks by preparing ourselves against the different types of malware is practically impossible because it has to done exhaustively. Not doing so puts the system at great danger which defeats the purpose of defending a malware attack. Also, when technology grows leaps and bounds, so does the capability of malware. True exhaustive analysis of malware cannot be done in theory or practice. The list of malwares is constantly growing and evolving. True preparation against attacks takes a whole different approach.

This approach is assessing vulnerability of the system rather than the attacker. If the attacker is constantly evolving and learning new techniques against the system's defense, then efforts to defend against certain types of attacks are futile. Hence, we try to predict an attack in a more generic sense before it has even happened by assessing the system itself. This theory is mainly based on the assumption that malware is targeted. Even though there are variants, a malware always targets a vulnerability or an exploit of the system to attack. If we can find out these weak points on the system and patch them before an attack happens, we can develop a very secure and malware proof security configuration.

Contribution to the Society

The only constant in technology is change. It is forever evolving and never stops improving. The growth is exponential and each new technological advancement is as fascinating as the previous era, if not more. However, it has its own downsides, namely technology being used in ways it should not be intended to. Malware, like technology has evolved. What started as pranks to annoy coworkers many years ago are now in a form that is aimed at causing massive destruction or theft of valuable information. Unlike tangible crimes, it is possible to perform a malware attack without leaving even so much as a digital fingerprint. Hence, protection against malware becomes a priority.

The value of data and information is forever increasing. With increase in data storage technology, cheaply available storage and computation resources and sudden increase in availability of the internet to the general public, every organization and civilian is trying to convert their valuable data as electronic data. Hence, it becomes vital to focus our efforts on protecting valuable assets.

The model we build will not only help the fight against malware a massive upper hand but also help with stopping an infection before it starts. By doing analysis on historic data available, the accuracy of the model can be improved because it gives us an insight into what makes a system vulnerable.

Signature of the Guide