Modularity Anomaly Types – Part II

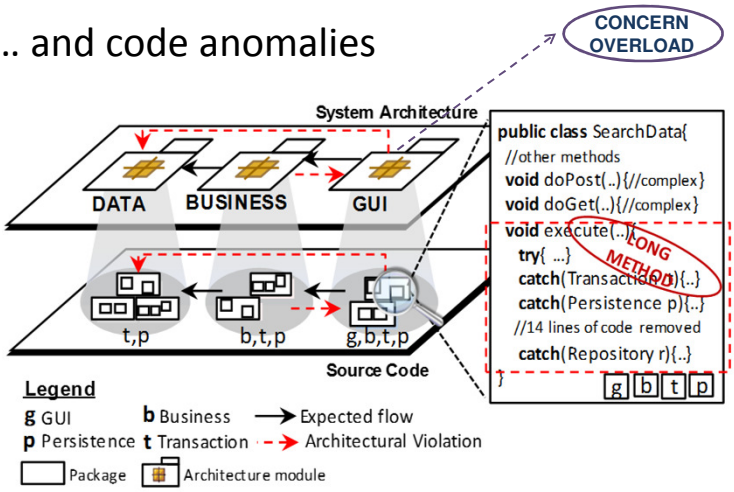
Alessandro Garcia

LES | DI |PUC-Rio - Brazil

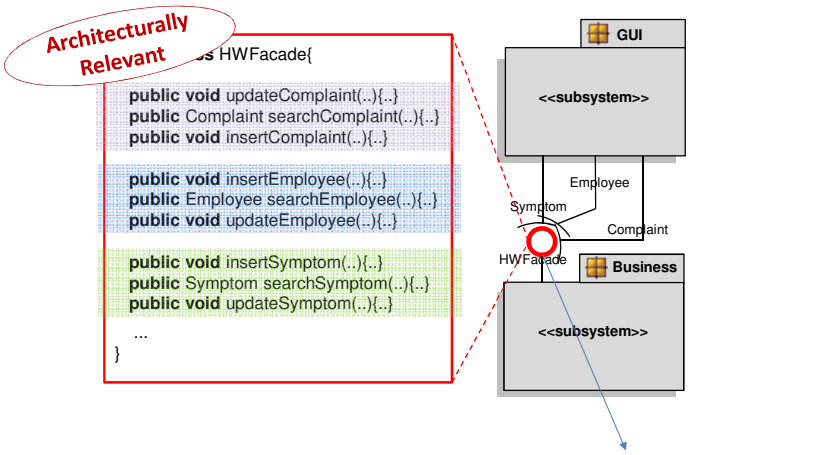
OPUS Group

Relation between architectural problems…

- … and code anomalies

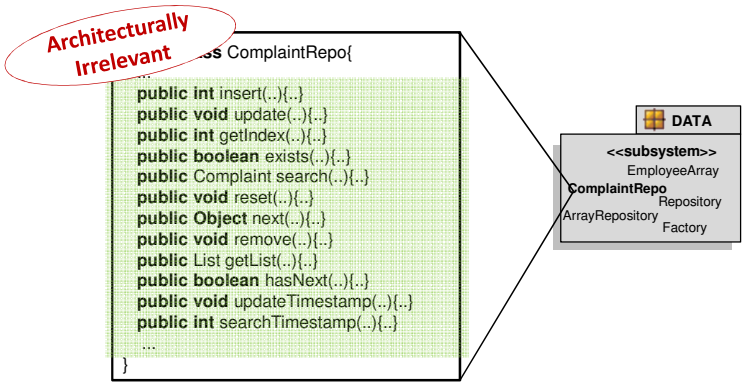


Relevance of Code Anomalies



Architectural Interface Bloat

Relevance of Code Anomalies



Three Questions

1 Are anomalous code elements and architecture problems related?

2 If so, which characteristics of the code anomaly are relevant for the architecture design?

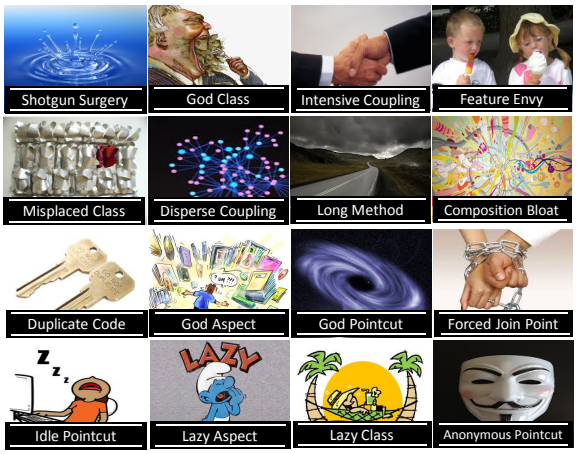
3 To what extent the applied refactorings actually addressed architecturally-relevant code anomalies?

5

Target systems

MIDAS	Mobile Media
PDP	Health Watcher
	Aspectual Watcher
	Aspectual Mobile

Code anomalies



Architectural anomalies

Target Systems

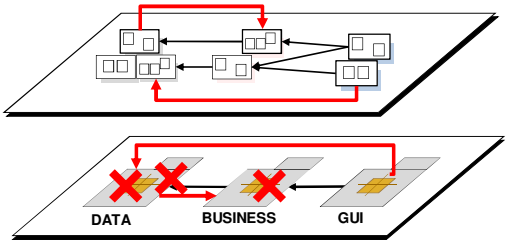
MIDAS	MM	HW	PDP
C++	Java	Java	C#
76 KLOC	54 KLOC	49 KLOC	22 KLOC
111 anomalies	170 anomalies	252 anomalies	175 anomalies

- 6 different systems
- 40 revisions
- Architecture information available

7

Analysis

- Recovering **Actual Architecture**
- Identifying Architecture Anomalies
- Detecting Code Anomalies
- Analyzing the Impact of Code Anomalies

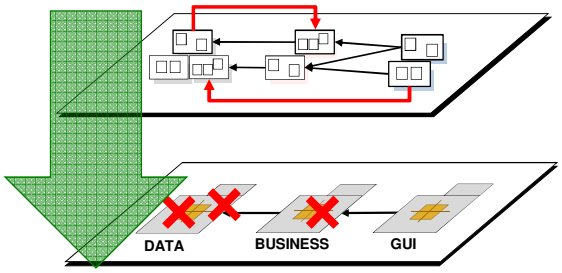


Q1 Are anomalous code elements and architecture problems related?

8

Analyzing the Impact of Code Anomalies

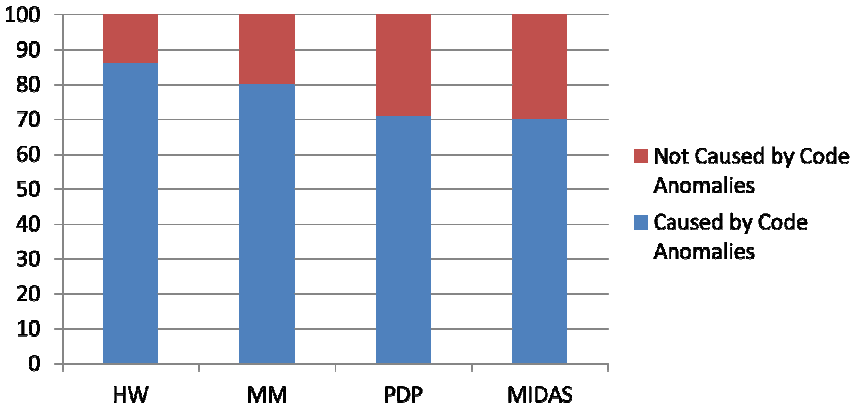
- Downstream Analysis
  - Which architecture problems are related to code anomalies



9

Analyzing the Impact of Code Anomalies

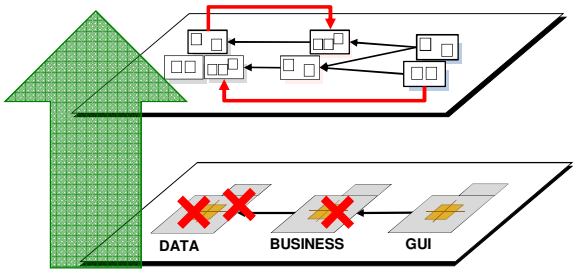
- Downstream Analysis



10

Analyzing the Impact of Code Anomalies

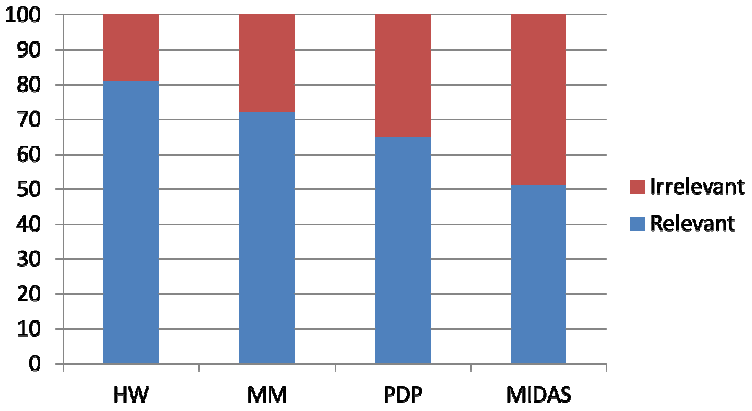
- Upstream Analysis
  - Which code anomalies are related to architecture problems



11

Analyzing the Impact of Code Anomalies

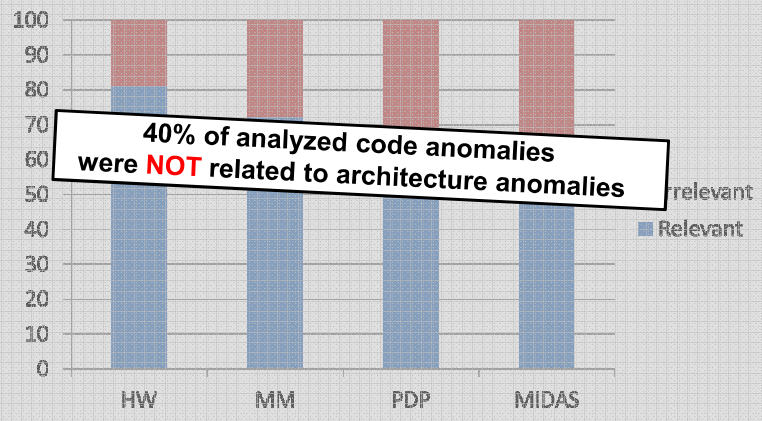
- Upstream Analysis



12

Analyzing the Impact of Code Anomalies

- Upstream Analysis



40% of analyzed code anomalies were NOT related to architecture anomalies

13

Three Questions

1 Are anomalous code elements and architecture problems related?

2 If so, which characteristics of the code anomaly are relevant for the architecture design?

3 To what extent the applied refactorings actually addressed architecturally-relevant code anomalies?

14

Identifying Relevant Code Anomalies

- The following characteristics were analyzed:
  - **Code anomaly type**
  - Change density
  - Error density
  - Anomaly density
  - And others...

Q2 which characteristics of code anomalies are architecturally-relevant?

15

Type of Code Anomaly

Kinds of Code Anomalies	AW	HW	AM	MM
Divergent Change	5	7	5	5
Feature	[illegible]	[illegible]	[illegible]	2
God Cl	[illegible]	[illegible]	[illegible]	5
Inappropriate Intimacy	6	[illegible]	[illegible]	8
Long Method	7	8	4	6
Long Parameter List	7	2	5	3
Misplaced Class	1	1	2	3
Shotgun Surgery	3	2	4	5
Small Class	4	7	3	6

None of the code anomalies was the best indicator across all analyzed systems

16

Identifying Relevant Code Anomalies

- The following characteristics were analyzed:
  - Change density
  - Error density
  - Anomaly density

Q2 which characteristics of code anomalies are architecturally-relevant?

Studying prioritization models

- Which other characteristics could be explored for detecting architecturally-relevant code anomalies ?
  - Change density
  - Error density
  - Anomaly density

Q2 which characteristics of code anomalies are architecturally-relevant?

R. Arcoverde et al – **RSSE/ICSE 2012**: Automatically Detecting Architecturally-Relevant Code Anomalies

R. Arcoverde et al – **SBES 2013**: ***Prioritization of Code Anomalies Based on Architecture Sensitiveness.*** SBES'13) Brasília, Brazil, September 2013.

Prioritization heuristics

Change density

System	# of Ranked CE	Arch. Relevant	%
HW	14	10	71%
MM	10	7	70%
PDP	**10**	**10**	**100%**

Error density

System	# of Ranked CE	Arch. Relevant	%
HW	**14**	**12**	**85%**
MM	10	8	70%
PDP	10	8	70%

Anomaly density

System	# of Ranked CE	Arch. Relevant	%
HW	10	7	60%
MM	10	9	70%
PDP	10	8	70%
MIDAS	**10**	**6**	**90%**

19

Priorization Factors

.. 1. **There is no 'universal' prioritization model**
2. **Prioritization models: satisfactory results too late**

May 16 20

Prioritization heuristics

Change density

System	# of Ranked CE	Arch. Relevant	%
HW	14	10	
MM	10	7	
PDP	**10**	**10**	**100%**

Version 12

Error density

System	# of Ranked CE	Arch. Relevant	%
HW	**14**	**12**	**85%**
MM	10	8	
PDP	10	8	

Version 9

Anomaly density

System	# of Ranked CE	Arch. Relevant	%
HW	10	7	
MM	10	9	
PDP	10	8	
MIDAS	**10**	**6**	**90%**

Version 10

21

Earliness of Anomaly

- Early anomaly: appears in the 1st version of each system

18%

Of all architecturally-relevant code anomalies were identified as **early anomalies**

22

Earliness of Anomaly

- Early anomaly: appears in the 1st version of each system

18%

Of all architecturally-relevant code anomalies were identified as **early anomalies**

and were related to more than

37%

of all architecture problems

However, we observed that...

- ... architecturally-relevant code anomalies tend to "flock together" even in the 1st version of each system:
  - anomaly agglomerations within a class or syntatically-related classes can be used as good indicator
- Continuous detection of emerging anomaly agglomerations?

Three Questions

1 Are anomalous code elements and architecture problems related?

2 If so, which characteristics of the code anomaly are relevant for the architecture design?

3 To what extent the applied refactorings actually addressed architecturally-relevant code anomalies?

25

Refactoring of Relevant Anomalies

- 658 refactorings
  - 33% high-level
    - Move member (16%)
    - Extract class or superclass (12%)
  - 67% low-level
    - Rename (32%)
    - Extract local variable (16%)
- **37%** of all architecture-relevant anomalies were refactored
- Isolated versions concentrated most of the refactoring efforts

Q3 architecturally-relevant refactorings?

26

Concluding Remarks

Most architecture problems are related to anomalous code elements

Concluding Remarks

Most architecture problems are related to anomalous code elements

Anomaly agglomerations are more frequently related to architecture problems

Concluding Remarks

Most architecture problems are related to anomalous code elements

Anomaly agglomerations are more frequently related to architecture problems

Architecturally-relevant anomalies are not frequently refactored

29

Concluding Remarks

Most architecture problems are related to anomalous code elements

Anomaly agglomerations are more frequently related to architecture problems

Architecturally-relevant anomalies are not frequently refactored

Detecting early code anomalies is a needed asset for assisting developers when prioritizing refactoring efforts

30