Computer Standards & Interfaces 48 (2016) 30–48

Contents lists available at ScienceDirect

Computer Standards & Interfaces

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

Intelligent software product line configurations: A literature review

Uzma Afzal[a,*], Tariq Mahmood[b,1], Zubair Shaikh[a,2]

[a] National University of Computer & Emerging Sciences, Department of Computer Science, Karachi, Pakistan
[b] Karachi Institute of Economics and Technology, College of Computing & Information Sciences, Karachi, Pakistan

ARTICLE INFO

Article history:
Received 18 December 2015
Received in revised form 11 March 2016
Accepted 28 March 2016
Available online 18 April 2016

Keywords:
Software product line
Literature review
Automated feature selection
Inconsistencies
Artificial intelligence
Industrial SPL tools
Predictive analytics

ABSTRACT

A software product line (SPL) is a set of industrial software-intensive systems for configuring similar software products in which personalized feature sets are configured by different business teams. The integration of these feature sets can generate inconsistencies that are typically resolved through manual deliberation. This is a time-consuming process and leads to a potential loss of business resources. Artificial intelligence (AI) techniques can provide the best solution to address this issue autonomously through more efficient configurations, lesser inconsistencies and optimized resources. This paper presents the first literature review of both research and industrial AI applications to SPL configuration issues. Our results reveal only 19 relevant research works which employ traditional AI techniques on small feature sets with no real-life testing or application in industry. We categorize these works in a typology by identifying 8 perspectives of SPL. We also show that only 2 standard industrial SPL tools employ AI in a limited way to resolve inconsistencies. To inject more interest and application in this domain, we motivate and present future research directions. Particularly, using real-world SPL data, we demonstrate how predictive analytics (a state of the art AI technique) can separately model inconsistent and consistent patterns, and then predict inconsistencies in advance to help SPL designers during the configuration of a product.

1. Introduction

A software product line (SPL) is a set of industrial software-intensive systems for configuring similar software products in which personalized feature sets are configured by different business teams [24,40,48,60,84]. It is typically implemented within the context of enterprise resource planning (ERP) systems. SPLs have seen global acceptance in industries such as Boeing, Nokia, Samsung, Ericsson, HP, General Motors, LG Electronics, Lucent Technologies, Phillips, Toshiba, Siemens and Robert Bosch [48,64,66]. feature modeling (FM) is the standard SPL design approach which configures a business product through feature sets, i.e., collections of features. An SPL system comprises different modules which are configured individually. Each configuration is assigned to a business team, which selects features manually from large-scaled feature sets (comprising hundreds to thousands of features), or modifies existing selection rules based on personal experience. Amongst other issues, this often generates inconsistencies (conflicts) when different feature sets are integrated in the final product. For instance, a business unit working on a Car Assembly SPL can add a manual gear to a car product, while another adds automatic gear; this is an inconsistency as both gears cannot co-exist.

The adoption of an inconsistency resolution process is critical for SPL management. Otherwise, business teams can continue to work in conflict with each other, leading to reduced productivity and morale, misuse of business assets, and non-compliance with customer requirements. This can become complicated with changing product requirements of SPL customers. Also, the increasing size of feature repositories are generating more inconsistencies, leading business teams to take risks in selecting features [35,63,65]. These problems have already been highlighted in different case studies and reports by various SPL industries [3,11,44,48,74,79]. Hence, a key SPL challenge is to ensure an autonomous, inconsistency-free configuration in the face of complex feature models, customer requirements and conflicting configuration viewpoints of business teams [48].

Artificial intelligence (AI) techniques [38,57] can provide the best solution to this problem. AI focuses on the design and implementation of intelligent programs which attempt to reproduce the way humans act rationally. AI techniques have seen widespread success in various industries, e.g., healthcare, engineering and other

* Corresponding author at: Federal Urdu University of Arts Science & Technology, Department of Computer Science, Karachi, Pakistan. Tel.: +92 21 99244141 9.
E-mail addresses: k123501@nu.edu.pk (U. Afzal), mahmood@pafkiet.edu.pk (T. Mahmood), zubair.shaikh@nu.edu.pk (Z. Shaikh).
[1] Tel.: +92 21 35091114 7; fax: +92 21 35091118.
[2] Tel.: +92 21 111 128 128; fax: +92 21 34100549.

http://dx.doi.org/10.1016/j.csi.2016.03.003
0920-5489/

applications [6,15,67]. AI applications to SPL feature modeling have the potential to facilitate business teams by autonomously identifying and resolving issues with more efficient product configurations, lesser inconsistencies and more optimized resources. This paper is focused on a literature review of AI applications to SPL configuration issues. Our motivation is three-fold: 1) determine the frequency of these AI applications and evaluate their validity, 2) identify knowledge gaps which demand further research, and 3) identify avenues of future research to potentially increase the frequency and usability of these AI applications. We conduct our review by answering three relevant research questions:

- Q1: What are the research and industrial applications which employ AI to SPL configuration issues?
- Q2: What are potential research avenues to increase the frequency and validity of AI applications to SPL configuration issues?
- Q3: What impact can predictive analytics [25,58,70] (a state of the art AI technique) have on inconsistency detection and resolution in large-scale software product lines?

The motivation for Q1 is to determine the frequency of AI applications, gauge their benefits to SPL configuration issues, and identify their limitations. We answer Q1 in Section 5 by extracting 173 research articles from online digital sources, out of which we identified only 19 relevant ones. These works employ limited feature models having 30–50 features and use the following traditional AI techniques to resolve inconsistencies and other issues: logic-based knowledge representation and reasoning, ontological modeling and reasoning, constraint satisfaction techniques, and optimization methods. We categorized these works in a typology of 8 SPL feature modeling perspectives which we identified using affinity diagram process. These are the following: 1) identification of inconsistencies in feature model, 2) debugging (inconsistencies removal) of conflicting feature models, 3) management of variability in selecting features, 4) void feature model and errors, 5) management of priorities in selecting features, 6) testing (verification) of feature models, 7) derivation of products from SPLs, and 8) optimization of feature model through combinatorial optimization. Moreover, we reviewed several industrial SPL tools and found that only Gears and Purevariants tools use logic-based techniques to identify and resolve a subset of inconsistencies [4,50].

These results show a limited effort in research and application of AI to SPL feature modeling. The research efforts have used small non-representative SPL feature sets without any experiments in real-world SPL system, and application in industry is confined to only 2 tools. Also, all applications use traditional AI techniques with no focus on novel ones. This is our motivation for Q2. Considering the concrete benefits which AI can potentially bring to SPL industry, we answer Q2 by providing several future research directions to motivate further research and application in this domain. Specifically, we 1) propose application of AI techniques across unapplied SPL feature modeling perspectives, 2) propose a comparison of different constraint satisfaction algorithms for feature modeling application, and 3) propose application of novel optimization algorithms [12] to determine optimal configurations for large representative feature sets.

[71] presents predictive analytics as an omnipresent science which has an impact on everything, more over [69] emphasizes the need to incorporate predictive analytics into information systems. In SPL product configurations, patterns of feature selection exist due to the regularity of human behavior. So, for instance, a particular business team could configure each SPL product in a similar way by making the decisions of selecting the similar sequences of features in each product which generates potentially useful data based on the prior knowledge and these decisions are hidden in the form of patterns in the datasets, learning (PA) is the process which turns this process around [80]. Our motivation for Q3 is to identify (mine) these hidden sequences, or patterns, in order to understand the particular patterns leading to inconsistent, or consistent configurations. Such insights could provide the appropriate decision support to the business teams, for instance, by indicating current patterns which could potentially lead to future inconsistencies. Predictive analytics has already seen a global adoption in the industrial sector [25,58,70]. We demonstrate how PA can learn and classify patterns of selection of feature sets for both consistent and inconsistent feature models. Using a real SPL dataset of the vendor master module, we use decision tree (a specific PA algorithm) to learn these patterns in previously configured products, and show how they can be used to indicate potentially inconsistent and consistent configurations in advance.

Based on the above discussion, this paper has four contributions:

1. The first literature review of research articles which have used AI to address SPL configuration issues.
2. A review of AI applications within standard industrial SPL tools to identify and resolve feature modeling issues.
3. A proposal of pragmatic future research directions to stimulate further research in this domain.
4. The first application of predictive analytics to predict consistent and inconsistent SPL product configurations in advance.

This paper is structured as follows. In Section 2, we describe our research methodology. To improve the readability of the literature review and provide readers an insight to the working of AI techniques and SPL issues we add two more sections before the main review, i.e. Sections 3 and 4. In Section 3, we discuss the SPL functionality with respect to feature model design and illustrate the problems of SPL configuration using the vendor master module example. In Section 4, we briefly review the AI techniques relevant to the literature review. After that, we describe the results of the review in Section 5 and Section 6. In Section 7, we describe the future research directions and finally, we conclude our paper in Section 8.

2. Research methodology

We now present our research methodology for extracting the relevant articles which employ AI techniques to resolve SPL configuration issues. Our methodology is adapted from guidelines discussed in [30]. To develop queries for searching the relevant articles, we extracted three representative keywords from our research questions, i.e., “software product lines”, “feature model inconsistencies” and “predictive analytics”. We also identified their related terms commonly used in research literature. For instance, “software product family” and “product family” are related terms of “software product lines” and “data mining” is a related term of “predictive analytics”. We combined keywords and their related words using Boolean “OR” operator. We combined different types of representative keywords using Boolean “AND” operator. We initially determined state of the art domains of AI, which are as follows [37,51,54,56]: logical AI, problem search, pattern recognition, knowledge representation, inference and reasoning, constraint satisfaction, planning, ontology, heuristics, genetic programming, evolutionary computation, and learning from experience. We then brainstormed each of these domains separately in order to determine its power in solving the SPL feature modeling issue.

- Logic techniques, which can be used to model the process of feature selection by the SPL team designers.
- Knowledge representation and ontology, which can be used to represent different processes of feature selection.