Artifact

BinCrypto: Binary Cryptographic Function Identification via Similarity Analysis with Path-insensitive Emulation

Here is the artifact of BinCrypto. It is provided as a Docker image based on Linux, containing the sample binaries and compiled executables of BinCrypto's prototype. The file tree of the prototype is:

```
|--- binCrypto/
|    |--- bin/                                                  # the dir
ectory of binaries
|    |    |--- x64_libcl347_gcc114_O0
|    |    |--- x64_libcl347_gcc114_O3
|    |    |--- ...
|    |--- data/                                                 # the dir
ectory of binaries' preprocess data
|    |    |--- x64_libcl347_gcc114_O0/
|    |    |--- x64_libcl347_gcc114_O3/
|    |    |--- ...
|    |--- scripts/                                              # the dir
ectory of Python scripts
|    |--- signatures/                                           # the dir
ectory of extracted code features
|    |    |--- x64_libcl347_gcc114_O0/
|    |    |--- x64_libcl347_gcc114_O3/
|    |    |--- ...
|    |--- output/                                               # the dir
ectory of similarity scores
|    |    |--- x64_libcl347_gcc114_O3VSx64_libcl347_gcc114_O0/
|    |--- comparison                                            # the exe
cutable to compare two binaries for similarity scores
|    |--- emulation                                             # the exe
cutable to emulate one binary for code features
|    |--- random.txt                                            # the pre
-defined random input values for emulation
|    |--- Dockerfile                                            # the con
figuration file for Docker
```

0. Start with the Docker Image

```
$ docker pull ruixiongh/bincrypto:latest
```

```
$ docker run --name bincrypto -it ruixiongh/bincrypto:latest
```

- Then, it is under the root path of the prototype.

1. Emulate the Sample Binaries of x64_libcl347_gcc114_O3 and x64_libcl347_gcc114_O0

```
# emulate to extract function signatures
$ python3 scripts/emulate_binary.py x64_libcl347_gcc114_O3
$ python3 scripts/emulate_inary.py x64_libcl347_gcc114_O0
```

- **x64_libcl347_gcc114_O3/O0** is cryptolib v3.4.7 which is compiled with GCC v11.4.0 -O3/-O0 for x64.
- The signature sequence of each function is recorded in the file under the path **(root)/signatures/x64_libcl347_gcc114_O3/** and **(root)/signatures/x64_libcl347_gcc114_O0/**.

2. Compare the Sample Binaries

```
# REF: x64_libcl347_gcc114_O0
# TAR: x64_libcl347_gcc114_O3
# compare each function signature of the TAR to those of the REF
$ python3 scripts/compare_binaries.py x64_libcl347_gcc114_O3 x64_libcl347
_gcc114_O0
```

- The results of similarity scores are under the path **(root)/output/x64_libcl347_gcc114_O3VSx64_libcl347_gcc114_O0/**
  - Each file is named as that of a TAR function.
  - Each line of a file is REF function and the similarity score with the TAR function.