FASTEN: Fine-grained Analyses of Software Ecosystems as Networks

Sebastian Proksch, Delft University of Technology
The leftpad incident

• A developer removed an NPM library, consisting of just 11 lines of code, over a naming dispute.

• The web broke in response.

https://www.theregister.co.uk/2016/03/23/npm_left_pad_chaos/
The equifax disaster

- Security breach through vulnerability in Apache Struts dependency
- Details stolen for 143M user accounts
- Estimates of >$4B damages
- Patch was available for more than two months

https://www.wired.com/story/equifax-breach-no-excite/
Untamed Use of Dependencies

```
package.json

1  {  
2      "name": "csv-parser",
3      "version": "1.9.3",
4      "description": "Streaming CSV parser that aims for maximum speed",
5      "repository": {  
6          "type": "git",
7          "url": "git+https://github.com/mgrintosh/csv-parser.git"
8      },
9      "dependencies": {  
10         "generate-function": "^1.0.1",
11         "generate-object-property": "^1.0.0",
12         "inherits": "^2.0.1",
13         "minimist": "^1.2.0",
14         "ndjson": "^1.4.0"
15      }
16  }
```

Library

```
package.json

1  {  
2      "name": "ndjson",
3      "version": "1.5.0",
4      "description": "streaming newline delimited json parser + serial",
5      "main": "index.js",
6      "scripts": {  
7          "test": "tape test.js",
8      },
9      "bin": {  
10         "ndjson": "cli.js"
11      },
12      "author": "max ogden",
13      "license": "BSD-3-Clause",
14      "dependencies": {  
15         "json-stringify-safe": "^5.0.1",
16         "minimist": "^1.2.0",
17         "split2": "^2.1.0",
18         "through2": "^2.0.3"
19      }
20  }
```

Dependencies

```
package.json

1  {  
2      "name": "ndjson",
3      "version": "1.5.0",
4      "description": "streaming newline delimited json parser + serial",
5      "main": "index.js",
6      "scripts": {  
7          "test": "tape test.js",
8      },
9      "bin": {  
10         "ndjson": "cli.js"
11      },
12      "author": "max ogden",
13      "license": "BSD-3-Clause",
14      "dependencies": {  
15         "json-stringify-safe": "^5.0.1",
16         "minimist": "^1.2.0",
17         "split2": "^2.1.0",
18         "through2": "^2.0.3"
19      }
20  }
```

Transitive Dependencies

```
package.json

1  {  
2      "name": "ndjson",
3      "version": "1.5.0",
4      "description": "streaming newline delimited json parser + serial",
5      "main": "index.js",
6      "scripts": {  
7          "test": "tape test.js",
8      },
9      "bin": {  
10         "ndjson": "cli.js"
11      },
12      "author": "max ogden",
13      "license": "BSD-3-Clause",
14      "dependencies": {  
15         "json-stringify-safe": "^5.0.1",
16         "minimist": "^1.2.0",
17         "split2": "^2.1.0",
18         "through2": "^2.0.3"
19      }
20  }
```

«depends»
The Sorry State of the Art

• Not much beyond simple package version matches
• No support for assessing updates
• No support for making decisions on which libraries to use
• No support for maintainers

We need to do better than that!
H2020 EU Project: FASTEN

Delft University of Technology

Athens University of Economics and Business

University of Milano

Software Improvement Group

XWiki

Endocode

OW2
FASTEN: Revolutionize Dep. Management

• Improve **Vulnerability Detection**
• Reliable **Impact Analysis**
• Better **License Management**
Dependency Networks

Package Dependency Network

Call Dependency Network
Our Research Goal: Static Analyses at Scale

• Scale Call-Graph Generation
• Use Cases
  • Licensing
  • Impact analysis
  • Vulnerability Detection
• Infrastructure
Incremental Call-Graph Generation

1. Resolve Dependencies
2. Compute (and cache) partial results
3. Merge Into Complete CG

Comparative Accuracy to State of the Art Approaches

61% Speed-up on Cached Call-Graph Generation
Data Processing Pipeline

Kubernetes Cluster
• 4 Machines
• 120 Cores
• 1.25 TB RAM
• 233 TB HDD

Apache Kafka

We have Call-Graphs for 3M+ Maven Packages
Browsable Dependency Information

Browsable Dependency Information
Build-System Integration
Stay Tuned & Get Ready for Public Testing

@FastenProject

https://github.com/fasten-project

https://www.fasten-project.eu/
Summary
Dependency Networks

Incremental Call-Graph Generation

1. Resolve Dependencies
2. Compute (and cache) partial results
3. Merge Into Complete CG

Build-System Integration

Sebastian Proksch
S.Proksch@tudelft.nl
Delft University of Technology
www.fasten-project.eu