Software Supply Chains

Gail Murphy
Univ. of British Columbia
Tasktop Technologies

@gail_murphy
California: Design

Source: Supply Chain 24/7, 09/14
iPhone Supply Chain

Source: Supply Chain 24/7, 09/14
Software Supply Chains
Loose Software Supply Chain

17.2B component requests

>105K suppliers

>834K total components

2014: Central Repository of Java open source components

2015 State of the Software: Supply Chain Report (Sonatype)
Tight Software Supply Chain
All is good?
Outline
Outline

Naïve View
Outline

Naïve View

Reality: Loose Supply Chain
Outline

Naïve View

Reality: Loose Supply Chain

Reality: Tight Supply Chain
Key points:

(re)use is not free
controlled transparency

Caveats:

challenges over solutions
Naïve View
Supply Chain: suppliers, parts, manufacturers, finished goods...

specialized excellence
lower costs
higher quality
Naïve View

Software Supply Chain Spectrum

Bouncy Castle used >> 10K organizations
(vast) majority of developers are part of a software supply chain

suppliers → components → YOU! → software
Naïve View

Loose

Apache TomEE™

>105K
suppliers

>834K
total components

central repository

GitHub project dependences
Loose

+ build products (and other components) faster
+ higher-quality components
+ low cost to (re)use
+ ongoing updates
Naïve View

Tight

multiple tiers of contractually-obligated suppliers

Boeing → General Electric → Hydro-Aire
Naïve View

- higher-quality components
- on-time production
- lower overall product cost
Software Supply Chains

open

faster, better, cheaper

closed

Loose

Tight
Reality View:
Loose Supply Chain
Reality View / Loose Supply Chain

Two Parts
Reality View / Loose Supply Chain

Two Parts

Social
Two Parts

Social

Quality
Reality View / Loose Supply Chain

Two Parts

Social

Quality

S

Q
Social Implications of OSS Library Use

#1  How often does the use of an OSS library lead to a social link between projects?

#2  Do social contributions occur before or after a dependence is introduced on a library?

#3  What kind of social contributions occur?

Palyart and Murphy, 2015, under review
Terminology

user
project/repository

A

technical dependence

B

library
project/repository

social interactions
issue
comments
pull request
commit

Palyart and Murphy, 2015, under review
Data

- 1,125 GitHub repos

Palyart and Murphy, 2015, under review
<table>
<thead>
<tr>
<th><strong>Reality View / Loose Supply Chain</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
</tr>
<tr>
<td>- not a fork</td>
</tr>
<tr>
<td>- public</td>
</tr>
<tr>
<td>- forked at least twice</td>
</tr>
<tr>
<td>- use Maven</td>
</tr>
</tbody>
</table>

Palyart and Murphy, 2015, under review

- 23,059 repositories
- 1,125 GitHub repos
Reality View / Loose Supply Chain

Data

=1,125 GitHub repos

Palyart and Murphy, 2015, under review
Data

- 23,059: not a fork
  - public
  - forked at least twice
  - use Maven

- 17,900: depend on GitHub

- 1,409: > 20 issues
  - issues > 5% pull requests
  - handle account deletions

=1,125 GitHub repos

Palyart and Murphy, 2015, under review
Reality View / Loose Supply Chain

Data

23,059
- not a fork
- public
- forked at least twice
- use Maven

17,900
- depend on GitHub

1,409
- > 20 issues
- issues > 5% pull requests
- handle account deletions

1,227
- high confidence in correct library dependences

=1,125 GitHub repos

Palyart and Murphy, 2015, under review
Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

projects that often have a social link (28%)

Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

Projects that often have a social link (28%)

Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

Projects that often have a social link (28%)

Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

projects that sometimes have a social link (23%)

Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

projects that rarely have a social link (49%)

Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

projects that rarely have a social link (49%)

Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

projects that rarely have a social link (49%)

Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

Palyart and Murphy, 2015, under review
#1 - How often does library use lead to social links?

generally...

the more popular the library,
the less likely developers of a user project are to get involved

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

- Technical Link
- Social Link
#2 - When do social contributions occur related to library use?

- Technical Link
- Social Link

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

- Technical Link
- Social Link

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

- Technical Link
- Social Link

In only 61% of pairs, did technical precede social

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

social before technical

http://www.cs.ubc.ca/~mpalyart/stc_timeline/

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

social before technical

http://www.cs.ubc.ca/~mpalyart/stc_timeline/

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

social before technical

http://www.cs.ubc.ca/~mpalyart/stc_timeline/

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

social before technical

http://www.cs.ubc.ca/~mpalyart/stc_timeline/

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

**time to involvement**

- **social before technical**
  - most interactions within a few months

- **technical before social**
  - more interactions span a longer time

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

**duration of involvement**

- **social before technical**
  - either short involvement or quite long
- **technical before social**
  - most involvement under 5 days

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

**Number of contributions**

- **Social before technical**
  - More contributions, stronger communities?

- **Technical before social**
  - Mostly < 10 contributions

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

Palyart and Murphy, 2015, under review
#2 - When do social contributions occur related to library use?

When social before technical (39%)...
more often closely tied to technical and often more contributions

When technical before social (61%)...
may take a long time for interaction and then the interactions are often quick

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- **Technical Link**
- **Social Link**

User

A

Library

B

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- Technical Link
- Social Link

User

A

Forward

Library

B

Reality View / Loose Supply Chain

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- **Technical Link**
- **Social Link**

35% of pairs seeking help, feature requests, pull requests

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- **Technical Link**
- **Social Link**

User

Library

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- Technical Link
- Social Link

User

A

Library

B

Backward

Reality View / Loose Supply Chain

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- Technical Link
- Social Link

User A → Library B

Backward

30% of pairs

existing social community between projects

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- Technical Link
- Social Link

User

Library

A

B

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- **Technical Link**
- **Social Link**

User A and Library B are connected through Forward & Backward links.

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- Technical Link
- Social Link

User A

Library B

Forward & Backward

35% of pairs

user developers contribute to library
library developers later do pull-request to user project to update library

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

**# developers**

- **BO**: backward only
- **F&B**: forward & backward
- **FO**: forward only

**# social contributions**

- **BO**: backward only
- **F&B**: forward & backward
- **FO**: forward only

---

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

- **BO** = backward only
- **F&B** = forward & backward
- **FO** = forward only

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

---

**# developers**

- **BO**
- **F&B**
- **FO**

**# social contributions**

- **BO**
- **F&B**
- **FO**

**involvement time**

- **BO**
- **F&B**
- **FO**

---

- **= backward only**
- **= forward & backward**
- **= forward only**

---

Palyart and Murphy, 2015, under review
#3 - What kind of social contributions occur?

more backward social contributions than expected and their presence indicates a strong social link

- red = backward only
- green = forward & backward
- blue = forward only

Palyart and Murphy, 2015, under review
Loose Software Supply Chain

often a social cost to using a library

more often than expected cost to being a library
Reality / Tight Supply Chain

Two Parts

Social

Quality

S

Q
Reality / Tight Supply Chain

Two Parts

Social

Quality
Quality Implications of OSS Library Use

- **17.2B** component requests
- **>105K** suppliers
- **>834K** total components

2014: Central Repository of Java open source components

2015 State of the Software: Supply Chain Report (Sonatype)
Quality Implications of OSS Library Use

- 17.2B component requests
- >105K suppliers
- >834K total components

constant updating ~ 3.5 times / yr

2014: Central Repository of Java open source components

2015 State of the Software: Supply Chain Report (Sonatype)
Quality Implications of OSS Library Use
Quality Implications of OSS Library Use

Almost too Big to Fail, Geer and Corman, USENIX 2014
Quality Implications of OSS Library Use

Almost too Big to Fail, Geer and Corman, USENIX 2014
Quality Implications of OSS Library Use

only 41% of vulnerable dependencies remediated
Quality Implications of OSS Library Use

only 41% of vulnerable dependencies remediated
mean-time-to-repair of these was 390 days
Quality Implications of OSS Library Use

Only 41% of vulnerable dependencies remediated. The mean-time-to-repair of these was 390 days. CVSS level 10 - still 224 days to repair.

Almost too Big to Fail, Geer and Corman, USENIX 2014
Quality Implications of OSS Library Use

CVE-2013-2251
CVSS 9.3
Exploitability 10

since identification...

4,076 organizations have downloaded the vulnerable component 179,050 times

2015 State of the Software: Supply Chain Report (Sonatype)
Quality Implications of OSS Library Use

The Legion of the Bouncy Castle

CVE-2007-6721
CVSS 10
Exploitability 10

since identification...

11,236 organizations have downloaded the vulnerable component 214,484 times

2015 State of the Software: Supply Chain Report (Sonatype)
Quality Implications of OSS Library Use

of 240,757 component downloads by large financial or technology firms in 2014…

- 7.5% were of known defective part
- 66% and or those with a defective part, the defects were older than 2013

2015 State of the Software: Supply Chain Report (Sonatype)
Loose Software Supply Chain

(re)use is not free

social and upgrade costs to use
Reality View:
Tight Supply Chain
Reality / Tight Supply Chain

Two Parts
Two Parts

Social
Reality / Tight Supply Chain

Two Parts

Social  Quality

S   Q
Reality / Tight Supply Chain

Two Parts

Social  Quality

S  Q
Tight Software Supply Chain
Tight Software Supply Chain

contractual agreement

contractual agreement
Tight Software Supply Chain

Boeing \[\rightarrow\] General Electric \[\rightarrow\] Hydro-Aire

contractual agreement

contractual agreement
Communication

contractual agreement

contractual agreement

contractual agreement
Communication

contractual agreement

restricted information flow

restricted information flow
Communication

contractual agreement

contractual agreement
Communication

contractual agreement

contractual agreement

Req Change #1
Communication

contractual agreement

contractual agreement

Req Change #1

Req Change #2
Communication

contractual agreement

Req Change #1

Req Change #2

Test Result #3
Communication

contractual agreement

contractual agreement
Communication

contractual agreement

contractual agreement

Req Change #1
Communication

contractual agreement

contractual agreement

Req Change #1

Req Change #2
Communication

contractual agreement

contractual agreement

Req Change #1

Req Change #2

Test Result #3

Reality / Tight Supply Chains
Reality / Tight Supply Chains

Communication

contractual agreement

contractual agreement
Communication

contractual agreement

contractual agreement
Communication

Doors RTC HP Quality Center

Blueprint RTC HP Quality Center

VersionOne Eclipse HP Quality Center
Communication

Doors RTC
HP Quality Center

Schema Mappings

Blueprint RTC
HP Quality Center

Schema Mappings

VersionOne Eclipse
HP Quality Center
Tight Software Supply Chain

need tools to facilitate appropriate communication
Reality / Tight Supply Chain

Two Parts

Social

Quality
Reality / Tight Supply Chain

Two Parts

Social

S

Quality

Q
Tight Software Supply Chain

Boeing → General Electric → Hydro-Aire

ability to verify the brake software wasn’t built in
Tight Software Supply Chain

Boeing → General Electric → Hydro-Aire

full transparency full opacity
Tight Software Supply Chain

controlled transparency

balance need to share with protection of intellectual property
Open Problems

Illustration copyright Ai825/Shutterstock
Open Problems

**Loose Software Supply Chains**

can we....

- measure and predict social cost of component use?
- determine when backward social contributions are needed?
- assess when a component upgrade is needed?
- lower the cost of quality and security upgrades?
Open Problems

**Tight Software Supply Chains**

can we....

- automatically apply IP filters to information exchange?
- provide white-box information without revealing secret sauce?
- cost-effectively manage multi-tiered supply chains?
- effectively handle arrangements of tight and loose supply chains?
Summary
Thanks to many post-docs, students and industrial collaborators over the years for their insights.

Thanks to NECSIS colleagues (particularly Jo Atlee, Marsha Chechik and Mark Lawford) for conversations.

Thanks to Sonatype for an analysis of the Central Repository.
Summary

Software Supply Chains
Software Supply Chains

Naïve
Tight
Loose
Open
Summary

Software Supply Chains

Naïve
Better, faster, cheaper

Loose Supply Chain
Reuse is not free

Tight Supply Chain
Controlled transparency

Open Problems
Technical and ecosystem
Summary

Software Supply Chains

“supply chain” conjures up thoughts of organized, managed flows

for software supply chains, the reality is different (chaotic? brittle?)

(re)use is not free

controlled transparency

Photo copyright Wierink/Shutterstock

@gail_murphy
"supply chain" conjures up thoughts of organized, managed flows

for software supply chains, the reality is different (chaotic? brittle?)

(re)use is not free
controlled transparency

@gail_murphy
Software Supply Chains

Gail Murphy
Univ. of British Columbia
Tasktop Technologies

@gail_murphy

Photo copyright Wierink/Shutterstock
With exception of pictures and icons