

Chapter 33

Mining Lifecycle Event Logs for Enhancing Service- Based Applications

Schahram Dustdar

Vienna University of Technology, Austria

Philipp Leitner

Vienna University of Technology, Austria

Franco Maria Nardini

ISTI-CNR, Pisa, Italy

Fabrizio Silvestri

ISTI-CNR, Pisa, Italy

Gabriele Tolomei

ISTI-CNR, Pisa, Italy

ABSTRACT

Service-Oriented Architectures (SOAs), and traditional enterprise systems in general, record a variety of events (e.g., messages being sent and received between service components) to proper log files, i.e., event logs. These files constitute a huge and valuable source of knowledge that may be extracted through data mining techniques. To this end, process mining is increasingly gaining interest across the SOA community. The goal of process mining is to build models without a priori knowledge, i.e., to discover structured process models derived from specific patterns that are present in actual traces of service executions recorded in event logs. However, in this work, the authors focus on detecting frequent sequential patterns, thus considering process mining as a specific instance of the more general sequential pattern mining problem. Furthermore, they apply two sequential pattern mining algorithms to a real event log provided by the Vienna Runtime Environment for Service-oriented Computing, i.e., VRESCo. The obtained results show that the authors are able to find services that are frequently invoked together within the same sequence. Such knowledge could be useful at design-time, when service-based application developers could be provided with service recommendation tools that are able to predict and thus to suggest next services that should be included in the current service composition.

DOI: 10.4018/978-1-4666-2455-9.ch033

INTRODUCTION

The vast majority of nowadays software-based systems, ranging from the simplest, i.e., small-scale, to the most complex, i.e., large-scale, record massive amounts of data in the form of logs. Such logs could either refer to the functioning of the system as well as keep trace of any possible software or human interaction with the system itself. For this reason, logs represent a valuable source of hidden knowledge that can be exploited in order to enhance the overall performances of any software-based system.

Well-known examples of systems that have started trying to improve their performances by analyzing event logs are surely Web Search Engines (SEs).

Roughly, SEs are increasingly exploiting past user behaviors recorded in query logs in order to better understand people search intents, thus, for providing users with better search experiences. Indeed, by accurately recognizing and predicting actual user information needs, SEs are now able to offer more sophisticated functionalities (e.g., query suggestion) as well as better relevant result sets in response to a specific query (e.g., query diversification).

Moreover, there are plenty of modern enterprise software systems that need to operate in highly dynamic and distributed environments in a standardized way. Such systems implement their business logic according to the Service-Oriented Architecture (SOA) principles, thus, assembling their business processes as the composition and orchestration of autonomous, protocol-independent, and distributed logic units, i.e., software services.

Service-based systems and applications (SBAs) require proper run-time environments where their composing services can be searched, bound, invoked, monitored and managed. Therefore, SBA's run-time support might keep track of what is going on during the whole application lifecycle by roughly recording all such events to log files, i.e., service event logs.

Analysis of such service event logs could reveal interesting patterns, which in turn might be exploited for improving the overall performances of SOA's run-time frameworks as well as supporting SBA designers during the whole application lifecycle.

The main contribution of this work concerns the application of data mining techniques to a real-life service event log collected by the VRESCo SOA run-time framework. Our aim is to analyze the historical events stored on VRESCo in order to discover software services that are frequently invoked and composed together, i.e., process mining.

Although traditional process mining refers to a set of techniques and methodologies whose aim is to distill a structured process description from a set of actual traces of executions recorded in event logs, here we treat it as an instance of the sequential pattern mining problem.

The remaining of the work is structured as follows. First, we start describing background concepts and past work that somehow concerns with service event log analysis. Section 1 describes the information collected by SOA lifecycle event logs, in particular focusing on the VRESCo run-time framework. In Section 2, we propose how VRESCo event log may be analyzed for approaching our research challenge. Therefore, Section 3 shows the experiments we conduct on a real VRESCo log data set. Finally, we summarize the contributions we provide in this work together with any further idea that could be better investigated as future work.

BACKGROUND AND RELATED WORK

In this work, we present a use case for event log mining in service-based systems. This idea bears some resemblance to the established idea of business activity management (BAM) (Kochar, 2005). BAM considers the event-driven governance of

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/mining-lifecycle-event-logs-enhancing/73463

Related Content

A Survey of Parallel and Distributed Data Warehouses

Pedro Furtado (2009). *International Journal of Data Warehousing and Mining* (pp. 57-77).

www.irma-international.org/article/survey-parallel-distributed-data-warehouses/1826

An Approach to Improve Generation of Association Rules in Order to Be Used in Recommenders

Hodjat Hamidiand Elnaz Hashemzadeh (2017). *International Journal of Data Warehousing and Mining* (pp. 1-18).

www.irma-international.org/article/an-approach-to-improve-generation-of-association-rules-in-order-to-be-used-in-recommenders/188487

Knowledge Graph-Enhanced Interleaved Multi-Head Attention Knowledge Tracing

Zehan Guo, Honghai Guan, Chungang He, Ye Xuand Rui Liu (2025). *International Journal of Data Warehousing and Mining* (pp. 1-20).

www.irma-international.org/article/knowledge-graph-enhanced-interleaved-multi-head-attention-knowledge-tracing/377619

Impediments to Exploratory Data Mining Success

Jeff Zeanah (2004). *Organizational Data Mining: Leveraging Enterprise Data Resources for Optimal Performance* (pp. 280-299).

www.irma-international.org/chapter/impediments-exploratory-data-mining-success/27922

A Smart Recyclable Waste Collection System Based on IoT and an Application

Nadide Caglayanand Sule Itir Satoglu (2019). *Optimizing Big Data Management and Industrial Systems With Intelligent Techniques* (pp. 148-169).

www.irma-international.org/chapter/a-smart-recyclable-waste-collection-system-based-on-iot-and-an-application/218744