Chapter 5
Mining Patterns Using Business Process Management

Ishak H. A. Meddah
Oran University of Science and Technology – Mohamed Boudiaf, Algeria

Khaled Belkadi
Oran University of Science and Technology – Mohamed Boudiaf, Algeria

ABSTRACT

Process mining provides an important bridge between data mining and business process analysis. This technique allows for the extraction of information from event logs. In general, there are two steps in process mining: correlation definition or discovery and then process inference or composition. Firstly, the authors mine small patterns from log traces of two applications; those patterns are the representation of the execution traces of a business process. In this step, the authors use existing techniques. The patterns are represented by finite state automaton or their regular expression. The final model is the combination of only two types of small patterns that are represented by the regular expressions (ab)* and (ab*c)*. Secondly, the authors compute these patterns in parallel and then combine those small patterns using the composition rules. They have two parties. The first is the mine, where the authors discover patterns from execution traces, and the second is the combination of these small patterns. The pattern mining and the composition is illustrated by the automaton existing techniques.

INTRODUCTION

Many techniques have been proposed that mine patterns from execution traces. However; existing techniques mine only simple patterns, or a single complex pattern that is limited to a particular set of manually selected events.

Authors (Ammons, G and al., 2002) has recognized that patterns can be specified as regular languages, That allows the compact representation of patterns as regular expressions or finite state automata, and it allows the characterization of the pattern mining as a language learning problem.

Process mining approaches are fundamentally similar; each takes as input a static program or a dynamic traces or profile and produces one or more compact regular languages that specify the pattern representation or the workflow. However; the individual solutions differ in key ways.

In this paper, we present a new general approach to patterns mining that addresses several of the limitations of current techniques. Our insight is twofold. First, we recognize that instances of smaller patterns can be composed in parallel into larger patterns. Second, we observed also that the composition of small pattern can be in parallel.

We then leverage this insight to divide our work into two parties; The first one, we use a technique how we can mine two types of small patterns and we compose them by using standard algorithms for finite state automaton manipulation, and some special rules using by M.Gabel and Z.Su [(Gabel, M., & Su, Z, 2008), (Ammons, G and al., 2002)], the mining is also performed by symbolic mining algorithm [(Gabel, M & Su, Z, 2008, May), (Gao, X, 2013), (Guyet, T., & Quiniou, R, 2011)].

Our approach is an amelioration of existing work in pattern mining, it has been implemented in the java programming language with more log files of two application; the SKYPE and VIBER applications. The size of those applications log files is 10 GB, 15 GB, 18 GB, and 20 GB which are generated by log file generator.

The executions traces in our applications are the user's actions like call, the answer, and the messages…

RELATED WORK

Many techniques are suggested in the domain of process mining, we quote:

M.Gabel and al (Gabel, M., & Su, Z, 2008) present a new general technique for mining temporal specification, they realized their work in two steps; firstly they discovered the simple patterns using existing techniques, then combine these patterns using the composition and some rules like Branching and Sequencing rules.

Temporal specification expresses formal correctness requirement of an application's ordering of specific actions and events during execution, they discovered patterns from traces of execution or program source code; The simples patterns are represented using regular expression (ab)* or (ab*c)* and their representation using finite state automaton, after they combine simple patterns to construct a temporal specification using a finite state automaton.

G.Greco and al (Greco, G and al., 2006) discovered several clusters by using a clustering technique, and then they calculate the pattern from each cluster, they combine these patterns to construct a final model, they discovered a workflow scheme from, and then they mine a workflow using a Mine Workflow Algorithm, after they define many clusters from a log traces by using clustering technique and Process Discover Algorithm and some rules cluster.

Then they use a Find Features Algorithm to find a patterns of each cluster, finally they combine these patterns to construct a completely hierarchical workflow model.

In their clustering algorithm, clusters reflect only structural similarities among traces; they say that in future works extending their techniques to take care of the environment so that clusters may reflect not only structural similarities among traces, but also information about, e.g., users and data values.

H.R.Motahari-Nezhed and al (2008) use a service conversation log; first they split a log into several partitions, 2nd they discovered a model from each partition, and 3rd, they annotate the discover protocol model with various metadata to construct a protocol model from real-word service conversation logs.

10 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-patterns-using-business-process-management/197696

Related Content

Multilingual Web Content Mining: A User-Oriented Approach
Rowena Chauand Chung-Hsing Yeh (2004). *Intelligent Agents for Data Mining and Information Retrieval (pp. 88-100).*
www.irma-international.org/chapter/multilingual-web-content-mining/24157

Comparative Analysis of Machine Learning-Based Algorithms for Detection of Anomalies in IIoT
Bhupal Naik D. S., Venkatesulu Dondetiand Sivadi Balakrishna (2022). *International Journal of Information Retrieval Research (pp. 1-55).*
www.irma-international.org/article/comparative-analysis-of-machine-learning-based-algorithms-for-detection-of-anomalies-in-iiot/298647

A Study on Web Searching: Overlap and Distance of the Search Engine Results
Shanfeng Zhu, Xiaotie Deng, Qizhi Fangand Weimin Zhang (2004). *Intelligent Agents for Data Mining and Information Retrieval (pp. 208-225).*
www.irma-international.org/chapter/study-web-searching/24165

Interactive IR in OPAC Environments
Iris Xie (2008). *Interactive Information Retrieval in Digital Environments (pp. 29-52).*
www.irma-international.org/chapter/interactive-opac-environments/24523

Indexing Outline Shapes Using Simple Polygons
Saliha Aouat (2015). *International Journal of Information Retrieval Research (pp. 19-35).*
www.irma-international.org/article/indexing-outline-shapes-using-simple-polygons/128276