Applied Soft Computing xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

The experimental applications of search-based techniques for model-based testing: Taxonomy and systematic literature review

Aneesa Saeed *, Siti Hafizah Ab Hamid, Mumtaz Begum Mustafa

Department of Software Engineering, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia

A R T I C L E I N F O

Article history:
Received 20 October 2015
Received in revised form 25 April 2016
Accepted 16 August 2016
Available online xxx

Keywords:
Software testing
Systematic literature review
Model-based testing
Search-based techniques
Taxonomy
Test case generation

A B S T R A C T

Context: Model-based testing (MBT) aims to generate executable test cases from behavioral models of software systems. MBT gains interest in industry and academia due to its provision of systematic, automated, and comprehensive testing. Researchers have successfully applied search-based techniques (SBTs) by automating the search for an optimal set of test cases at reasonable cost compared to other more expensive techniques. Thus, there is a recent surge toward the applications of SBTs for MBT because the generated test cases are optimal and have low computational cost. However, successful, future SBTs for MBT applications demand deep insight into its existing experimental applications that underlines stringent issues and challenges, which is lacking in the literature.
Objective: The objective of this study is to comprehensively analyze the current state-of-the-art of the experimental applications of SBTs for MBT and present the limitations of the current literature to direct future research.
Method: We conducted a systematic literature review (SLR) using 72 experimental papers from six data sources. We proposed a taxonomy based on the literature to categorize the characteristics of the current applications.
Results: The results indicate that the majority of the existing applications of SBTs for MBT focus on functional and structural coverage purposes, as opposed to stress testing, regression testing and graphical user interface (GUI) testing. We found research gaps in the existing applications in five areas: applying multi-objective SBTs, proposing hybrid techniques, handling complex constraints, addressing data and requirement-based adequacy criteria, and adapting landscape visualization. Only twelve studies proposed and empirically evaluated the SBTs for complex systems in MBT.
Conclusion: This extensive systematic analysis of the existing literature based on the proposed taxonomy enables to assist researchers in exploring the existing research efforts and reveal the limitations that need additional investigation.

1. Introduction

Model-based testing (MBT) aims to produce executable test cases by consistently analyzing the behavioral design models of a software system (such as a Finite State Machine) by following a test strategy (for example: a coverage criterion which the target is to cover specific model features). The model is an abstract representation of the software behavior. To fully automate MBT, three tasks are required: constructing models for testing, defining a suitable testing strategy and adequacy criteria, and generating test data for executing test cases. Recently, MBT gained increasing interest in both industry and academia. This is visible from several academic studies [1–3] and industrial projects [4,5] on MBT. Many strategies have been applied for MBT, such as combinatorial testing [6], model checking [7], and SBTs [8]. The latest endeavor is to deploy SBTs to MBT and this recently become a field of interest as reported in [9,10]. One reason among the advantages is the capability of SBTs to find the optimal set of test cases in terms of maximum coverage criteria among all possible test cases at minimum cost. Specifically, the process of the test case generation can be formulated as an optimization process: The output of the test case generation could be hundreds of thousands of test cases for a certain system under test (SUT). From this context, there is a need to select systematically those that adhere to particular coverage criteria at a reasonable cost and that are predicted to be fault detecting. Thus, the generation of test cases can be reformulated as a search problem that aims to find the required or optimal set of test cases from the space of the

* Corresponding author.
E-mail addresses: aneesa.saeed@siswa.um.edu.my (A. Saeed), sitihafizah@um.edu.my (S.H. Ab Hamid), mumtaz@um.edu.my (M.B. Mustafa).

http://dx.doi.org/10.1016/j.asoc.2016.08.030

all possible test cases. Studies applied SBTs for MBT showed their significant performance compared to other techniques. For example, studies concluded that SBTs outperformed model checking for testing dynamic systems [11], and embedded real-time systems [12]. Another study declared that the generating test cases using model checkers is more expensive than using heuristic techniques [13]. There are still several problems to be properly understood and tackled for such SBTs to work because the MBT has different characteristics based on chosen application domains and type of models that makes it different from the code-based testing domain. Deep insights into the existing work can help in the future research.

SBTs are a group of generic algorithms that utilize heuristics to obtain solutions for optimization problems at an affordable computational time cost. Researchers apply SBTs for automatic test case generation based on a test objective (adequacy criteria) that is formulated as a fitness function [14]. The applications of SBTs for test case generation is referred to as search-based software testing (SBST) [15], and these techniques are not applied necessarily to MBT. The first study on SBST [16] is further probably to be the first study on SBTs for software engineering, called search-based software engineering (SBSE). The critical key in the applications of SBTs is the fitness function. The aim of using the fitness function is to guide the search for the test data that maximizes the achievement of the test objective. Therefore, different fitness functions were proposed to capture different test purposes, such as structural testing [17–25], functional testing [26,27], stress testing [28,29], and non-functional properties testing [30]. Most commonly studies of SBTs in software testing have focused on structural testing, while the recent researchers have paid their attentions to functional testing since 2003 [27]. With respect to structural testing, the target of generating test cases is to cover the internal structure of the system source code or model. For functional testing, the generated test cases are constructed from the specification of system behavior to detect the faults in the system functionality.

Automatic test case generation is an attempt to automate the software testing process and reduce its cost. The lower the cost of test case generation, the lower the cost of the software testing process and software development will be. The cost of software testing consumes almost half of the entire software development cost [31]. SBTs and MBT have received much interest in the research studies in recent times [32]. Anand et al. [32] conducted a survey on test case generation approaches based on the expert knowledge of each specific approach, in which they presented the current state-of-the-art and future challenges of MBT and SBTs separately, as a part of the survey. This survey did not give an overview of the application of SBTs for MBT. Extensive review papers presented the state-of-the-art of each field (SBTs and MBT) separately as the following: the MBT field has been intensively reviewed in [33–35,10,36,37]. Alternately, comprehensive studies [38–41] have reviewed generally SBTs for the test case generation. Another review study [15] extensively surveyed SBTs for non-functional testing. A recent study presented the achievement, challenges and open problems in SBST domain [42]. With respect to reviewing SBTs for MBT, Harman et al. [43] surveyed the applications of SBTs for MBT as apart of the SBSE survey. Another study [44] has noted the opportunities and future challenges for applying SBTs for the test case generation from Event-B models. Concerning the taxonomy, only the survey conducted by Utting et al. [10] presented a taxonomy to classify MBT approaches, but their taxonomy is not generic enough to classify all the detailed concepts in SBTs for MBT. In conclusion, all these surveys have provided a good overview of the current state-of-the-art of each domain separately; however, none of them enable to provide comprehensive inspections of the existing applications of SBTs for MBT. We believe that analyzing the research activity in this field is important to explore the available research opportunities for further investigations based on the studies limitations. Moreover, the proposed taxonomy can be considered as a basic framework to classify existing work in this field. The taxonomy may also become useful in developing and judging new work. This study should be of interest to several stakeholders. First, the MBT research community since this review indicates a discrepancy between MBT aspects (such as model transformation, constraint, and coverage criteria) and SBTs. Second, search-based testing community since it shows how far SBST investigated in MBT. Finally, individuals and academia community which interested in applying SBTs for MBT because this study provides a comprehensive overview and open a gateway for new research opportunities.

To link the knowledge and provide a comprehensive overview to the state-of-the-art, the main contributions of this study are:

- Propose a basic classification framework in form of taxonomy to classify existing experimental applications of SBTs for MBT.
- Provide an overview of the current state-of-the-art in the SBTs for MBT domain.
- Highlights several limitations in SBTs for MBT in order to pave the way for future research efforts.

To conduct our SLR, we analyzed 72 experimental papers out of 546 retrieved papers from six data sources (from year 2001 to 2013). We then constructed a taxonomy by applying content analysis. Specifically, we iteratively gathered and analyzed a set of keywords from the papers to propose the taxonomy. We frequently updated the taxonomy until all the papers were covered and we utilized the taxonomy to classify the papers in terms of how they formulated the problem, how they proposed a solution, and how they assessed their solution. Consequently, this classification provided a comprehensive overview of the current experimental research on SBTs for MBT. Quantitative and cross analysis were then performed on the data to come up with intrinsic deficiencies and future directions.

The rest of this paper is organized as follows: Section 2 presents the methodology that we follow in this study. Section 3 presents and discusses the results and finding. The related surveys are described in Section 4. Threats to validity of this study is presented in Section 5. Finally, conclusion and future work are presented in Section 6.

2. Methodology

The methodology by which this SLR study was conducted is based on the guidelines proposed by Kitchenham et al. [45,46]. The guidelines organize the steps of conducting a SLR into three stages, planning, conducting, and reporting, as shown in Fig. 1. The aim of the first stage (planning) is to develop the review protocol, which encompasses: identifying research questions, search strategy, inclusion/exclusion criteria, data collection, and methods of synthesis. In the second stage (conducting), the focus is on executing the review protocol. The last stage (reporting) concerns how to elaborate on the final report.

2.1. Research questions

Defining the research questions is a potential step in defining the review protocol. We derived three questions to embody this study sub-objectives and form the basis of this SLR:

RQ 1: How can a basic classification framework be devised based on the current research on SBTs for MBT?
RQ 2: What is the current state of SBTs for the MBT research area?
RQ 3: What can be concluded from the current results that will help to preside future directions?