Universiti Teknologi MARA

Keylogger Detection Analysis Using Machine Learning Algorithm

Muhammad Faiz Hazim Bin Abdul Rahman

Thesis submitted in fulfilment of the requirements for Bachelor of Computer Science (Hons.) Data Communication and Networking Faculty of Computer and Mathematical Sciences

July 2022

SUPERVISOR'S APPROVAL

KEYLOGGER DETECTION ANALYSIS USING MACHINE LEARNING ALGORITHM

By

MUHAMMAD FAIZ HAZIM BIN ABDUL RAHMAN
2019819702

This thesis was prepared under the supervision of the project supervisor, Ts Dr Abidah Hj Mat Taib. It was submitted to the Faculty of Computer and Mathematical Sciences and was accepted in partial fulfilment of the requirements for the degree of Bachelor of Programme's Name.

Approved by:

Ts Dr Abidah Hj Mat Taib
Thesis Supervisor

Approved by:

Nor Alifah Bt Rosaidi
Project Co-Supervisor

JULY ,2022

STUDENT DECLARATION

I certify that this thesis and the project to which it refers is the product of my own work and that any idea or quotation from the work of other people, published or otherwise are fully acknowledged in accordance with the standard referring practices of the discipline.

……

MUHAMMAD FAIZ HAZIM BIN ABDUL RAHMAN

2019819702

JULY 14, 2022

ABSTRACT

Malware is one of the most harmful forms of attack on computers because of its passive approach and hidden execution. The most widespread type of malicious software that discreetly monitors user activities and logs keystrokes is called keylogging malware. Accordingly, the goal of this study are to create a detection model based on both supervised machine learning on keylogger dataset.Plus, to analyse the efficiency of a detection model on keylogger dataset by evaluating a selection of attributes.Besides, to test the accuracy of detection models on keylogger dataset comparing two machine learning algorithms. This study is carried out through the utilisation of two machine learning techniques, namely Decision Tree and Naive Bayes, on Jupyter Notebook in order to conduct an analysis of the Keylogger Detection dataset obtained from a trustworthy website known as Kaggle. There are a few outcomes that have been achieved to decide between those two machine learning methods that have better accuracy to carry out analysis on the dataset which of the two, but rather Decision Tree, have the greater accuracy. Early identification of a keylogger malware attack could prevent hackers from accessing personal user data and reduce the likelihood of infiltration, which could reveal account information, credit cards, usernames, passwords, and other data. In this way, we can decrease the likelihood of being the victim of a spyware attack and losing our information. It is intended that this initiative would deliver benefits to all of the users and be useful to them.

Table of Contents