Static Analysis of Security Issues of the Python Packages Ecosystem

Authors : Adam Gorine, Faten Spondon

Abstract : Python is considered the most popular programming language and offers its own ecosystem for archiving and maintaining open-source software packages. This system is called the python package index (PyPI), the repository of this programming language. Unfortunately, one-third of these software packages have vulnerabilities that allow attackers to execute code automatically when a vulnerable or malicious package is installed. This paper contributes to large-scale empirical studies investigating security issues in the python ecosystem by evaluating package vulnerabilities. These provide a series of implications that can help the security of software ecosystems by improving the process of discovering, fixing, and managing package vulnerabilities. The vulnerable dataset is generated using the NVD, the national vulnerability database, and the Snyk vulnerability dataset. In addition, we evaluated 807 vulnerability reports in the NVD and 3900 publicly known security vulnerabilities in Python Package Manager (pip) from the Snyk database from 2002 to 2022. As a result, many Python vulnerabilities appear in high severity, followed by medium severity. The most problematic areas have been improper input validation and denial of service attacks. A hybrid scanning tool that combines the three scanners bandit, snyk and dlint, which provide a clear report of the code vulnerability, is also described.

Keywords : Python vulnerabilities, bandit, Snyk, Dlint, Python package index, ecosystem, static analysis, malicious attacks
Conference Title : ICCSPS 2023 : International Conference on Computer Science, Programming and Security
Conference Location : London, United Kingdom
Conference Dates : March 16-17, 2023