Tuesday, 27 September 2022

Analyzing the Relationship between Community and Design Smells in Open-Source Software Projects: An Empirical Study

Haris Mumtaz[1], Paramvir Singh[2], Kelly Blincoe[1]

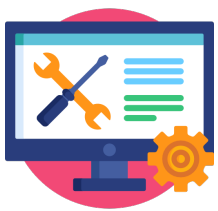
[1]Department of Electrical, Computer, and Software Engineering
[2]School of Computer Science

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

ENGINEERING

Motivation

- Conway's Law: "Organizations, which design systems, are constrained to produce designs which are copies of the communication structures of these organizations" [1].



Technical Design Structure

Communication Structure

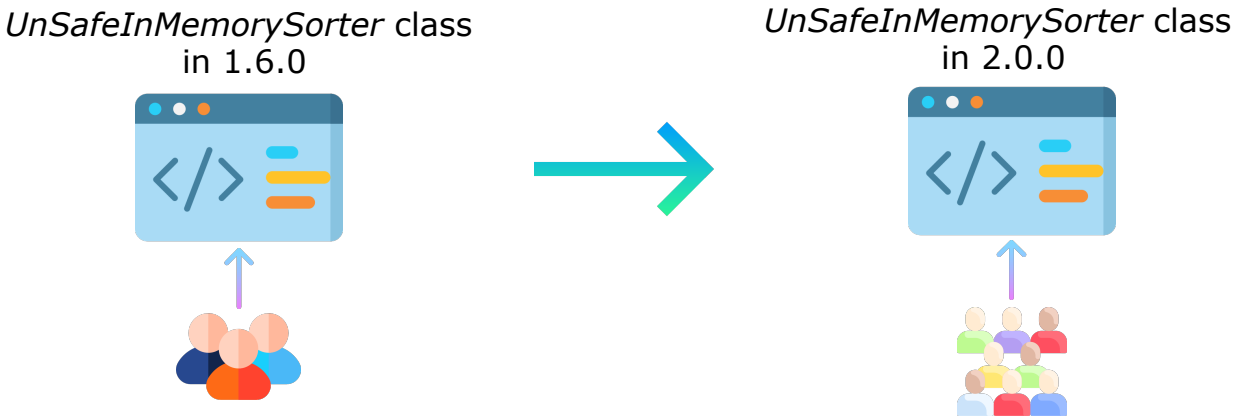
Design Issues

Communication Issues

[1] M. E. Conway. How do committees invent. *Datamation*, 14(4):28–31, 1968.

Motivation

- Large teams in social design → communication issues [2].
- Large components in technical design → Maintainability issues [3].



[2] G. Catolino, F. Palomba, D. A. Tamburri, A. Serebrenik, and F. Ferrucci. Refactoring community smells in the wild: The practitioner's field manual. *ICSE* 2020.
[3] G. Suryanarayana, G. Samarthyam, and T. Sharma. Refactoring for software design smells. *ACM SIGSOFT Software Engineering Notes*, 2015.

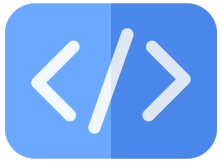
Research Question

RQ

Is there a relationship between community and design smells in software projects?

Data

- Extract design and community smells from 100 releases of 10 open-source Apache projects (10 recent releases each).



109—6234 classes

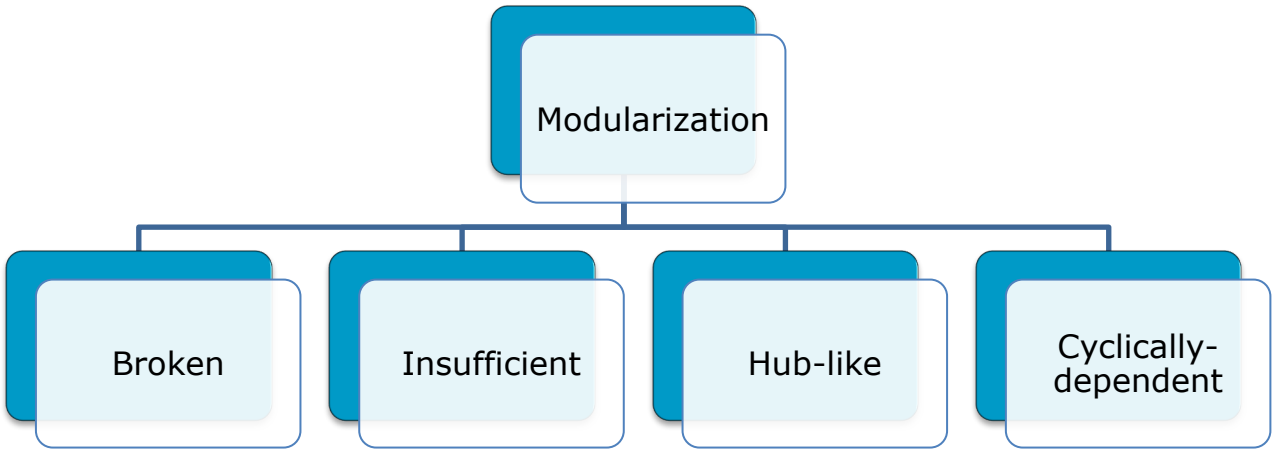
26—1,786 developers

3.2k—34.9k commits

10—22 years

Design Smells

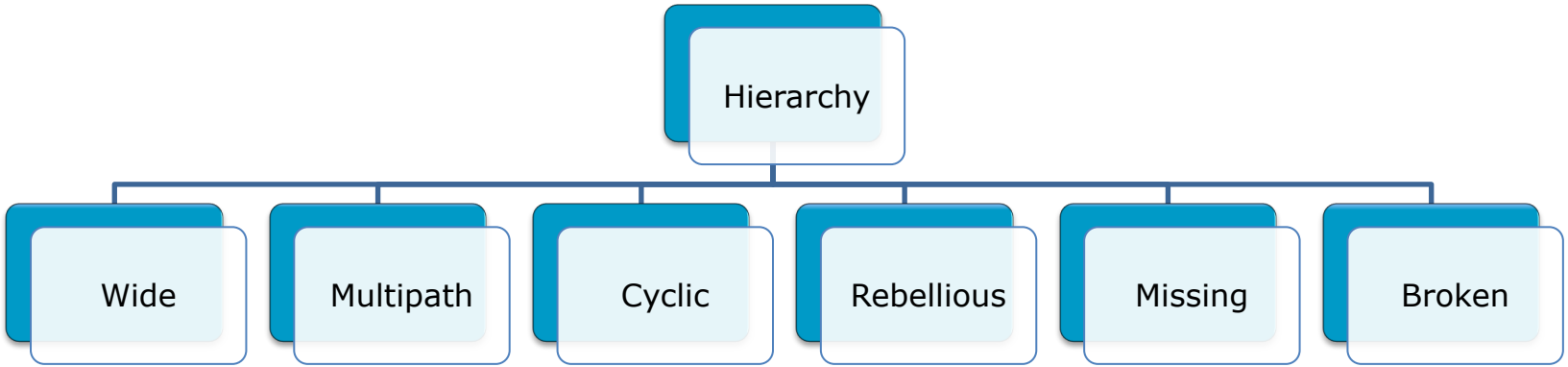
- Sub-optimal patterns in software design [3].
- Analyze common modularization smells in software design [3].



[3] G. Suryanarayana, G. Samarthyam, and T. Sharma. Refactoring for software design smells. *ACM SIGSOFT Software Engineering Notes*, 2015.

Design Smells

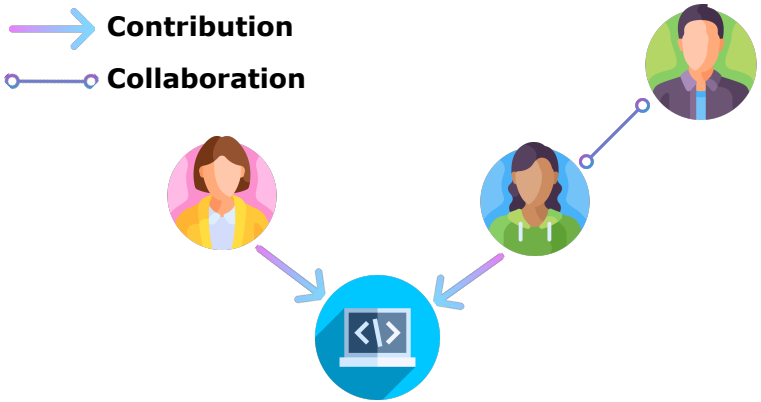
- Analyze common hierarchy smells in software design [3].



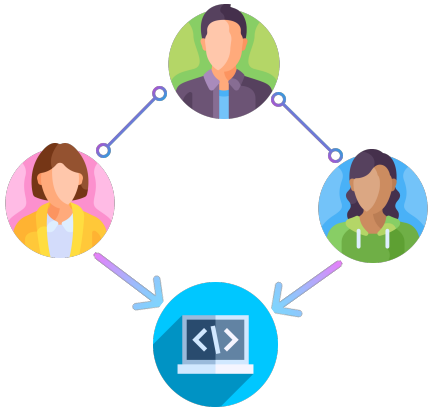
[3] G. Suryanarayana, G. Samarthyam, and T. Sharma. Refactoring for software design smells. *ACM SIGSOFT Software Engineering Notes*, 2015.

Community Smells

- Sub-optimal patterns in organizational and communication structure [4].
- Investigate frequently occurring community smells in projects [4].



Organizational Silo

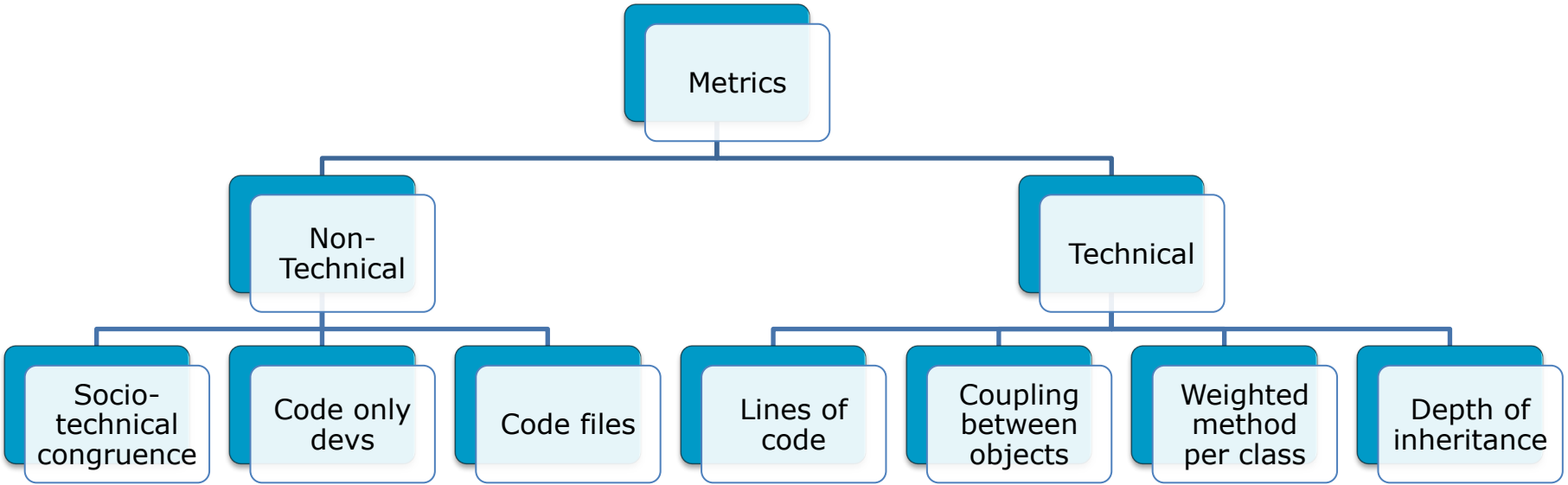


Missing Links

[4] D. A. Tamburri, P. Kruchten, P. Lago, and H. v. Vliet. Social debt in software engineering: Insights from industry. *JISA*, 2015.

Additional Metrics

- Non-technical (community) metrics [5].
- Technical (non-community) maintainability metrics.



[5] F. Palomba and D. A. Tamburri. Predicting the emergence of community smells using socio-technical metrics: A machine-learning approach. *JSS*, 2021.

Data Collection

- ***Designite*** tool to collect design smells.

https://www.designite-tools.com/

- ***Kaiaulu*** tool to compute community smells.

https://github.com/sailuh/kaiaulu

Methods

- Correlation Analysis (**Spearman's correlation**)
  - Relationship between community and design smells.
- Trend Analysis (**Mann-Kendall test**)
  - Similarity in the trends of community and design smells.
- Statistical Modeling (**Information Gain Analysis**)
  - Extent of the dependency of community smells on design smells.

Design smell: community smells + community metrics + technical metrics

Control factors

Results-Correlation

Modularization Smells

	Broken		Insufficient		Hub-like		Cyclically-dependent	
Project	**Organizational Silo**	**Missing Links**	**Organizational Silo**	**Missing Links**	**Organizational Silo**	**Missing Links**	**Organizational Silo**	**Missing Links**
Ant	0.78**	0.74*	0.81**	0.78**	0.74*	0.74*	0.94***	0.89***
Cassandra	0.14	0.71*	0.5	0.76*	0.28	0.75*	0.2	0.47
Jackrabbit	0.74*	0.64*	0.22	0.34	0.74*	0.64*	0.31	0.21
Jena	0.25	-0.02	0.29	-0.08	-0.25	-0.55	0.61*	0.41
JMeter	-0.04	-0.26	0.32	0.5	0.35	0.53	0.56	0.7*
Karaf	0.74*	0.71*	0.77**	0.81**	0	0	0.75*	0.76*
Spark	0.62*	0.7*	0.78**	0.78**	0.9***	0.78**	0.79**	0.77**
CloudStack	-0.44	-0.65*	0.44	0.63*	0.41	0.41	0.34	0.58
CXF	0.01	0.76*	0.22	0.73*	0.12	0.78**	-0.13	0.73*
Nutch	0.09	-0.09	-0.03	-0.14	0	0	0.07	-0.17

**p <0.05, **p <0.01, ***p <0.001*

Results-Correlation

Hierarchy Smells

	Wide		Multipath		Cyclic		Rebellious		Missing		Broken	
Project	**O_S**	**M_L**	**O_S**	**M_L**	**O_S**	**M_L**	**O_S**	**M_L**	**O_S**	**M_L**	**O_S**	**M_L**
Ant	0.85**	0.79**	0.86**	0.81**	0.77**	0.72*	0.92***	0.87**	0.59	0.65*	0.89***	0.83**
Cassandra	0.54	0.84**	0.53	0.29	0.56	0.78**	0.81**	0.94***	0.66*	0.91***	0.51	0.76*
Jackrabbit	0.74*	0.64*	0	0	0.74*	0.64*	0.74*	0.64*	0.25	0.32	-0.12	0.13
Jena	0.73*	0.4	-0.14	-0.43	0.31	0.17	-0.32	-0.07	0.47	0.47	0.22	-0.04
JMeter	0.45	0.64*	0.04	0.26	0.45	0.64*	-0.4	-0.35	0	0	0.44	0.67*
Karaf	0.6*	0.69*	0.65*	0.57	0.65*	0.65*	0	0	0	0	0.8**	0.79**
Spark	0.67*	0.67*	0.52	0.52	0.76*	0.61*	0.9***	0.78**	0.85**	0.77**	0.8**	0.79**
CloudStack	0.65*	0.65*	0.3	0.53	0	0	0	0.21	0	0	0.58	0.75*
CXF	-0.11	-0.82**	0.29	0.51	-0.01	0.18	0.31	0.69*	0.21	0.86**	0.22	0.73*
Nutch	0	0	0	0	0.04	-0.03	0	0	-0.1	-0.21	0.24	-0.02

***O_S** is Organizational Silo and **M_S** is Missing Links*

**p <0.05, **p <0.01, ***p <0.001*

Results-Trends

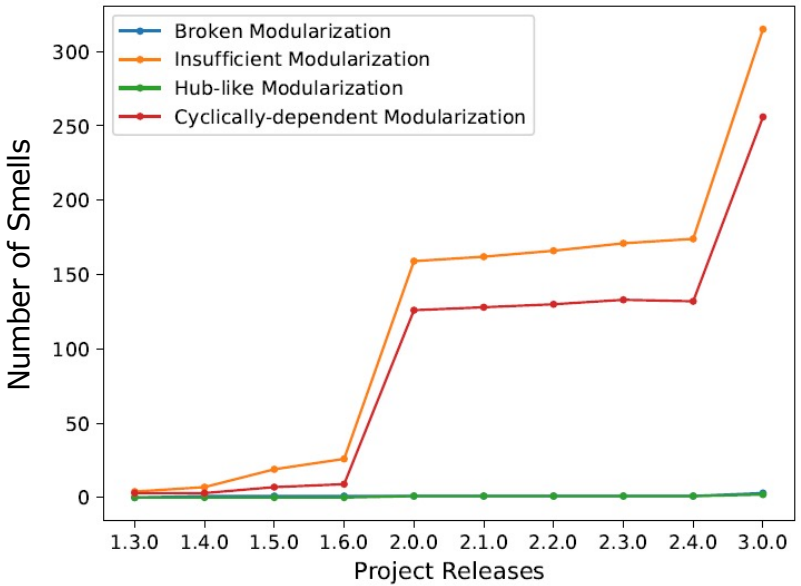
Smell	Ant	Cassandra	Jackrabbit	Jena	JMeter	Karaf	Spark	CloudStack	CXF	Nutch
Broken Modularization	29[↑]*	[–]	[–]	24[↑]*	[–]	33[↑]**	17[↑]*	[–]	35[↑]**	26[↑]*
Insufficient Modularization	45[↑]***	42[↑]***	29[↑]**	28[↑]*	33[↑]**	45[↑]***	45[↑]***	43[↑]***	45[↑]***	[–]
Hub-like Modularization	32[↑]**	27[↑]*	[–]	[–]	[–]	[–]	29[↑]**	[–]	40[↑]***	[–]
Cyclically-dependent Modularization	38[↑]***	29[↑]**	[–]	[–]	[–]	41[↑]***	42[↑]***	41[↑]***	34[↑]**	29[↑]*
Wide Hierarchy	35[↑]**	29[↑]**	[–]	[–]	33[↑]**	37[↑]***	23[↑]*	[–]	22[↓]*	[–]
Multipath Hierarchy	33[↑]**	[–]	[–]	24[↑]*	[–]	24[↑]*	[–]	35[↑]**	31[↑]**	[–]
Cyclic Hierarchy	39[↑]***	41[↑]***	[–]	[–]	33[↑]**	21[↑]*	27[↑]*	[–]	27[↑]*	[–]
Rebellious Hierarchy	28[↑]*	[–]	[–]	-24[↓]*	[–]	[–]	29[↑]**	24[↑]*	39[↑]***	[–]
Missing Hierarchy	24[↑]*	25[↑]*	[–]	[–]	[–]	[–]	32[↑]**	[–]	24[↑]*	[–]
Broken Hierarchy	33[↑]**	44[↑]***	35[↑]**	[–]	37[↑]***	41[↑]***	43[↑]***	37[↑]**	45[↑]***	27[↑]**
Organizational Silo	30[↑]**	[–]	[–]	[–]	[–]	28[↑]*	27[↑]*	[–]	[–]	[–]
Missing Links	28[↑]*	25[↑]*	[–]	[–]	[–]	31[↑]**	31[↑]**	[–]	23[↑]*	[–]

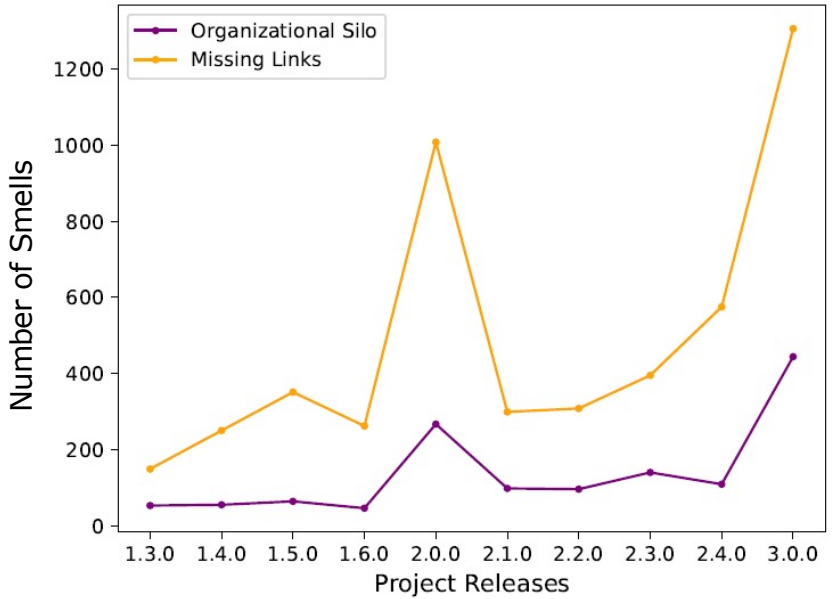
[↑] = *Increasing trend;* [↓] = *Decreasing trend;* [–] = *No trend*

**p <0.05, **p <0.01, ***p <0.001*

Results-Trends

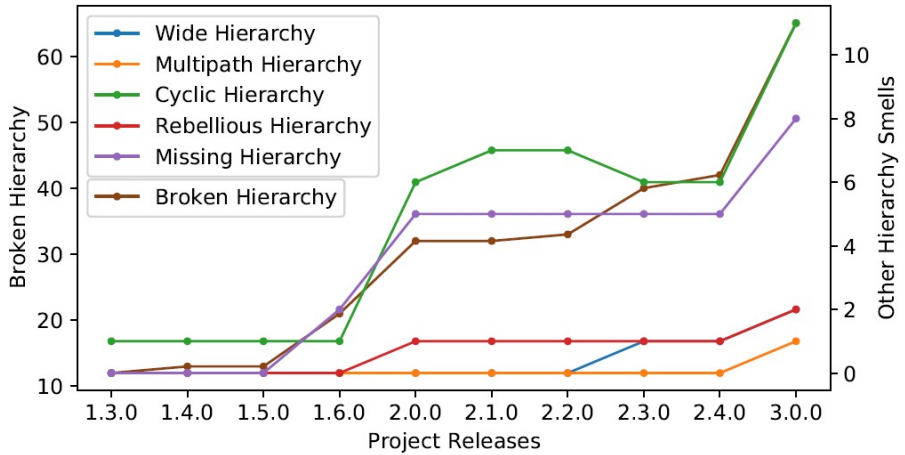
Trends of Apache Spark project





Results-Trends

Trends of Apache Spark project



Results-Gain Analysis

Design smell: community smells + community metrics + technical metrics

Control factors

- Missing Links is the most dependent community smell on design smells.

Answer to RQ

RQ

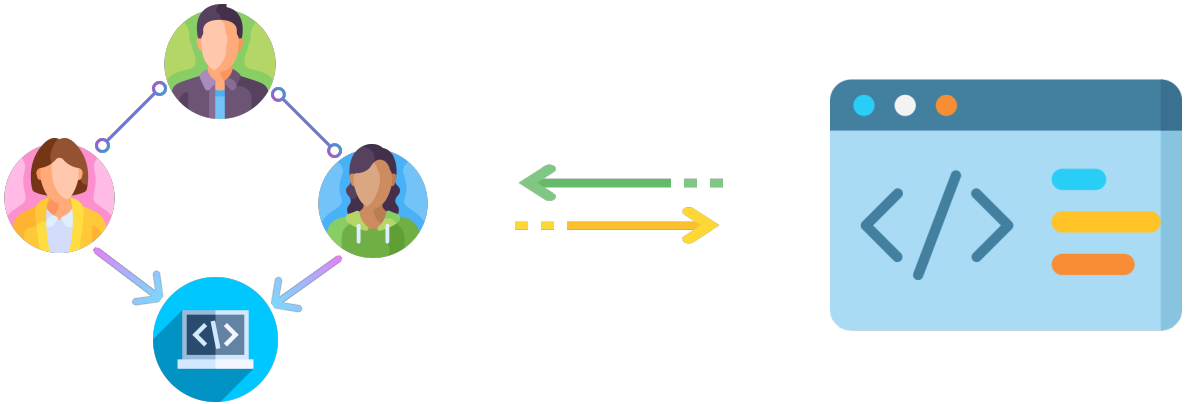
Is there a relationship between community and design smells in software projects?

Answer

- The Missing Links smell has demonstrated more significant correlations and trend similarities with the design smells in the analyzed projects.

Main Takeaways

- Missing Links has demonstrated relationships with the design structures that are either not implemented at all (i.e., missing) or implemented when not required (i.e., broken).

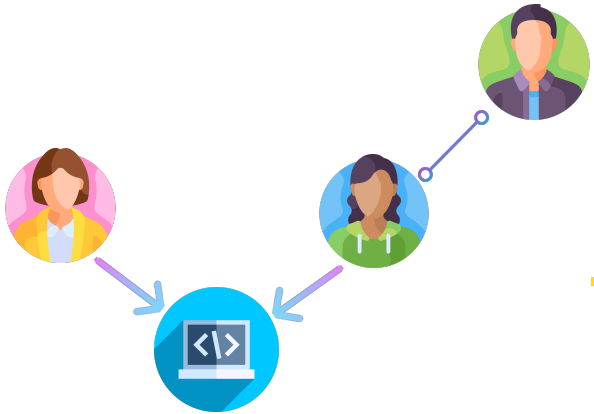


Non-coordinating devs

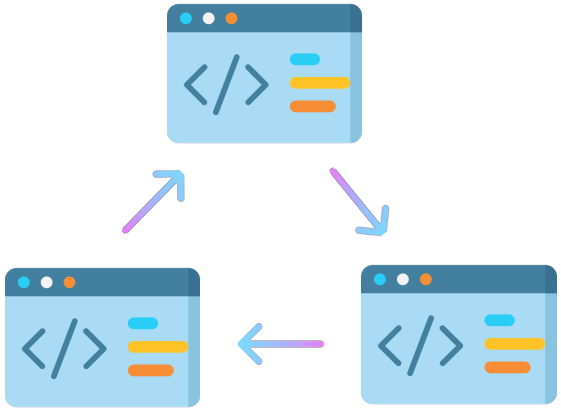
Insufficient Modularization

Main Takeaways

- Design smells that focus on cycles between the classes (e.g., Cyclically-dependent Modularization and Cyclic Hierarchy) have a relationship with Organizational Silo.



Disjoint sub-communities

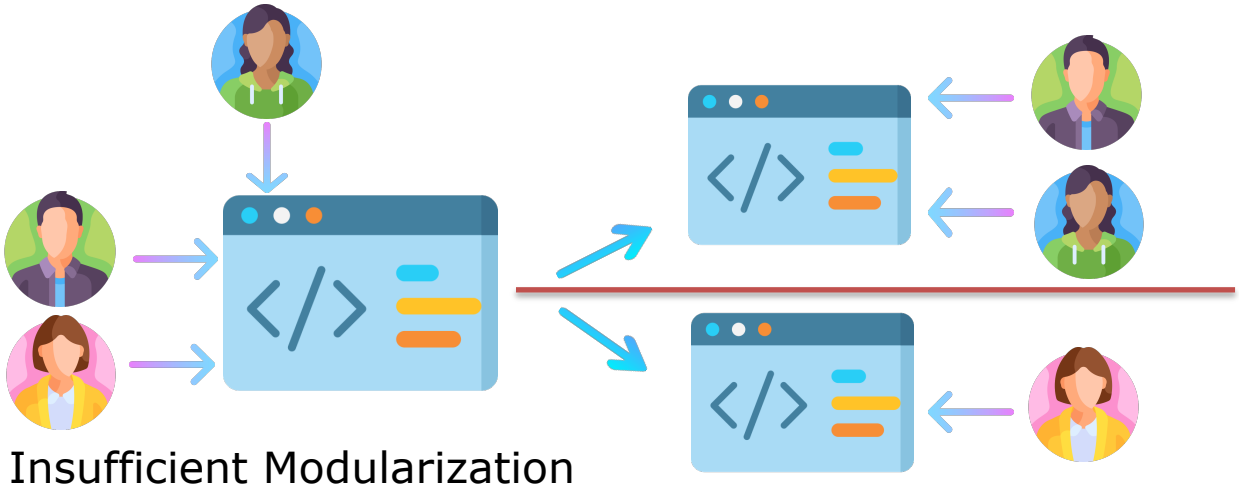


Cycles between classes

Main Takeaways

- Community-aware development is also important alongside technical development.

- Community and technical issues should be refactored together.



Insufficient Modularization

Paper QR

Conclusion

- Conway’s law suggests correlation between social and technical aspects.
- Previous studies had analyzed the relationship at code and architecture levels.
- We investigate such relationship in software design using smells.
- Our results show relationship between design and community smells.
- We propose collective refactoring of social and technical (design) aspects.

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

ENGINEERING

Thank you!

Haris Mumtaz
hmum126@aucklanduni.ac.nz
HarisMumtaz_

Paramvir Singh
p.singh@auckland.ac.nz
ParamvirSingh_

Kelly Blincoe
k.blincoe@auckland.ac.nz
KellyBlincoe

Paper QR