

An empirical study on the interplay between semantic coupling and co-change of software classes

Extended Abstract

Nemitari Ajenka

Department of Computer Science
Edge Hill University
United Kingdom
Nemitari.Ajenka@edgehill.ac.uk

Andrea Capiluppi

Department of Computer Science
Brunel University London
United Kingdom
Andrea.Capiluppi@brunel.ac.uk

Steve Counsell

Department of Computer Science
Brunel University London
United Kingdom
Steve.Counsell@brunel.ac.uk

ABSTRACT

The evolution of software systems is an inevitable process which has to be managed effectively to enhance software quality. Change impact analysis (CIA) is a technique that identifies impact sets, i.e., the set of classes that require correction as a result of a change made to a class or artefact. These sets can also be considered as ripple effects and typically non-local: changes propagate to different parts of a system.

Two classes are considered logically coupled if they have co-changed in the past; past research has shown that the precision of CIA techniques increases if logical and semantic coupling (i.e., the extent to which the lexical content of two classes is related) are both considered. However, the relationship between semantic and logical coupling of software artefacts has not been extensively studied and no dependencies established between these two types of coupling. Are two often co-changed artefacts also strongly connected from a semantic point of view? Are two semantically similar artefacts bound to co-change in the future? Answering those questions would help increase the precision of CIA. It would also help software maintainers to focus on a smaller subset of artefacts more likely to co-evolve in the future.

This study investigated the relationship between semantic and logical coupling. Using Chi-squared statistical tests, we identified similarities in semantic coupling using class corpora and class identifiers. We then computed Spearman's rank correlation between semantic and logical coupling metrics for class pairs to detect whether semantic and logical relationships co-varied in OO software. Finally, we investigated the overlap between semantic and logical relationships by identifying the proportion of classes linked through both coupling types. Our empirical study and results were based on seventy-nine open-source software projects. Results showed that: (a) measuring the semantic similarity of classes by using their identifiers is computationally efficient; (b) using identifier-based coupling can be used interchangeably with semantic similarity based on their corpora, albeit not always; (c) no correlation between the strengths of semantic and change coupling was found. Finally, (d) a directional relationship between the two was identified; 70% of

semantic dependencies are linked through change coupling but not vice versa.

Based on our findings, we conclude that identifying more efficient methods of semantic coupling computation as well as a directional relationship between semantic and change dependencies could help to improve CIA methods that integrate semantic coupling information. This may also help to reveal implicit dependencies not captured by static source code analysis.

CCS CONCEPTS

• **General and reference** → *Metrics*; • **Software and its engineering** → *Object oriented development*;

KEYWORDS

Semantic coupling, co-change, Java software, metrics

ACM Reference Format:

Nemitari Ajenka, Andrea Capiluppi, and Steve Counsell. 2018. An empirical study on the interplay between semantic coupling and co-change of software classes: Extended Abstract. In *ICSE '18: ICSE '18: 40th International Conference on Software Engineering*, May 27-June 3, 2018, Gothenburg, Sweden. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/3180155.3190833>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ICSE '18, May 27-June 3, 2018, Gothenburg, Sweden

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5638-1/18/05.

<https://doi.org/10.1145/3180155.3190833>