

Similarity Hashing of EVM Bytecode

Raphael Nußbaumer

October 2020

1 Introduction

- ”Identifying almost identical files using context triggered piecewise hashing” [1].
In the paper ”Automated labeling of unknown contracts in Ethereum” [3]
they used the mean of three distance measures (Levenshtein, Jaccard and
Sorensen) like described in the referenced Papers
”Detection of Malicious Portable Executables Using Evidence Combinational
Theory with Fuzzy Hashing” [2]
and ”Static analyzer of vicious executables (SAVE)” [4]

References

- [1] Jesse Kornblum. ”Identifying almost identical files using context triggered piecewise hashing”. In: *Digital investigation* 3 (2006), pp. 91–97.
- [2] Anitta Patience Namanya et al. ”Detection of malicious portable executables using evidence combinational theory with fuzzy hashing”. In: *2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud)*. IEEE. 2016, pp. 91–98.
- [3] Robert Norvill et al. ”Automated labeling of unknown contracts in ethereum”. In: *2017 26th International Conference on Computer Communication and Networks (ICCCN)*. IEEE. 2017, pp. 1–6.
- [4] Andrew H Sung et al. ”Static analyzer of vicious executables (save)”. In: *20th Annual Computer Security Applications Conference*. IEEE. 2004, pp. 326–334.