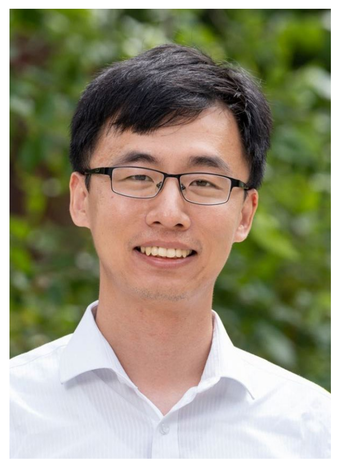
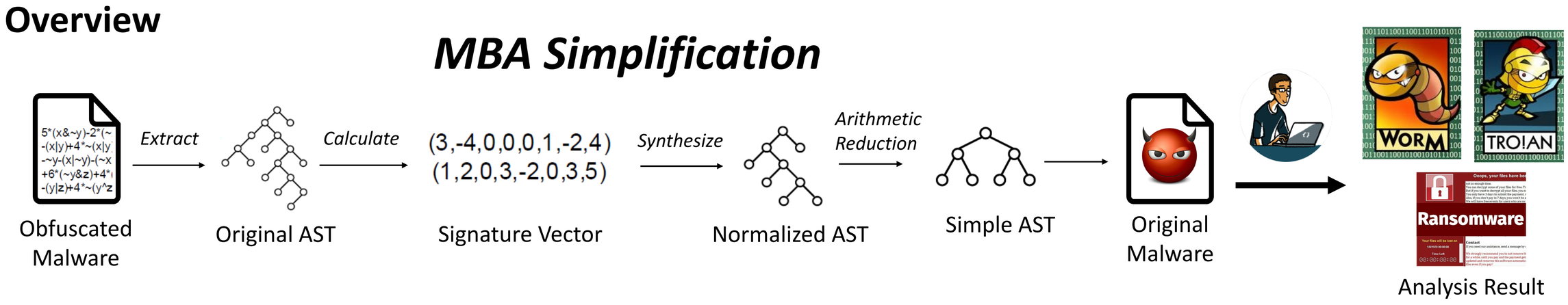
CRII: SaTC: Simplification of Mixed Boolean-Arithmetic Obfuscated Expression



Dongpeng Xu, University of New Hampshire

Project URL (open-source): https://github.com/softsec-unh/MBA-Blast, https://github.com/softsec-unh/MBA-Solver

Overview



Challenges:

- Many malware packers and obfuscators use Mixed-Boolean-Arithmetic (MBA) expressions, e.g., $x+y \rightarrow 2(x \vee y)-(\neg x \wedge y)-(x \wedge \neg y)$
- Existing simplification methods cannot handle MBA

```
int fun(int x,int y,int z)
{
  int c;
  c = x+y;

  return c;
}
```

(a) Original program.

```
int fun(int x,int y,int z)
{
  int c;
  c = 4*(~x&y)-(x^y)-(x|y)
      +4*~(x|y)-~(x^y)-~y-
      (x|~y)+1+6*x+5*~z+
      (~(x^z))-(x|z)-2*~x-
      4*(~(x|z))-4*(x&~z)
      +3*(~(x|~z));

  return c;
}
```

(b) MBA obfuscated program.

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

30TH USENIX SECURITY SYMPOSIUM

ARTIFACT EVALUATED usenix ASSOCIATION PASSED

PLDI 2021

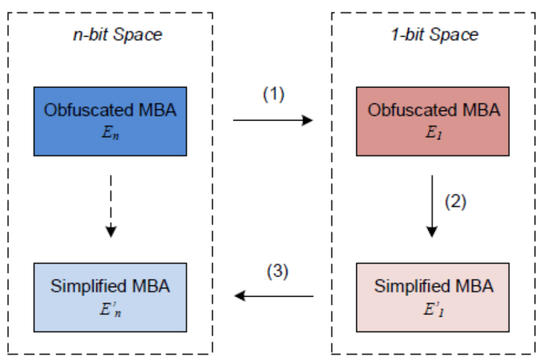
EMNLP 2021 Online & in the Dominican Republic — Conference on Empirical Methods in Natural Language Processing

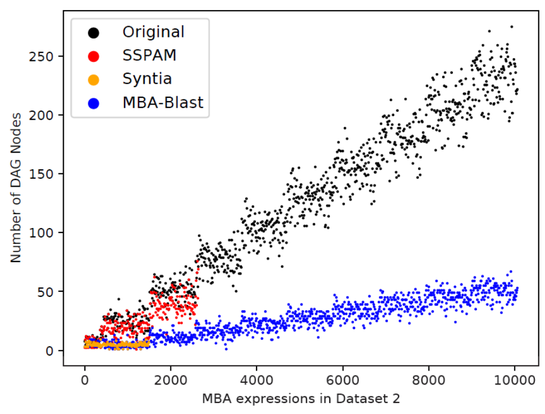
Artifacts Available acm

Artifacts Evaluated acm Reusable

Solutions:

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





Broader Impact: Society

- Advance software de-obfuscation
- Fight against packed malware
- Open-source, publicly available

Tigress

VMProtect software

Z3 Theorem Prover

STP constraint solver
Simple Theorem Prover SMT solver

VIRUSTOTAL

Broader Impact: Education and Outreach

- GenCyber K-12 summer camp
- Undergraduate courses: CS 527, CS 727
- Graduate courses: CS 827, CS 927