DAFL: Directed Grey-box Fuzzing Guided by Data Dependency

Tae Eun Kim, Jaeseung Choi, Kihong Heo, Sang Kil Cha
Background

Fuzzing
- Testing a program with randomly generated inputs
- Successful achievements
  - e.g., AFL, Google’s OSS Fuzz project
Background

Fuzzing
• Testing a program with randomly generated inputs
• Successful achievements
  • e.g., AFL, Google’s OSS Fuzz project

Directed Fuzzing
• Aims to reach the given target location(s)
  • Generate crashing inputs from bug reports
    (e.g., static analysis alarms)
Directed Grey-box Fuzzing (DGF)

Seed Pool

Achieved new coverage?

Target Program
Directed Grey-box Fuzzing (DGF)

What is the distance of this seed?

More Mutants!!

Achieved new coverage?
Limitations of DGF

1. **Noisy** seed distance based on Control Flow Graph (CFG)
   - Complex control structures (e.g., loops) introduce noise in the seed distance
Limitations of DGF

Noisy CFG-based Seed distance

```c
if (/ * Sanitize f */)
    break;

f = fopen(...);

if (/ * Sanitize f */)
    break;

type = fgetc(f);
if (/* type is valid */)

... 

blockParse(f, type);

... 

Crash(f, type);
```

*CVE-2017-7578 in swftophp*
Limitations of DGF

Noisy CFG-based Seed distance

f = fopen(...);

if (*Sanitize f*)

break;

blockParse(f, type);

type = fgetc(f);
if (*type is valid*)

Noisy CFG-based Seed distance

Limitations of DGF

*CVE-2017-7578 in swftophp*
Limitations of DGF

Noisy CFG-based Seed distance

```c
if (/ * Sanitize f */)
  break;

f = fopen(...);

if (/ * Sanitize f */)
  ...

blockParse(f, type);

break;

Crash(f, type);

type = fgetc(f);
if ( /* type is valid */ )
  Noisy CFG-based Seed distance

Limitations of DGF

*CVE-2017-7578 in swftophp*
Limitations of DGF

Noisy CFG-based Seed distance

**AFLGo:** Distance based on all nodes in CFG (Lower is Better)  

34.5  Average(34, 35)
Limitations of DGF

Noisy CFG-based Seed distance

**AFLGo**: Distance based on all nodes in CFG  (Lower is Better)

- **34.5**  Average(34, 35)
- **37.5**  Average(34, 35, 36, 45)

---

*CVE-2017-7578 in swftophp*
Limitations of DGF

Noisy CFG-based Seed distance

**AFLGo:** Distance based on all nodes in CFG  (Lower is Better)
- 34.5
- 37.5

**WindRanger:** Distance based on Diverging nodes in CFG  (Lower is Better)
- 34

---

*CVE-2017-7578 in swftophp*
# Limitations of DGF

## Noisy CFG-based Seed distance

**AFLGo:** Distance based on all nodes in CFG  (Lower is Better)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>![like]</td>
<td>![dislike]</td>
<td>34.5</td>
<td>37.5</td>
</tr>
</tbody>
</table>

**WindRanger:** Distance based on Diverging nodes in CFG  (Lower is Better)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>![like]</td>
<td>![dislike]</td>
<td>34</td>
<td>45</td>
</tr>
</tbody>
</table>

---

*CVE-2017-7578 in swftophp*
DAFL’s Solution

DUG-based Semantic Relevance Score

DAFL utilizes Definition-Use Graph (DUG) to calculate

Semantic Relevance Score
DAFL’s Solution

DUG-based Semantic Relevance Score

**AFLGo**: Distance based on all nodes in CFG  (Lower is Better)

- ![Thumbs Up] 34.5
- ![Thumbs Down] 37.5

**WindRanger**: Distance based on Diverging nodes in CFG  (Lower is Better)

- ![Thumbs Up] 34
- ![Thumbs Down] 45

---

*CVE-2017-7578 in swftophp*
DAFL’s Solution

DUG-based Semantic Relevance Score

**AFLGo:** Distance based on all nodes in CFG  (Lower is Better)

- 34.5
- 37.5

**WindRanger:** Distance based on Diverging nodes in CFG  (Lower is Better)

- 34
- 45

**DAFL:** Semantic Relevance Score based on DUG  (Higher is Better)

![Diagram](image)
DAFL’s Solution

DUG-based Semantic Relevance Score

**AFLGo**: Distance based on all nodes in CFG  (Lower is Better)

- 34.5
- 37.5

**WindRanger**: Distance based on Diverging nodes in CFG  (Lower is Better)

- 34
- 45

**DAFL**: Semantic Relevance Score based on DUG  (Higher is Better)

- 1  Sum(1)

---

*CVE-2017-7578 in swftophp*
DAFL’s Solution

DUG-based Semantic Relevance Score

AFLGo: Distance based on all nodes in CFG (Lower is Better)
- 34.5
- 37.5

WindRanger: Distance based on Diverging nodes in CFG (Lower is Better)
- 34
- 45

DAFL: Semantic Relevance Score based on DUG (Higher is Better)
- 1 Sum(1)
- 2 Sum(1, 1)

*CVE-2017-7578 in swftophp*
Limitations of DGF

1. **Noisy** seed distance based on Control Flow Graph (CFG)
   - Complex control structures introduce noise in the seed distance

2. **Negative** Coverage Feedback
   - Generate seeds that cover irrelevant program locations
Limitations of DGF

Negative coverage feedback

- Case: CVE-2017-7578 in swftophp
Limitations of DGF

Negative coverage feedback

- Case: CVE-2017-7578 in swftophp
- The promising seed is easily outnumbered.
DAFL’s Solution

Selective Coverage Instrumentation

- Case: CVE-2017-7578 in swftophp
- The promising seed is easily outnumbered.

DAFL selectively receives coverage feedback by Selective Coverage Instrumentation
DAFL’s Solution

Selective Coverage Instrumentation

• Case: CVE-2017-7578 in swftophp
• The promising seed is easily outnumbered.

No seeds are generated from irrelevant functions

Seed Pool

parseSWF_R
parseSWF_RGBA
parseSWFDEFINEMORPHSHAPE
blockParse
outputBlock
SWF_warn

filelen_check 80 other functions

main

parseSWF_RGBA

*CVE-2017-7578 in swftophp
DAFL Overview

Static Analysis Phase

- Static Analyzer
- Relevant Functions
- Selective Coverage Instrumentation

Fuzzing Phase

- Def-Use Graph
- Semantic Relevance Scoring
- Semantic Relevance Score
- Seed Scheduling
- Mutants
- Instrumented Execution

Program with a Target Location

Instrumented Program
DAFL Overview

Out of 41 target bugs,
- Shows the best performance in **27** cases.
- Reproduces bugs **4.99** times faster than SOTA directed grey-box fuzzers.
- Finds **6** more bugs within a given timeout (24H).
Evaluation

Crash Reproduction

• Benchmark
  • 41 CVEs from 10 programs

• Baselines
  • 1 Undirected Fuzzer
    • AFL
  • 3 Directed Fuzzers
    • AFLGo
    • WindRanger
    • Beacon

• Criteria
  • Median time of 40 iterations to reproduce the target bug
Evaluation

Crash Reproduction

Best performance in 27 cases.
4.99 times faster than the baseline DGF
Finds 6 more bugs within a given timeout (24H).
Summary

• Directed Grey-box Fuzzing
  • Limitations
    • Noisy Seed Distance
    • Negative Coverage Feedback

• Solution: DAFL
  • Directed Grey-box Fuzzing Guided by Data Dependency

• Key Concepts of DAFL
  • Semantic Relevance Scoring
  • Selective Coverage Instrumentation

→ Achieves 4.99 times performance boost against the SOTA Directed Grey-box Fuzzers

Link to our artifact!!