Predicting Security of Data Protocols

by Matus Korman, KTH (korman@kth.se)

One-page abstract for SWITS 2016

Context. Designing and writing software is an error-prone process, which almost unavoidably leads to the introduction of defects. A subset of such software defects not only threaten the reliability of software products; they also pose vulnerabilities (with respect to cyber security). A software vulnerability, alone or in combination with other vulnerabilities or special circumstances, can be misused by an attacker to cause undesired behavior, malfunction, or remote execution of arbitrary code, which can result in the execution of a legitimate program being suddenly taken over or complemented by the execution of illegitimate harmful code "injected" by an attacker through a piece of data. A famous example event that tends to lead to such conditions is a *buffer overflow*. The National Vulnerability Database[1] (NVD) indexes about 75 thousand software vulnerabilities, and over a dozen new ones are being published on a daily basis. However, those numbers represent just the tip of an iceberg, since only a few of all actual software vulnerabilities are discovered and made public.

Introduction and the problem. Vulnerable pieces of software are often related to routines of data manipulation, especially parsing. However, there are different data protocols that are being implemented by software – protocols used in handling web content, e-mails, remote terminal access, time synchronization, telemetry and industrial control, financial transactions etc; however, not necessarily limited to those networked ones – data compression formats, media codecs, etc. Such protocols have distinctions – for example different complexity, different types and characteristics of communication (data exchange), usage in different environments and by different user groups. All of these attributes may affect the security of the specific operation of a software implementation of a data protocol. However, we currently do not have a means to predict security from the attributes of data protocols.

Aim of the study. This study intends to explore the relations between complexity and other attributes of data protocols on the one side, and security of the usage of data protocols on the other. This study has the ambition to formulate a predictive model that infers the latter from the former.

Current state and preliminary insights. The work is currently ongoing and I intend to make this theme the topic of my dissertation. Some preliminary results have shown discernible differences in the spectra of software weaknesses (using Common Weakness Enumeration[2] (CWE)) found related to a selection of nine different data protocols (related through software implementations of the protocols and software vulnerabilities (using Common Vulnerability Enumeration[3] (CVE))).

Significance of the work. The significance of the work can be seen as two-fold. Firstly, such a predictive model might aid architects, security analysts and defenders in knowing where to most probably expect software vulnerabilities, and how to minimize that probability in their systems and architectures. Secondly, such a predictive model might aid attackers in making attack decisions (e.g., choice of targets, choice of attack tactics) that lead to a higher probability of success.

[1] https://nvd.nist.gov

[2] https://cwe.mitre.org

[3] https://cve.mitre.org