Code Signing Project Proposal

Mason Hyman and Julia Steed

***Abstract*— This is a proposal for a project analyzing current code-signing public key infrastructure (PKI). The goal of this project is to analyze common errors make by Certificate Authorities (CAs) when issuing certificates and their impact on PKI overall.**

I. OBJECTIVE

The objective of this project is to evaluate code-certificate misissuance by certificate providers. Certificate Authorities (CAs) regularly make errors when issuing certificates that can affect the overall security of public key infrastructure (PKI). In order to quantify these errors, our project will utilize linting procedures to comb through a vast data set of code-signing certificates. The intended result of this project would be to display the most common errors made by CAs and their larger impact on the security of PKI.

II. MOTIVATION

The security of digital assets and software integrity is of extreme importance in today's world. Code-signing certificates represent a critical aspect of data security. They are the digital signatures that vouch for the authenticity and trustworthiness of software and data sources.

Certificate authorities (CAs) are responsible for verifying the identities of websites and issuing digital certificates as part of PKI. CA's often fail to correctly issue certificates due to implementation errors and indifference. We plan to analyze the errors authorities make and consider their impact on PKI security as a whole. We aim to deepen our understanding of data science, PKI challenges, and ways to improve digital security.

III. DATA SOURCES

The data used in this project will come from the following sources:

- **VirusTotal**: VirusTotal provides access to a diverse set of malware samples, including those with code-signing certificates. It can analyze file and URLs for viruses, worms, trojans, and other malicious content.
- **AlienVault** AlienVault offers valuable threat intelligence data, including malware-related information. It analyzes security event data from sources such as network traffic, log files, and host-based data to identify suspicious or malicious activity. We will use AlienVault's capabilities to collect, analyze, and monitor data related to code-signing certificate revocations and security events.
- **Dr. Doowon Kim's Code Signing Certificate Dataset** Dataset provided by Dr. Kim

IV. MEMBER RESPONSIBILITIES

- **Mason Hyman**
  - Data collection and analysis
  - Project report and presentation
- **Julia Steed**
  - Data collection and analysis
  - Project report and presentation

V. MILESTONES

Milestone 1

- **Description** Analyze current accepted baselines for the format of code-signing certificates.
- **Due Date** 10/15/23
- **Responsible Members** All

Milestone 2

- **Description** Establish linting program to parse basic fields of each certificate.
- **Due Date** 10/29/23
- **Responsible Members** All

Milestone 3

- **Description** Complete linting program for all fields.
- **Due Date** 11/12/23
- **Responsible Members** All

Milestone 4

- **Description** Perform Analysis on data derived from lint program.
- **Due Date** 11/26/23
- **Responsible Members** All

Milestone 5

- **Description** Create project presentation and write report.
- **Due Date** 12/3/23
- **Responsible Members** All

VI. EXPECTED OUTCOME

The expected outcome of the project would be to create a lint that can successfully detect and categorize errors made in the certificate signing process. Additionally, an analysis will be performed on the data from the linting process such that meaningful results are drawn. At a high-level, this project will give insight into how CA's impact the security of code-signing PKI through misissuance errors.