M. Tichy, E. Bodden, M. Kuhrmann, S. Wagner, J.-P. Steghöfer (Hrsg.): SE 2018,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2018 91

A framework for capturing, statistically modeling and analyzing the evolution of software models

Hamed Shariat Yazdi[1] Lefteris Angelis[2] Timo Kehrer[3] Udo Kelter[4]

Abstract: In this work, we report about a recently developed *framework for capturing, statistically modeling and analyzing the evolution of software models*, published in [Ya16]. The framework captures the changes between revisions of models in terms of both low-level (internal) and high-level (developer-visible) edit operations applied between revisions. Evolution is modeled statistically by using ARMA, GARCH and mixed ARMA-GARCH time series models. Forecasting and simulation aspects of these time series models are thoroughly assessed, and the suitability of the framework is shown by applying it to a large set of design models of real Java systems. A main motivation for, and application of, the resulting statistical models is to control the generation of realistic model histories which are intended to be used for testing model versioning tools. Further usages of the statistical models include various forecasting and simulation tasks.

Keywords: Model-Driven Engineering, Software model evolution analysis, Time series analysis Forecasting, Simulation

Summary

Model-Driven Engineering (MDE) strongly depends on comprehensive, high-quality tool support. MDE tool suites should offer not only basic services such as editing, checking and translation of models, but also support team collaboration and project management by services for model versioning and history analysis. This paper concentrates on the latter group of services, which includes services for comparing, merging, and patching models and for analyzing model histories, e.g. in order to assess the past development process or to forecast the future amount of changes. A common substantial problem in this context is the shortage or the complete lack of adequate test models for evaluating novel solutions. One way out of this problem is to generate realistic model histories, which leads back to the problem to understand how models evolve and to express the evolution, e.g. using statistical methods. However, less is known about how models of software systems evolve over time and which statistical model can describe their evolution.

This paper makes two main contributions: (1) A methodological framework for capturing, statistically modeling, forecasting and simulating the evolution of software models within a project. Such statistical models, forecasts and simulations can be used for a variety of purposes, including forecasting of the amount of changes in the next releases of the system and software cost/effort estimation methods [JS07]. (2) A test model generator which is able

[1] Computer Science Institute III, University of Bonn, Germany, shariat@cs.uni-bonn.de
[2] Department of Informatics, Aristotle University of Thessaloniki, Greece, lef@csd.auth.gr
[3] Department of Computer Science, Humboldt-University Berlin, Germany, timo.kehrer@informatik.hu-berlin.de
[4] Software Engineering Group, University of Siegen, Germany, kelter@informatik.uni-siegen.de

to produce artificial histories of model revisions which „behave“ as specified by a statistical model. These artificial model histories are intended to be used as benchmarks for various model versioning tools [PYK11; PYK12].

State-of-the-art approaches to understand the evolution of models of software systems are based on software metrics and similar static properties; the extent of the changes between revisions of a software system is expressed as differences of metrics values, and further statistical analyses are based on these differences (see, e.g. [VSN07]). Unfortunately, such approaches do not properly reflect the dynamic nature of changes. In contrast to this, our methodological framework assumes changes to be expressed structurally by means of edit operations which modify the abstract syntax graph (ASG) of a model. We support two levels of abstraction of edit operations, namely „low-level“ graph editing operations, which include creations and deletions of nodes and edges in an ASG, and „user-level“ edit operations, which are offered by modern model editors and which typically comprise many ASG changes. At both levels of abstraction, we were able to statistically model the evolution of design models of real world Java projects using two kinds of time series models: ARMA and mixed ARMA-GARCH models. We finally addressed the non-trivial problem of how to simulate the evolution in the artificially generated test models and how to generate more statistically realistic histories of test models. In this regard, we used our simulated time series sequences to configure our model generator. The simulated sequences were interpreted as sizes of model differences which were applied between subsequent revisions of models in the history. The inclined readers will find more information about other mathematical properties of models' evolution in [Ya13]. A more detailed work in presented in [Ya15].

References

[JS07] Jorgensen, M.; Shepperd, M.: A systematic review of software development cost estimation studies. IEEE Transactions on software engineering 33/1, 2007.

[PYK11] Pietsch, P.; Yazdi, H. S.; Kelter, U.: Generating realistic test models for model processing tools. In: IEEE/ACM Intl. Conf. on Automated Software Engineering. 2011.

[PYK12] Pietsch, P.; Yazdi, H. S.; Kelter, U.: Controlled Generation of Models with Defined Properties. In: Software Engineering. Pp. 95–106, 2012.

[VSN07] Vasa, R.; Schneider, J.-G.; Nierstrasz, O.: The inevitable stability of software change. In: IEEE Intl. Conf. on Software Maintenance. 2007.

[Ya13] Yazdi, H. S.; Pietsch, P.; Kehrer, T.; Kelter, U.: Statistical Analysis of Changes for Synthesizing Realistic Test Models. In: Software Engineering. 2013.

[Ya15] Yazdi, H. S.: Statistical Analysis and Simulation of Design Models Evolution, PhD thesis, University of Siegen, 2015.

[Ya16] Yazdi, H. S.; Angelis, L.; Kehrer, T.; Kelter, U.: A framework for capturing, statistically modeling and analyzing the evolution of software models. Journal of Systems and Software 118/, 2016.