Simplification of Mixed Boolean-Arithmetic Obfuscated Expression

University of New Hampshire

Dongpeng Xu

PLDI 2021 · ARTIFACT EVALUATED usenix ASSOCIATION PASSED · Artifacts Available · Artifacts Evaluated Reusable

30TH USENIX SECURITY SYMPOSIUM · EMNLP 2021 Online & in the Dominican Republic Conference on Empirical Methods in Natural Language Processing

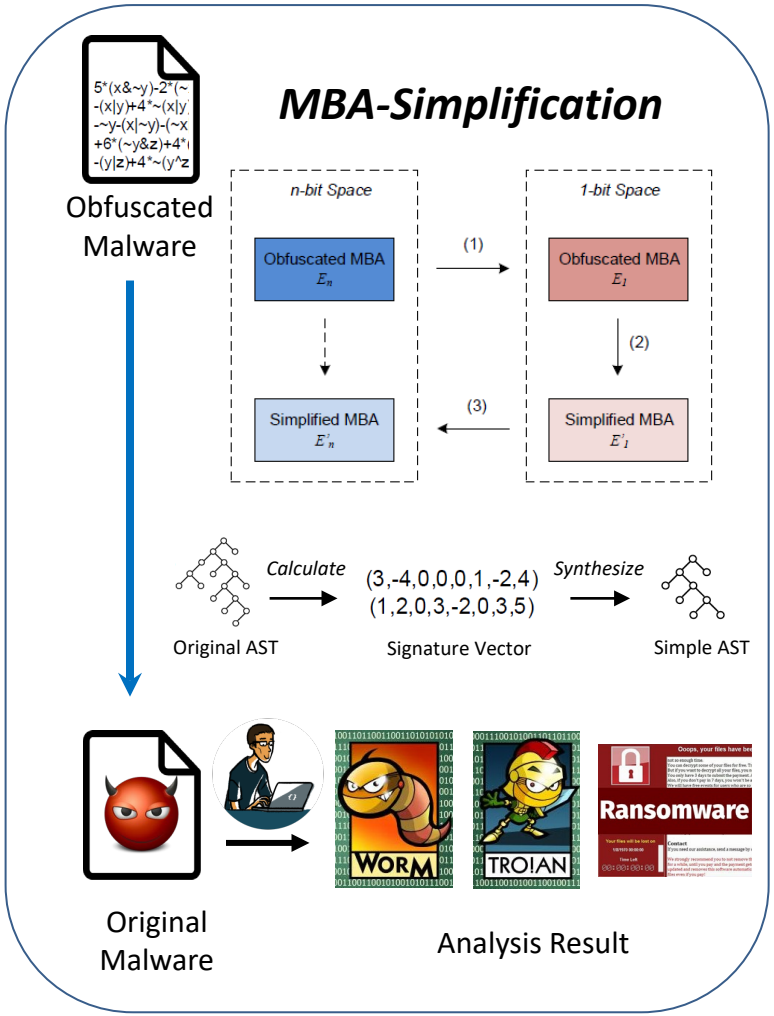
Challenge:

- Many malware packers and obfuscators use Mixed-Boolean-Arithmetic (MBA) expressions
- Existing simplification methods cannot handle MBA

$$x+y \rightarrow 2(x \vee y)-(\neg x \wedge y)-(x \wedge \neg y)$$

Solution:

- Discover important math features: n-bit to 1-bit equivalent transformation
- Semantic preserving translation to reduce MBA-alternation



Scientific Impact:

- Help malware analyzers understand packed malware
- Largely advance the security community's understanding of MBA's inner mechanism
- Boost SMT solver's performance on solving MBA expressions

Broader Impact and Broader Participation:

- Fight against malware
- Open-source
- Research methods and findings have been used in GenCyber K-12 summer camp and university courses

Project Number: 1948489
Project Type: CRII
Contact: dongpeng.xu@unh.edu

NATIONAL SECURITY AGENCY · NSF · GenCyber