university of groningen

Assessing Change Proneness at the Architecture Level: An Empirical Validation

Elvira-Maria Arvanitou, Apostolos Ampatzoglou,
Konstantinos Tzouvalidis, Alexander Chatzigeorgiou,
Paris Avgeriou, Ignatios Deligiannis

university of groningen

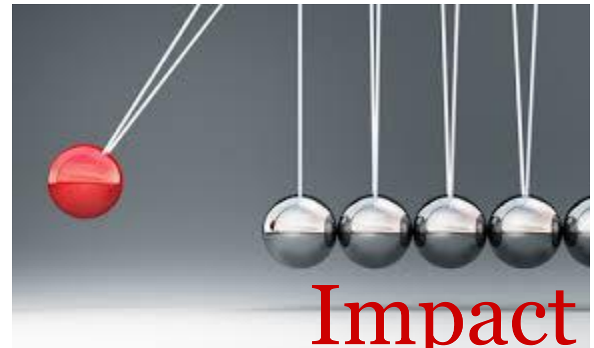
HELLENIC REPUBLIC
UNIVERSITY OF MACEDONIA
ECONOMIC & SOCIAL SCIENCES

Outline

- › Context
- › Module Change Proneness Measure
- › Empirical Validation
- › Results
- › Discussion

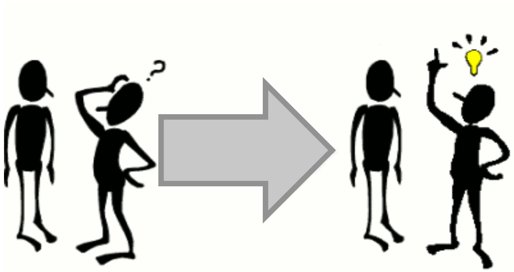
university of groningen

Change Impact Analysis



Benefits of "Predicting" Change Proneness

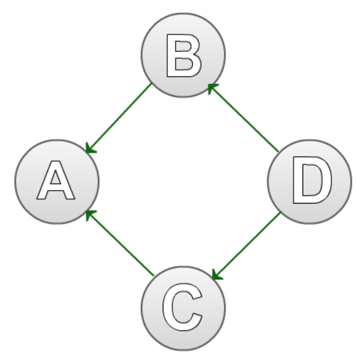
Before... CHANGE After...



Program Comprehension

test cases

Effort Estimation



Component Dependency Identification

Change Impact Analysis and Architecture

- Not any metrics at the Architecture Level

- Design and Code Level: Coupling & Size

  The degree to which one class is connected to other classes of the system

- Alternatives?

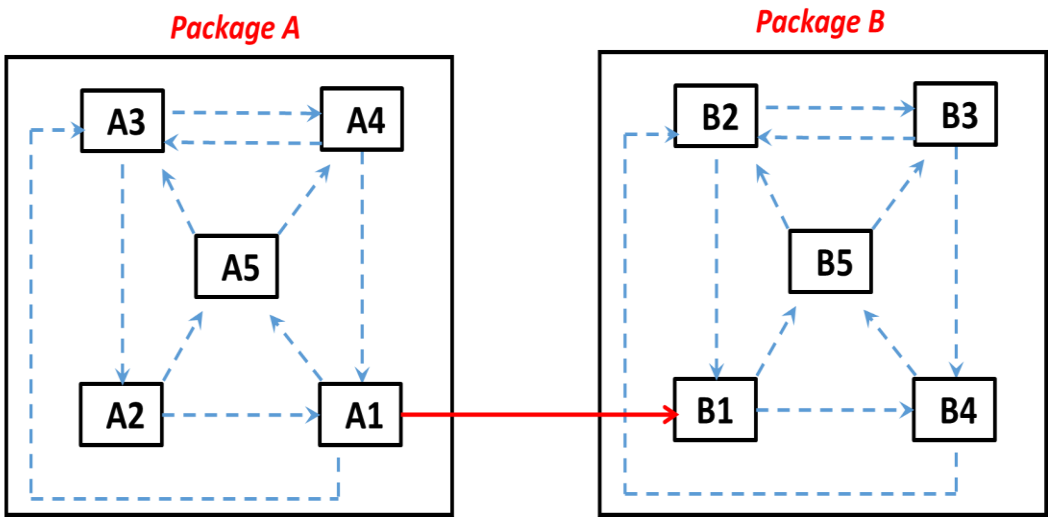
  Propose a new metric

  Aggregate an existing one from design level

Aggregation of Design Metrics

› How can the strength be aggregated?

- Intra-module
- Inter-module

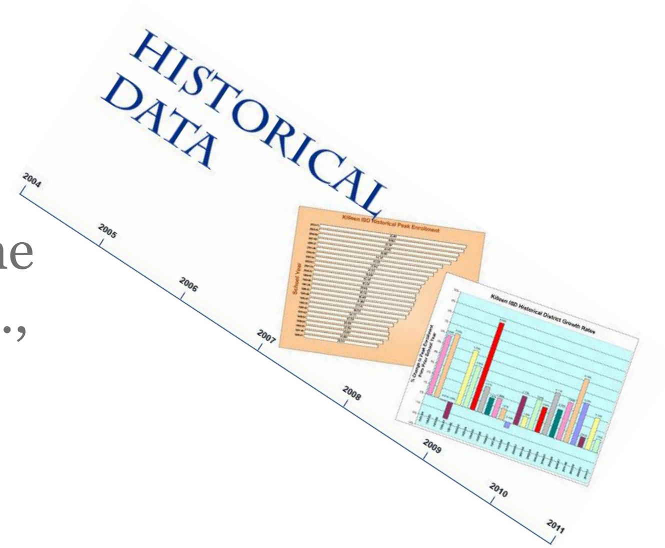


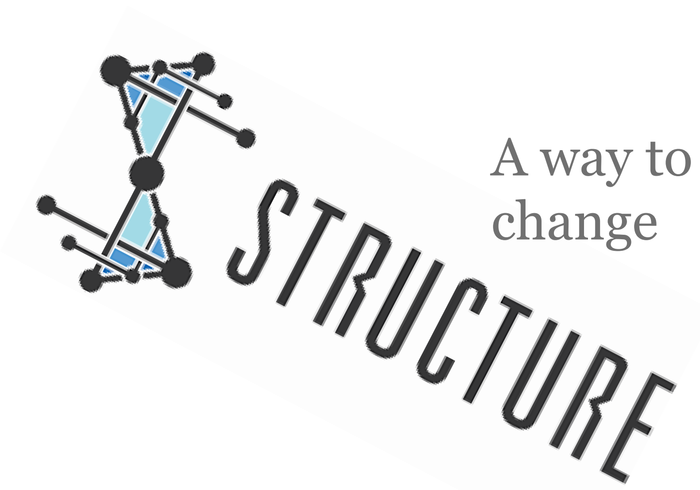
university of groningen

Change Proneness Components

› What is needed?

The source of the ripple effect (i.e., change)



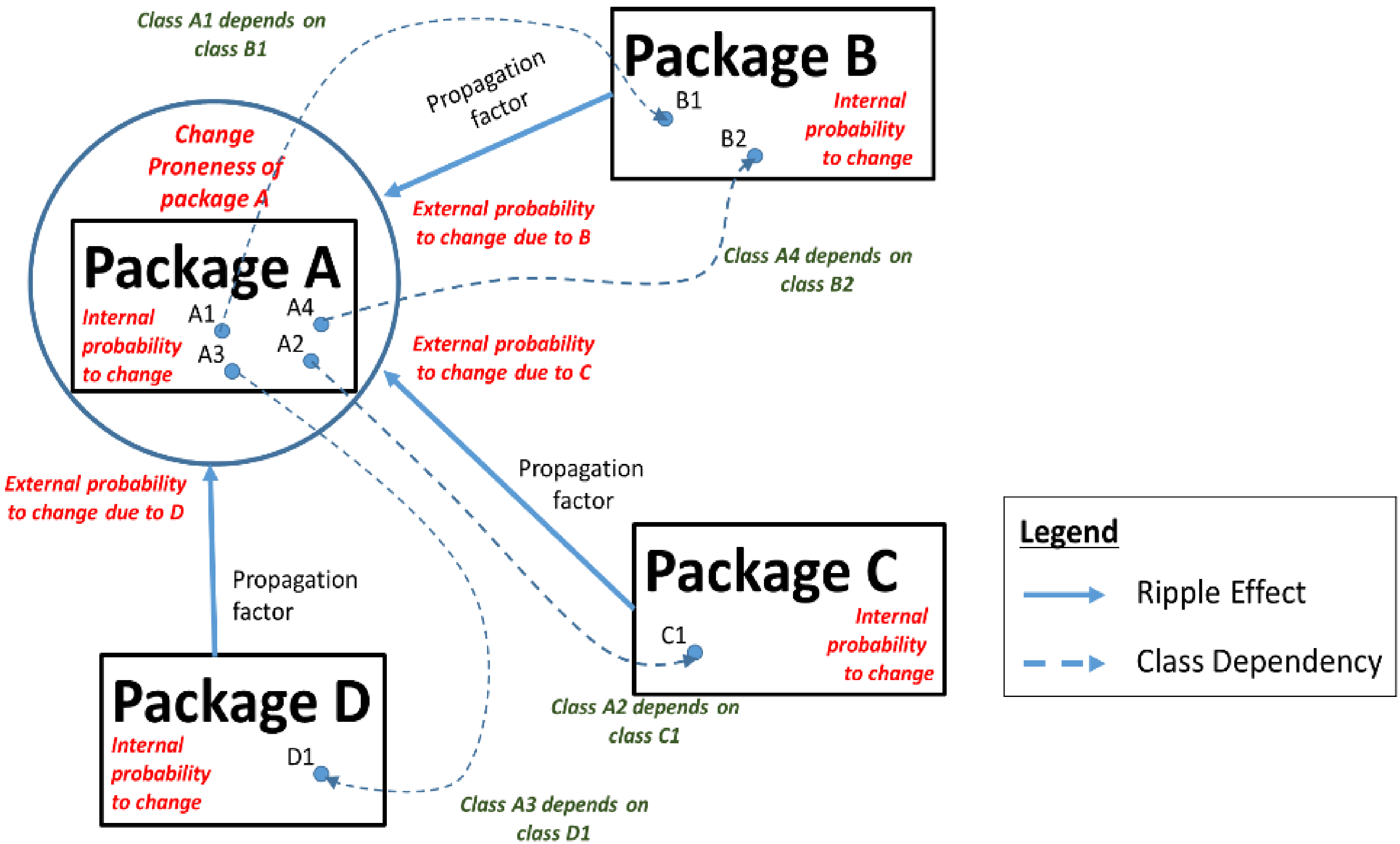


A way to transfer the change

Outline

- Context
- Module Change Proneness Measure
- Empirical Validation
- Results
- Discussion

Change Scenario



Putting the pieces together

› Joint Probability of All Events

Even one change from any dependency is enough for the module to change

MCPM(A) = *Joint Probability* {P(A), P(A:external$_B$), P(A:external$_C$), P(A:external$_D$)}

P(A:external$_B$) = P(A|B) • P(B)

*P(A|B) is the **propagation factor** between module B and A (i.e., the probability that a change made in B is emitted to A).*

*P(B) refers to the **internal probability** of changing module B.*

Putting the pieces together: Propagation Factor

$$REM(\mathrm{B}\rightarrow\mathrm{A}) = \sum_{i=0}^{i<dependencies(\mathrm{A}\rightarrow\mathrm{B})} \frac{NPrA(Bi) + NOP(Bi) + NDMC(\mathrm{Ai}\rightarrow\mathrm{Bi})}{NOM(Bi) + NOA(Bi)}$$

NDMC: Number of direct method calls

NOM: Number of methods

NOA: Number of attributes

NPrA: Number of protected attributes (only for inheritance)

NOP: Number of polymorphic methods (only for inheritance)

Outline

- Context
- Module Change Proneness Measure
- Empirical Validation
- Results
- Discussion

Case Study Design

Goal of this study:

Compare the validity of MCPM to other metrics

Three package-level coupling metrics

- Afferent Coupling (Ca)
- Efferent Coupling (Ce)
- Instability (I)

university of groningen

Case & Data Collection

APACHE WICKET

Wro4j

guava-libraries
Guava: Google Core Libraries for Java 1.6+

Apache Commons™
http://commons.apache.org/

JODA-TIME API

- Demographics
  - Project
  - Version
  - Package name
- Assessors
  - MCPM
  - Ca
  - Ce
  - I
- Actual changes

160 Java packages ~ Units of analysis
~30 packages per project

university of groningen

Data Analysis

<table>
<tr><th>Criterion</th><th>Test</th><th>Variables</th></tr>
<tr><td>Correlation</td><td>Pearson correlation</td><td>Independent: Assessors
Dependent: Actual changes
(last version of the projects)</td></tr>
<tr><td>Consistency</td><td>Spearman correlation</td><td>Independent: Assessors
Dependent: Actual changes
(last version of the projects)</td></tr>
<tr><td>Tracking</td><td>Spearman correlation</td><td>Independent: Assessors
Dependent: Actual changes
(across all versions)</td></tr>
<tr><td>Predictability</td><td>Linear Regression</td><td>Independent: Assessors
Dependent: Actual changes
(last version of the projects)</td></tr>
<tr><td>Discriminative Power</td><td>Kruskal Wallis Test</td><td>Testing: Assessors
Grouping: Actual changes
(last version of the projects)</td></tr>
<tr><td>Reliability</td><td colspan="2">all the aforementioned tests
(separately for each project –across all versions)</td></tr>
</table>

Outline

- Context
- Module Change Proneness Measure
- Empirical Validation
- Results
- Discussion

university of groningen

Results

TABLE V. CORRELATION ANALYSIS

Project	MCPM	Ca	Ce	I
wro4j	***.348***	.288	.346	.102
Guava	***.487***	.407	.272	.109
commons-lang	***.805***	.156	*.754*	-.166
joda-time	**.205**	.090	-.187	-.409
Wicket	*.476*	*.412*	***.791***	-.016
% sig.	**80%**	40%	40%	0%

TABLE VI. CONSISTENCY ANALYSIS

Project	MCPM	Ca	Ce	I
wro4j	***.398***	.052	*.379*	.110
Guava	.437	***.484***	.409	.197
commons-lang	**.306**	.110	.242	.013
joda-time	***.378***	.321	-.161	-.400
Wicket	*.419*	-.069	***.623***	.059
% sig.	**60%**	20%	40%	0%

TABLE VII. TRACKING ANALYSIS

Project	MCPM	Ca	Ce	I
wro4j	***.390***	.050	*.375*	.105
Guava	.400	***.450***	.301	.150
commons-lang	**.301**	.106	.240	.010
joda-time	**.370**	.317	-.155	-.395
Wicket	*.410*	-.064	***.618***	.052
% sig.	**40%**	20%	**40%**	0%

university of groningen

Results

TABLE V. CORRELATION ANALYSIS

Project	MCPM	Ca	Ce	I
wro4j	***.348***	.288	.346	.102
Guava	***.487***	.407	.272	.109
commons-lang	***.805***	.156	*.754*	-.166
joda-time	**.205**	.090	-.187	-.409
Wicket	*.476*	*.412*	***.791***	-.016
% sig.	**80%**	40%	40%	0%

TABLE VI. CONSISTENCY ANALYSIS

Project	MCPM	Ca	Ce	I
wro4j	***.398***	.052	*.379*	.110
Guava	.437	***.484***	.409	.197
commons-lang	**.306**	.110	.242	.013
joda-time	***.378***	.321	-.161	-.400
Wicket	*.419*	-.069	***.623***	.059
% sig.	**60%**	20%	40%	0%

TABLE VII. TRACKING ANALYSIS

Project	MCPM	Ca	Ce	I
wro4j	***.390***	.050	*.375*	.105
Guava	.400	***.450***	.301	.150
commons-lang	**.301**	.106	.240	.010
joda-time	**.370**	.317	-.155	-.395
Wicket	*.410*	-.064	***.618***	.052
% sig.	**40%**	20%	**40%**	0%

Results

TABLE III. PREDICTIVE POWER

Project	MCPM	Ca	Ce	I
wro4j	**.030**	.031	**.030**	.032
Guava	***.104***	*.110*	.115	.119
commons-lang	***.067***	.112	*.075*	.125
joda-time	**.319**	.324	.320	.320
Wicket	*.012*	*.012*	***.008***	.013
% sig.	**60%**	40%	40%	0%

TABLE IV. DISCRIMINATIVE POWER

Project	MCPM	Ca	Ce	I
wro4j	***.008***	.587	*.049*	.613
Guava	**.059**	.277	.139	.835
commons-lang	.999	**.727**	.999	.889
joda-time	.381	**.190**	.571	.381
wicket	***.000***	.734	***.000***	.862
% sig.	**40%**	0%	**40%**	0%

university of groningen

Results

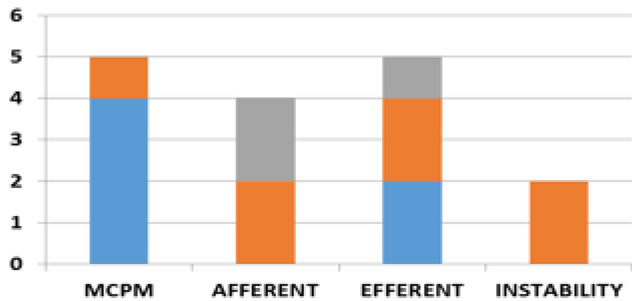
TABLE III. PREDICTIVE POWER

Project	MCPM	Ca	Ce	I
wro4j	**.030**	.031	**.030**	.032
Guava	***.104***	*.110*	.115	.119
commons-lang	***.067***	.112	*.075*	.125
joda-time	**.319**	.324	.320	.320
Wicket	*.012*	*.012*	***.008***	.013
% sig.	**60%**	40%	40%	0%

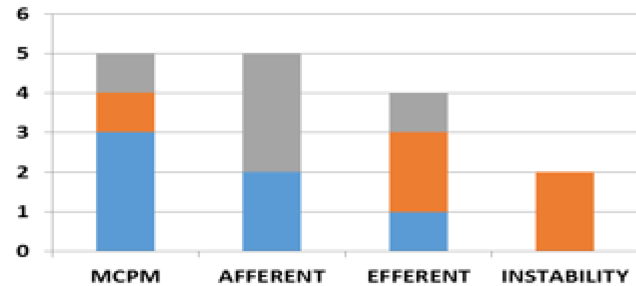
TABLE IV. DISCRIMINATIVE POWER

Project	MCPM	Ca	Ce	I
wro4j	***.008***	.587	*.049*	.613
Guava	**.059**	.277	.139	.835
commons-lang	.999	**.727**	.999	.889
joda-time	.381	**.190**	.571	.381
wicket	***.000***	.734	***.000***	.862
% sig.	**40%**	0%	**40%**	0%

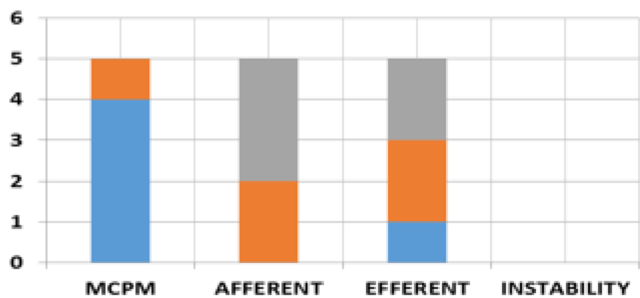
Results



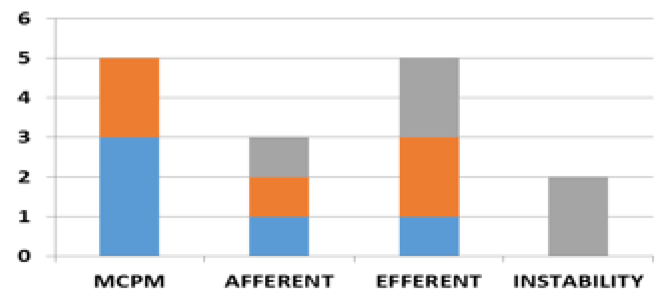
(a) Predictability Analysis



(b) Discriminative Power Analysis

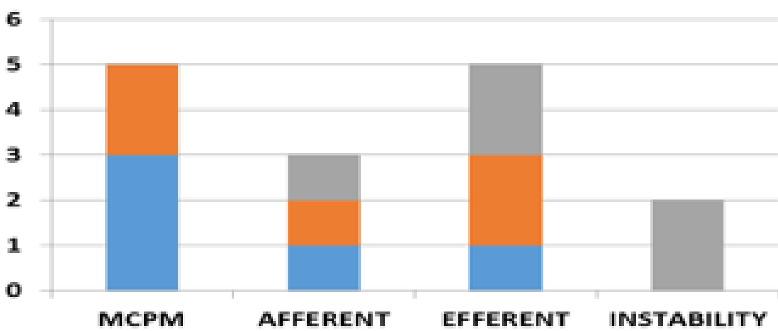


(c) Correlation Analysis



(d) Consistency Analysis

Results



(e) Tracking Analysis

TABLE VIII. Reliability Analysis

Criterion	MCPM	Ca	Ce	I
Corelation	**14**	7	9	0
Consistency	**13**	6	9	2
Tracking	**13**	6	9	2
Predictive Power	**14**	6	11	4
Discriminative Power	**12**	9	8	4
Total	**66**	34	46	12

university of groningen

Outline

- Context
- Module Change Proneness Measure
- Empirical Validation
- Results
- Discussion

Assessing Power

Combining both aspects

- Internal probability
- Ripple effects
- Strength of Dependencies

Tracking vs. Consistency

Efferent vs. Afferent Coupling

Implications

Practitioners

Use in quality monitoring processes

Test Case Prioritization

Tool Support

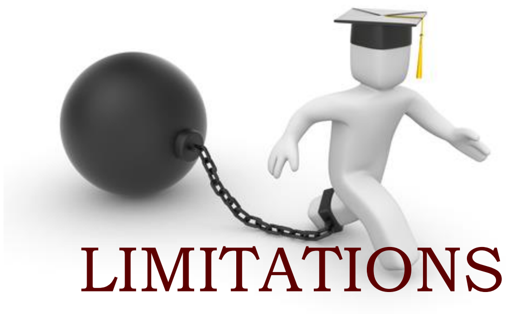
Researchers

Tailoring to a higher granularity

Explore usefulness for practitioners

Replicate with larger history

Threats to Validity



Construct Validity:

- Tool Accuracy
- Method Accuracy

Lack of Generalization to:

- Programming Language / Paradigm

Reliability:

- No research bias
- Public repositories

Questions?

Thank you for your attention!