Ministry of Higher Education and Scientific Research, Iraq

University of Babylon

information technology collage

Information Security Department

Morning Study

Detecting Keylogger Using Machine Learning

A Graduate Project Submitted to the Department of Information Security of the College of Information Technology, University of Babylon, in Partial Fulfilment of the Requirements for the Bachelor's degree in Information Security of Information Technology.

STUDENT'S NAME

Hussein Haider Ali

Supervised by

Assist. Lec. Hassan Abdulameer Hassan

2023-2024

ABSTRACT

In today’s world, the field of information technology is rapidly evolving. Maintaining security and privacy is a major problem for cyber professionals. According to studies, the quantity of new malware is rapidly increasing. A keylogger is a highly sophisticated malware that records every keystroke made on the machine, allowing the attacker the potential to steal enormous amounts of critically sensitive information invisibly without the authorization of the message’s owner. Identifying keylogger is important to avoid data loss and sensitive information leaking. Anti-viruses can detect keylogger via heuristic and behavior analysis, but if the keylogger is not a Known threat, antivirus or anti-malware software cannot detect it as a virus. Machine learning is effective in detecting malware. This project seeks to detect each application’s set of rights and storage levels and distinguish between programs with proper access and keylogger applications that can misuse permissions. This keylogger detection technique is fully black-box; it is based on behavioral traits that are universal to all keyloggers and do not rely on the keylogger’s internal structure. In this research, a keylogger detection model has been proposed using machine learning to detect the keylogger and spyware. The model has been trained on keylogger and spyware data-set to identify the host behavior during keylogger running on the system. The results will be evaluated by using several metrics and presented based on the classification report and confusion matrix to identify system success in detecting keylogger spyware.