CrossMark

A guest editorial: special issue on search based software engineering and data mining

Marouane Kessentini[1] · Tim Menzies[2]

The concept of "search" has been explored in many fields such as Search-Based Software Engineering (SBSE) and Data Mining (DM). Thus, there is a strong theoretical connection between the goals of Search-Based SE and data mining. Therefore, recent studies have explored combinations of DM and SBSE, using a range of techniques in several areas including: tuning control parameters for a genetic algorithm using data-mining; learning effective mutations; finding explanation of the search results obtained by heuristic search; the ill-defined fitness function problems; etc.

The goal of this special issue was to understand the cost/benefit tradeoffs in combining SBSE and DM based on real-world case studies covering several aspects of the software lifecycle. This special issue presents some of the latest innovative results in that direction.

In "An Exploratory Study for Software Change Prediction in Object-Oriented Systems using Hybridized Techniques", Malhotra et al. proposed a hybrid approach, combining SBSE and DM, for predicting change prone classes in six application packages of a popular operating system for mobile apps. The results of the study confirm that the use of hybridized techniques for developing models to identify change prone classes outperforms several traditional machine learning techniques.

A "Meta-learning Based Selection of Software Reliability Models" approach was proposed by Caiuta et al.. In this paper, the authors used Meta-learning technique for the selection software reliability models (SRMs). The technique includes three main

✉ Marouane Kessentini
   marouane@umich.edu

   Tim Menzies
   tim.menzies@gmail.com

[1] University of Michigan, Dearborn, MI, USA

[2] North Carolina State University, Raleigh, NC, USA

steps: meta-knowledge extraction, meta-learning and classification. The results show statistical difference between the Meta-learning approach and the choice of the worst performing model with a large stochastic difference.

A challenging problem was addressed by Aleti et al. in the "Analysing the Fitness Landscape of Search-Based Software Testing Problems" paper. The goal of the proposed approach is to give indications about how successful the search has been based on the fitness landscape characterization. Another interesting paper on the use of interactive evolutionary algorithms and machine learning for the next release problem was proposed by Araújo et al. in "An Architecture based on Interactive Optimization and Machine Learning applied to the Next Release Problem". A large scale empirical study with software confirms the feasibility of the proposed architecture to incorporate human knowledge during the optimization process.

The objectives of the special issue were reached in terms of advancing the state of the art of combining SBSE and DM techniques to challenging problems in software engineering. Several real-world problems in these areas were formulated for the first time as optimization problems and most of the proposed contributions showed very promising results that outperform existing studies. Some of the proposed approaches were also validated in an industrial setting that show the applicability of search techniques in realistic and scalable environments.

Please enjoy reading these novel contributions!