Pontifícia Universidade Católica do Rio de Janeiro

Isela Macía Bertrán

# On the Detection of Architecturally-Relevant Code Anomalies in Software Systems

**Tese de Doutorado**

Thesis presented to the Programa de Pós-Graduação em Informática of the Departamento de Informática, PUC-Rio as partial fulfillment of the requirements for the degree of Doutor em Informática.

Advisor: Prof. Arndt von Staa
Co-advisor: Prof. Alessandro Fabricio Garcia

Rio de Janeiro
March, 2013

Pontifícia Universidade Católica do Rio de Janeiro

**Isela Macía Bertrán**

# On the Detection of Architecturally-Relevant Code Anomalies in Software Systems

Thesis presented to the Programa de Pós-Graduação em Informática of the Departamento de Informática, PUC-Rio as partial fulfillment of the requirements for the degree of Doutor em Informática.

**Prof. Arndt von Staa**
Advisor
Departamento de Informática – PUC-Rio

**Prof. Alessandro Fabricio Garcia**
Co-advisor
Departamento de Informática – PUC-Rio

**Profa. Yuanfang Cai**
Drexel University

**Prof. Paulo Henrique Monteiro Borba**
Universidade Federal de Pernambuco (UFPE)

**Prof. Carlos José Pereira de Lucena**
Departamento de Informática – PUC-Rio

**Prof. Renato Fontoura de Gusmão Cerqueira**
Departamento de Informática – PUC-Rio

**Prof. José Eugenio Leal**
Coordinator of the Centro Técnico Científico da PUC-Rio

Rio de Janeiro, March 20th, 2013

**Isela Macía Bertrán**

She completed her undergraduate studies in Computer Science at the University of Havana, Cuba in 2006. She received her Master degree in Informatics from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio) in 2009.



Bibliographic data

Macía, Isela

On the Detection of Architecturally-Relevant Code Anomalies in Software Systems / Isela Macía Bertrán; orientador: Arndt von Staa; co–orientador: Alessandro Fabricio Garcia. — Rio de Janeiro : PUC–Rio, Departamento de Informática, 2013.

260 f: il.(color.) ; 30 cm

1. Tese (Doutorado em Informática) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013.

Inclui Bibliografia.

1. Informática – Tese. 2. Arquitetura de Software. 3. Degradação Arquitetural. 4. Anomalia de Código. 5. Anomalia de Código Arquiteturalmente Relevante. I. Staa, Arndt. II. Garcia, Alessandro. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. IV. Título.

CDD: 004

To my parents and husband

## Acknowledgements

First of all, I would like to thank my supervisors, Arndt von Staa and Alessandro Garcia. In particular, I want to express my deepest appreciation to Arndt for his amazing wisdom and for giving me freedom to shape my research. Arndt, thanks a lot for always keeping me on the right track and making me think. I must also thank Alessandro Garcia for teaching me how to do research, guiding me in all the process with his extensive knowledge, enthusiasm and motivation.

Special thanks to Nenad Medvidovic for his honest criticism, insightful discussions, questions and for helping me during my stay in Los Angeles.

I am also thankful to the members of my examining committee, who generously contributed with their time and expertise: Carlos Lucena, Renato Cerqueira, Yuanfang Cai and Paulo Borba.

Very special thanks to Julio for the unique way in which he loves, understands, and supports me day by day, to my parents and to my family who have believed in me and encouraged me since the first day .

Special thanks to Chico and Camila for the long talks and for supporting me in the bureaucratic tasks, and to Roberta for the hard work together and for reviewing the English in this thesis.

I thank Rosi for supporting me and helping me.

Thanks to my colleagues from the Laboratory of Software Engineering (LES) at PUC-Rio for the wonderful environment, to my colleagues from the Opus Research Group at PUC-Rio for the technical discussions and to my colleagues and staff members from the Department of Computer Science at USC for making my stay there an extraordinary experience.

Thanks to Vera for her constant smile and for lending me her phone booth.

Many thanks to the faculty and staff members from the Informatics Department at PUC-Rio, who helped me during the last six years.

My doctoral studies were financially supported by CNPq and CAPES. These institutions have all my gratitude.

## Abstract

Macía, Isela; Staa, Arndt; Garcia, Alessandro. **On the Detection of Architecturally-Relevant Code Anomalies in Software Systems**. Rio de Janeiro, 2013. 260p. DSc. Thesis – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Code anomalies can signal software architecture degradation. However, the identification of architecturally-relevant code anomalies (i.e. code anomalies that strongly imply architectural deficiencies) is particularly challenging due to: (i) lack of understanding about the relationship between code anomalies and architectural degradation, (ii) the focus on source code anomaly detection without considering how it relates to the software architecture, and (iii) lack of knowledge about how reliable these detection techniques are when revealing architecturally-relevant code anomalies. This thesis presents techniques for identifying architecturally-relevant code anomalies. Architecture-sensitive metrics and detection strategies were defined to overcome the limitations of conventional detection strategies. These metrics and strategies leverage traces that can be established between architectural views and system implementation. The thesis also documents code anomaly patterns (i.e. recurring anomaly relationships) that are strongly related to architectural problems. A tool, called SCOOP, was developed to collect the architecture-sensitive metrics, apply the new detection strategies, and identify the documented code anomaly patterns. Using this tool, we evaluated our technique in a series of empirical studies, comparing its accuracy with that of conventional detection techniques when identifying architecturally-relevant code anomalies.



## Keywords

Software Architecture; Architectural Degradation; Code Anomaly; Architecturally-Relevant Code Anomaly.

# Resumo

Macía, Isela; Staa, Arndt; Garcia, Alessandro. **Detecção de Anomalias de Código Arquiteturalmente Relevantes em Sistemas de Software**. Rio de Janeiro, 2013. 260p. Tese de Doutorado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Anomalias de código podem sinalizar a degradação da arquitetura de software. No entanto, a identificação de anomalias de código arquiteturalmente relevantes (ou seja, aquelas que implicam em deficiências arquiteturais) é particularmente difícil devido: (i) a falta de compreensão sobre a relação existente entre anomalias de código e degradação arquitetural, (ii) ao fato do processo de detecção de anomalias ter como foco somente o código fonte, sem considerar como ele se relaciona com sua arquitetura, e (iii) a falta de conhecimento sobre a confiabilidade das técnicas de detecção em revelar anomalias de código que são arquiteturalmente relevantes. Esta tese apresenta técnicas para identificar anomalias de código que são arquiteturalmente relevantes. Métricas sensíveis à arquitetura e estratégias de detecção foram definidas para superar as limitações das técnicas de detecção convencionais. Estas métricas e estratégias aproveitam rastros que podem ser estabelecidos entre as visões arquiteturais e a implementação dos sistemas. A tese também documenta padrões de anomalias de código (ou seja, relações recorrentes de anomalias) que estão relacionados com problemas arquiteturais. Uma ferramenta, chamada de SCOOP, foi desenvolvida para coletar as métricas sensíveis à arquitetura, aplicar as novas estratégias de detecção, e identificar os padrões de anomalias de código. Usando esta ferramenta, a técnica proposta foi avaliada em uma série de estudos empíricos, comparando sua acurácia com técnicas convencionais de detecção durante o processo de identificação de anomalias de código que são arquiteturalmente relevantes.

## Palavras-chave

Arquitetura de Software; Degradação Arquitetural; Anomalia de Código; Anomalia de Código Arquiteturalmente Relevante.

## Summary

## List of Figures

## List of Tables