MODEL AND METHOD FOR PREDICTING SOFTWARE VULNERABILITIES

mjr mgr inż. Artur STACHURSKI

The statistical analyses of vulnerabilities reveal their presence in almost all commonly available software. Nonetheless, despite having an extensive array of technical capabilities to detect vulnerabilities in software, establishing analytical dependencies is not a trivial process. Therefore, ensuring software security implies the need to have predictive information about its potential vulnerabilities.

This dissertation presents a description of the proprietary probabilistic models and methods for predicting software vulnerabilities, using selected properties of the Markov chain. It is limited to modelling the vulnerability prediction process for customised software, designed to meet the particular information needs of an organization. It is assumed that the collected data on vulnerabilities is largely the result of conducting a wide range of security tests. As a result of applying the presented models and methods, the following characteristics will be estimated:

- general software security level;
- intensity of critical vulnerability occurrence;
- effectiveness of countermeasures against vulnerabilities;

As an extension of the presented mathematical models, a machine learning-based approach for allocating newly discovered vulnerabilities to predefined vulnerability categories was proposed. To discover the optimal approach, several machine learning models were utilized and compared. These models were trained using data from the NVD (*National Vulnerability Database* - maintained by NIST) reference vulnerability database.

The use of the proposed mathematical models and methods is intended to improve the qualitative and quantitative analysis of the collected data. As a result they may allow for a more comprehensive and more accurate assessment of software security level, thus determining possible transitions of the software from a secure state to a vulnerable state.