Chapter 10
Recovery and Refinement of Business Process Models for Web Applications

Alessandro Marchetto
Independent Researcher, Italy

Chiara Di Francescomarino
Fondazione Bruno Kessler (CIT), Italy

ABSTRACT

Web Applications (WAs) have been often used to expose business processes to the users. WA moderniza-tion and evolution are complex and time-consuming activities that can be supported by software docu-mentation (e.g., process models). When, as often happens, documentation is missing or is incomplete, documentation recovery and mining represent an important opportunity for reconstructing or completing it. Existing process-mining approaches, however, tend to recover models that are quite complex, rich, and intricate, thus difficult to understand and use for analysts and developers. Model refinement approaches have been presented in the literature to reduce the model complexity and intricateness while preserving the capability of representing the relevant information. In this chapter, the authors summarize approaches to mine first and refine later business process models from existing WAs. In particular, they present two process model refinement approaches: (1) re-modularization and (2) reduction. The authors introduce the techniques and show how to apply them to WAs.

INTRODUCTION

In the last decade, Web applications (in short WAs) have been often used to expose business processes to the users. WAs, in fact, are complex applications mainly implemented by means of distributed server-side components (written, e.g., in JSP, Java, Php, Perl) that are exercised by client-side components (written, e.g., in HTML, Ajax, Flash) running on a web-browser user-interface. Web applications are event-based applications in which the user executes the underlying application processes by interacting with the application user-interface.

WAs maintenance, evolution and modernization are complex and time consuming activities that can be supported by software documentation, e.g., business process models underlying the applications, as well as by application and domain knowledge. Unfortunately, often, such documentation and knowledge do not exist or, if they exist, they are not aligned with the actual application implementation because in most of the cases application maintenance and evolution activities are not well-documented. In these situations, documentation recovery and mining techniques can represent an appealing opportunity to reconstruct and complete the missing and inaccurate documentation.

Existing techniques to recover and mine process models (e.g., (van der Aalst et al. 2003), (Zou et al. 2009)), however, tend to generate models that are quite complex, rich and intricate, thus difficult to understand and use for analysts and developers. Such process models are known in the literature as "spaghetti" processes (Veiga et al. 2009).

To limit process model complexity and intricateness (e.g., (Alves et al. 2006), (Mendling et al. 2007)) several approaches have been proposed in the literature. In this chapter, we will present some relevant types of process refinement techniques, that is state-of-the-art techniques that improve process model readability while preserving the capability of the process to represent the relevant information about the application. In particular, we will detail two different approaches to improve the process model readability: (1) process model re-modularization; and (2) process model reduction.

Process model re-modularization (e.g., (Bose et al. 2009), (Di Francescomarino et al. 2010), (Reijers et al. 2008)) consists of improving the recovered process models by increasing their modularity degree. To this aim, process elements are clustered together according to heuristics and then represented in the initial processes through subprocesses (i.e., complex activities).

Process model reduction (e.g., (van der Aalst et al. 2003), (Marchetto et al. 2011), (Marchetto et al. 2011b), (Tomasi 2012)), instead, consists of reducing process model complexity by removing process elements according to heuristics.

In this chapter, after a summarized description of context and problems related to WAs maintenance and modernization activities (e.g., (Matthijssen et al. 2010), (Patel et al. 2007)), we will introduce the use of dynamic analysis to recover and mine business process models, represented using the Business Process Modeling Notation (BPMN, http://www.omg.org/spec/BPMN/1.2) and describing the processes underlying WAs, i.e., the processes realized and implemented by WAs. We will hence describe the information that can be collected during such an analysis and we will detail one of the existing techniques for recovering process models from existing WAs.

We will then present approaches and techniques to refine business process models recovered from existing Web applications. In this context we will introduce both the architecture and the practical use of tools supporting analysts and developers in applying the above mentioned approaches for the software application maintenance and modernization.

Last but not least, we will report a case study in which real Web applications are analyzed and their business processes reconstructed by means of the presented recovery and refinement techniques. In the study, expert analysts have been asked to provide their feedback about the recovered processes.

The rest of the chapter is organized in the following sections:

- **Background:** The section introduces the notions used in the chapter and presents references useful for the reader.
- **Related Works:** The section summarizes the state-of-the-art in terms of process recovery and mining.

Related Content

Decoupling Computation and Result Write-Back for Thread-Level Parallelization
Hiroaki Hirataand Atsushi Nunome (2020). *International Journal of Software Innovation (pp. 19-34).*
www.irma-international.org/article/decoupling-computation-and-result-write-back-for-thread-level-parallelization/256234

Malware Analysis and Its Mitigation Tools
D. R. Janardhana, A. P. Manu, K. Shivannaand K. C. Suhas (2023). *Malware Analysis and Intrusion Detection in Cyber-Physical Systems (pp. 263-284).*
www.irma-international.org/chapter/malware-analysis-and-its-mitigation-tools/331308

Multi-Platform Bluetooth Remote Control: Implementation and Results
Jarle Hansenand Gheorghita Ghinea (2012). *Handbook of Research on Mobile Software Engineering: Design, Implementation, and Emergent Applications (pp. 880-898).*
www.irma-international.org/chapter/multi-platform-bluetooth-remote-control/66504

Modelling Information Demand in an Enterprise Context: Method, Notation, and Lessons Learned
Magnus Lundqvist, Kurt Sandkuhland Ulf Seigerroth (2011). *International Journal of Information System Modeling and Design (pp. 75-95).*
www.irma-international.org/article/modelling-information-demand-enterprise-context/55489

Creating Applications for Real-Time Collaboration with XMPP and Android on Mobile Devices
Daniel Schuster, István Koren, Thomas Springer, Dirk Hering, Benjamin Söllner, Markus Endlerand Alexander Schill (2012). *Handbook of Research on Mobile Software Engineering: Design, Implementation, and Emergent Applications (pp. 824-844).*
www.irma-international.org/chapter/creating-applications-real-time-collaboration/66501