Toward Softer Signatures

Dennis Goodlett

Thanks

- My employer Hurricane Labs
- My Wife
- Barton Rhodes
- 23 gr8 Escape author

Gr8 Escape

- Pwn challenge, day 23 of OverTheWire CTF
- ELF, 64-bit x86
- Statically compiled and stripped
- VM implementation
- Challenge code complex enough to make it hard to distinguish from libc
  - At least for me...

Zignature Introduction

- A zignature is a piece of metadata associated to a function which may be used to identify that function later
- Radare2 supports FLIRT databases
- Radare2 has it’s own zignature database in sdb files
- Radare2 has multiple zignature types

Signature Philosophy

- Magic Feature!
- Attack statically compiled programs
- Minimise false positives
  - Require a perfect match to apply information
- Gotta Go fast!

Signature Matching Stinks

Signature Fail when...

- Signatures created from wrong version, linux distro, compiler, etc
- One bit changing breaks some signatures:
  - rasm2 -d 31c0 -> xor eax, eax
  - rasm2 -d 33c0 -> xor eax, eax
- Padding added
- Compiler optimizations
- Bugs in r2
- Obfuscation/Encryption

Introducing Softer matching

`zb?`

Zb philosophy

- I know the function is here, where is the best place to look for it
- Focus on finding particular functions at the reverser request
- Relies on the reverse engineer to verify findings
- Find potential matches with a less than ideal signature database
- Display top 5 closest matches to give perspective on how well a signature matches compared to others

Example Time

Graph Signatures (G)

- Basic info on graphs, just 5 integers
  - Ex: za sym.malloc g cc=32 nbbs=45 edges=69 ebbs=4 bbsum=736
- Less particular then byte signature so it is already a bit fuzzy
- All the information are numbers, so similarity is done with some simple math
- The “simple” math, see libr/anal/sign.c:matchGraph
- Similarity can quickly be computed
- Each field is treated with equal weight.

Byte Signature (B)

- Usually, a lot more information then Graphs
- Levenshtein distance used for comparison
  - Number of edits, additions, deletions required to get from one signature to the other
  - Divided by max signature length to produce a number between 1.0 and 0.0
- Cleanly deal with misalignment (due to addition/deletion)
- SLOW...

Avoid Levenshtein as much as possible

- The difference in signature lengths acts as a lower bound for Levenshtein Distance.
- Lower bound can be combined with Graph similarity to check if Levenshtein is needed.
- Significant speed up in typical case.

Future Work

- Functionality to find best signature file
  - Online database with minimum signatures needed to identify libc version
  - Other ways to intelligently reduce keyspace
- ESIL CFG zignatures! (See Condret's 2019 talk)
- Diff assembly of byte signature with current function
- Rasign2 accept multiple file inputs
- Bug fixes, refactoring, performance