Pythia: Grammar-Based Fuzzing of REST APIs with Coverage-guided Feedback and Learning-based Mutations

w/ Roxana Geambasu, Baishakhi Ray (Columbia University) and Patrice Godefroid, Marina Polishchuk (Microsoft Research)
Challenge

➢ Most payload types have a fixed set of possible values
➢ Can we augment this set of possible values?

Example

Tests
Sequence 1
Sequence 2
...

Target Service

Sequence 1
Request types
Payload types
Byte-level representation

| Request 1 | Request 2 | Request 3 |
|-----------|-----------|-----------|
| a         | b         | b         |
| b1 b2 b3...bn | b1 b2 b3...bn' | b1 b2 b3...bn'' |

Sequence 2
Request types
Request payload types
Byte-level representation

| Request 1 | Request 4 | Request 3 |
|-----------|-----------|-----------|
| a         | b         | b         |
| a         | d         | c         |
| b1 b2 b3...bn | b1 b2 b3...bn' | b1 b2 b3...bn'' |
First approach: Random byte-level mutations

Sequence 1

| Request 1 | Request 2 | Request 3 |
|-----------|-----------|-----------|
| a         | b         | c         |
| b1 b2 b3...bn | b1 b2 b3...bn' | b1 b2 b3...bn'' |

Sequence 2

| Request 1 | Request 4 | Request 3 |
|-----------|-----------|-----------|
| a         | b         | c         |
| b1 b2 b3...bn | b1 b2 b3...bn' | b1 b2 b3...bn'' |

...
First approach: Random byte-level mutations

➢ Byte-level mutations may destroy the whole sequence :-(

| Sequence 1 | Sequence 2 |
|------------|------------|
| Request 1  | Request 1  |
| a b b b    | a b b b    |
| b1 b2 b3...bn | b1 b2 b3...bn |
| Request 2  | Request 2  |
| b d e c a  | d e c a    |
| b1 b2 b3...bn' | b1 b2 b3...bn' |
| Request 3  | Request 3  |
| a b b b    | b1 XX b3...bn |
| b1 b2 b3...bn'' | b1 b2 b3...bn'' |

Pythia
byte-level mutation engine

Target Service
First approach: Random byte-level mutations

- Old rule @382:'E'
- New rule @382: '\xa4'

Sent: 'POST /api/v4/projects/997190/repository/branches HTTP/1.1 \nPRIVATE-TOKEN: DRiX47nu\xa4 P2ARa4APFr\n{"ref":"master","branch":"string"}\n'

Recv: 'HTTP/1.1 401 Unauthorized\nDate: Thu, 31 Oct 2019 06:19:53 GMT...\n{"message":"401 Unauthorized"}''

Parsing exception:: Exception Parsing Response item: 'name'

- Byte-level mutations may destroy the whole sequence :-(
Second approach: AST-level mutations

**Sequence 1**

| Request 1 | Request 2 | Request 3 |
|-----------|-----------|-----------|
| a b b b d e c a |
| b1 b2 b3…bn b1 b2 b3…bn' b1 b2 b3…bn'' |

**Sequence 2**

| Request 1 | Request 4 | Request 3 |
|-----------|-----------|-----------|
| a b b a d c a |
| b1 b2 b3…bn b1 b2 b3…bn' b1 b2 b3…bn'' |

...
Second approach: AST-level mutations

Sequence 1

| Request 1 | Request 2 | Request 3 |
|-----------|-----------|-----------|
| a         | b         | b         |
| b1 b2 b3...bn | b1 b2 b3...bn' | b1 b2 b3...bn'' |

Sequence 2

| Request 1 | Request 4 | Request 3 |
|-----------|-----------|-----------|
| a         | b         | b         |
| b1 b2 b3...bn | b1 b2 b3...bn' | b1 b2 b3...bn'' |
Second approach: AST-level mutations

- All AST leafs are flipped randomly
- Are all rules equal?

**Pythia**
AST-level mutation engine

| Sequence 1 | Sequence 1’ |
|------------|-------------|
| Request 1  | Request 1   |
| a b         | a c          |
| b b         | b b b3...bn |
| d e         | b b b3...bn’|
| c a         | b b b3...bn''|

| Sequence 2 | Sequence 2’ |
|------------|-------------|
| Request 1  | Request 1   |
| a b         | a b b3...bn |
| b b         | b b b3...bn’|
| a d         | b b b3...bn''|
| c a         | b b b3...bn’|

Target Service

➢ All AST leafs are flipped randomly
➢ Are all rules equal?
Third approach: Probabilistic AST-level mutations

Sequence 1

| Request 1 | Request 2 | Request 3 |
|-----------|-----------|-----------|
| a b b d e c a |
| b1 b2 b3...bn b1 b2 b3...bn' b1 b2 b3...bn'' |

Sequence 2

| Request 1 | Request 4 | Request 3 |
|-----------|-----------|-----------|
| a b b a d c a |
| b1 b2 b3...bn b1 b2 b3...bn' b1 b2 b3...bn'' |

Production rules

1: S → S + R
2: R → R + P
3: P → P + a
4: a → a2
5: P → P + b
6: b → b3
7: b → b2
8: R → e
9: S → S + R
10: R → R + P
...
Third approach: Probabilistic AST-level mutations

Sequences
Seq1: 1,2,3,4,5,6,5,7,8,1,2...
Seq2: 1,2,3,9,5,10...

Fuzz seq1
Which leaf of seq1 should I flip?
"Probably" leaf1 or leaf2
Train a sequence to sequence model

Seq2seq model

Pythia
Probabilistic mutation engine

Hidden state
Encoder
Decoder

Pythia
Coverage monitor

Target Service

Target Service

Primitive types
Primitive values

S
R
R
R
a
b
b
d
e
c
a

a2
b3
b2
d1
e1
c1
a3
**Selected evaluation results**

- **Q1**: Are tests generated by Pythia increasing code coverage?
- **Q2**: Can Pythia find new errors?

**Case study: Four Gitlab APIs**

| API   | Total requests | Avg. values per primitive type | Avg. dependencies per request type | Path dependencies | Body dependencies |
|-------|----------------|--------------------------------|-----------------------------------|-------------------|-------------------|
| Commits       | 15             | 11                             | 1.7                               | Yes               | Yes               |
| Issues        | 25             | 20                             | 1.8                               | Yes               | No                |
| Groups        | 53             | 2                              | 1.4                               | Yes               | No                |
| Branches      | 8              | 2                              | 1.5                               | Yes               | No                |
Code coverage (Q1)

| API      | Total requests | Avg. values per primitive type | Avg. dependencies per request type | Path dependencies | Body dependencies |
|----------|----------------|-------------------------------|-----------------------------------|-------------------|------------------|
| Commits  | 15             | 11                            | 1.7                               | Yes               | Yes              |
| Issues   | 25             | 20                            | 1.8                               | Yes               | No               |
| Groups   | 53             | 2                             | 1.4                               | Yes               | No               |
| Branches | 8              | 2                             | 1.5                               | Yes               | No               |
Code coverage (Q1)

|   | API       | Total requests | Avg. values per primitive type | Avg. dependencies per request type | Path dependencies | Body dependencies |
|---|-----------|----------------|--------------------------------|-----------------------------------|-------------------|------------------|
|   | Commits   | 15             | 11                             | 1.7                               | Yes               | Yes              |
|   | Issues    | 25             | 20                             | 1.8                               | Yes               | No               |
|   | Groups    | 53             | 2                              | 1.4                               | Yes               | No               |
|   | Branches  | 8              | 2                              | 1.5                               | Yes               | No               |
Code coverage (Q1)

| API     | Total requests | Avg. values per primitive type | Avg. dependencies per request type | Path dependencies | Body dependencies |
|---------|----------------|--------------------------------|-----------------------------------|-------------------|------------------|
| Commits | 15             | 11                             | 1.7                               | Yes               | Yes              |
| Issues  | 25             | 20                             | 1.8                               | Yes               | No               |
| Groups  | 53             | 2                              | 1.4                               | Yes               | No               |
| Branches| 8              | 2                              | 1.5                               | Yes               | No               |
### Code coverage (Q1)

| API       | Total requests | Avg. values per primitive type | Avg. dependencies per request type | Path dependencies | Body dependencies |
|-----------|----------------|--------------------------------|-----------------------------------|-------------------|-------------------|
| Commits   | 15             | 11                             | 1.7                               | Yes               | Yes               |
| Issues    | 25             | 20                             | 1.8                               | Yes               | No                |
| Groups    | 53             | 2                              | 1.4                               | Yes               | No                |
| Branches  | 8              | 2                              | 1.5                               | Yes               | No                |
New errors (Q2)