Improving Interactivity in Multidimensional Process Mining: The Interactive PMCube Explorer Tool

Thomas Vogelgesang

Department of Computer Science
University of Oldenburg, Germany
thomas.vogelgesang@uni-oldenburg.de

Abstract. Process mining is a set of techniques that analyze event logs in order to discover and enhance process models or to check their conformance to the event logs. Multidimensional process mining (MPM) is an emerging concept that adopts the idea of data cubes and OLAP to process mining. In recent years, different approaches and tools for MPM have been proposed. Despite the explorative character of MPM, such tools are still quite limited in their interactivity. For example, they lack direct interaction with process models and restrict the dynamic analysis workflow by forcing the users to follow a predefined sequence of analysis steps. In this demo, we present a novel tool for MPM that aims to overcome these limitations in order to provide seamless interaction. It is based on a multilevel operator framework which enables the user to perform, undo, and redo the analysis steps in an arbitrary order. The definition of variation points in the generic view model allows the user to dynamically activate or deactivate different perspectives on the process models and to directly interact with them.

1 Introduction

Process mining is a set of techniques to automatically analyze (business) processes using event data which is recorded during process execution and stored in so-called event logs. Most process mining approaches aim to discover a descriptive process model from an event log (process discovery). However, there is also a wide range of other techniques. Conformance checking, for example, compares an event log to a process model either to identify deviations of the process execution from a normative process model or to measure the fitness of a discovered model. Process enhancement aims to analyze additional data stored in the event log in order to annotate the process model with additional information (e.g., waiting times) to provide further perspectives.

The notion of multidimensional process mining (MPM) is an emerging concept that adopts the concepts of data cubes and OLAP [2] to the field of process mining [3] in order to analyze the processes from multiple perspectives. By using OLAP operators like roll-up and slice, it is possible to change the granularity of the event data or to filter it. One major goal of MPM is the comparison of different processes or variants process of the same process which are represented by the cells of the data cube.

There are different approaches for MPM described in literature like Process Cubes [4,1] and PMCube [6,5]. Both approaches use OLAP operations to partition the event data into subsets (sublogs) which are independently analyzed using arbitrary process mining algorithms. The main differences are the way the data cube is organized and the definition of OLAP operators. Furthermore, PMCube introduces also additional concepts to support the comparison of the extracted models (e.g., difference visualization). However, both approaches are similar in their general work-flow.

Due to its explorative character, interactivity is vital for MPM. However, tools for MPM are typically limited in their interactivity to a certain extent.

1. *Interaction with process models:* Even though process models are compositions of nodes and edges, they are statically visualized similar to an image. Direct interaction with process models (e.g., clicking on a node to trigger a filter) can make the analysis more intuitive.
2. *Dynamic analysis workflow:* Current tools for MPM force users to follow a certain workflow step by step. Changes to previously performed steps require the users to repeat the subsequent steps. Consequently, even minor adjustments of the OLAP query require a lot of effort for the analysts.
3. *Undo/redo of analysis steps:* Current tools for MPM do not provide undo/redo functionality. This may restrain users from exploring the processes because returning to a previous view on the process may be laborious.
4. *Performance:* Long processing times may disrupt the workflow. Therefore, performance is crucial for interactivity, even though MPM is not a time-critical application.

In this paper, we introduce Interactive PMCube Explorer, a tool which aims to provide seamless interactivity to MPM. In Section 2, we present the underlying concepts of our tool and in Section 3, we describe its implementation. Section 4 gives an overview of the tool demonstration.

2 Architecture

The implementation of our tool is based on the PMCube Explorer prototype [5]. Besides several improvements of the software architecture, it provides a newly designed user interface and workflow to improve the interactivity of MPM. Therefor, it manages the different analysis steps of the MPM workflow by a novel operator framework which defines a stack of operation levels. Each level represents particular analysis steps of the MPM workflow (e.g., OLAP query, process discovery, process enhancement). Interactions with user interface are mapped onto operations of the related level. The framework individually manages these operations for each level. This allows the user to undo/redo operations of a particular analysis step without affecting other levels. Also changes on lower levels are propagated to the levels above to automatically update the process mining results. This ensures that previously defined analysis steps do not need to be redefined after executing new operations on lower levels. E.g., the user does not need to configure and apply a process discovery algorithm again after changing the underlying OLAP query. For a more detailed description of the underlying concept, we refer to [7].

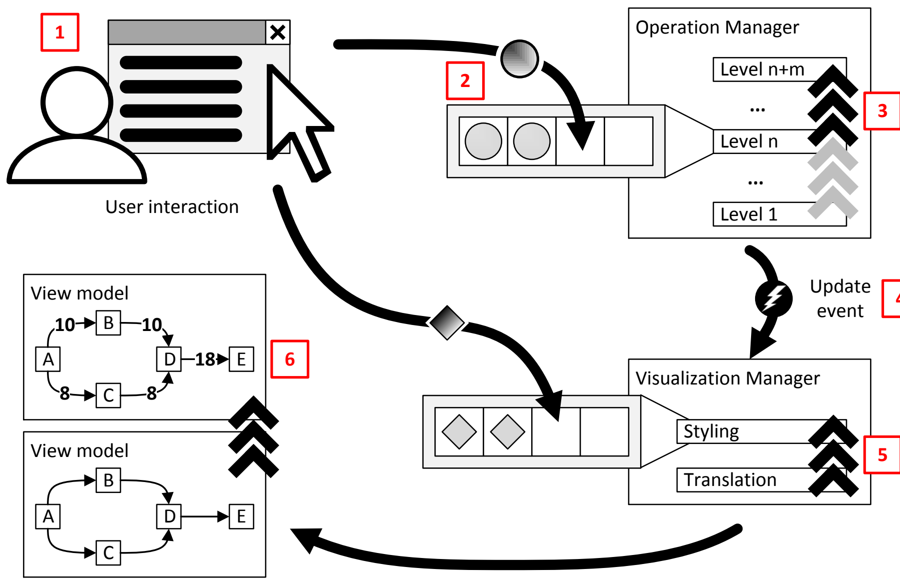

Fig. 1. General concept of the Interactive PMCube Explorer tool

Figure 1 shows the general concept of the operator framework. When the user interacts with the user interface, e.g. by applying a filter to the data (cf. item 1), the framework creates a new operation and adds it to the list of operations of the respective operation level (cf. item 2). Then, all operations of this and the subsequent levels are executed to propagate the changes (cf. item 3). After that, an update event triggers the visualization manager to update the currently displayed process models (cf. item 4). Therefor, the Visualization Manager creates a new view model and applies all activated styling operations to it (cf. item 5). Finally, the updated view model is shown in the user interface (cf. item 6).

3 Implementation

The Interactive PMCube Explorer tool is designed as a generic framework and is highly extensible. Most of its components like algorithms (e.g., for process discovery, conformance checking, process enhancement, consolidation, model difference calculation), process models, view models, database drivers, styling operations, and other operations like filters are integrated as plug-ins which are loaded during run-time. Currently, there are more than 70 plug-ins provided. Even though the tool is a prototypical implementation, it covers the entire MPM workflow. To show the feasibility of the operator framework, it provides several operations for almost every operation level. However, some features (e.g., interactive filters and styling) are currently only implemented as a proof of concept for particular process notations. Nonetheless, we plan to extend the tool by further plug-ins in the future. The tool and a manual containing a list of all plug-ins are available for download on our website[1].

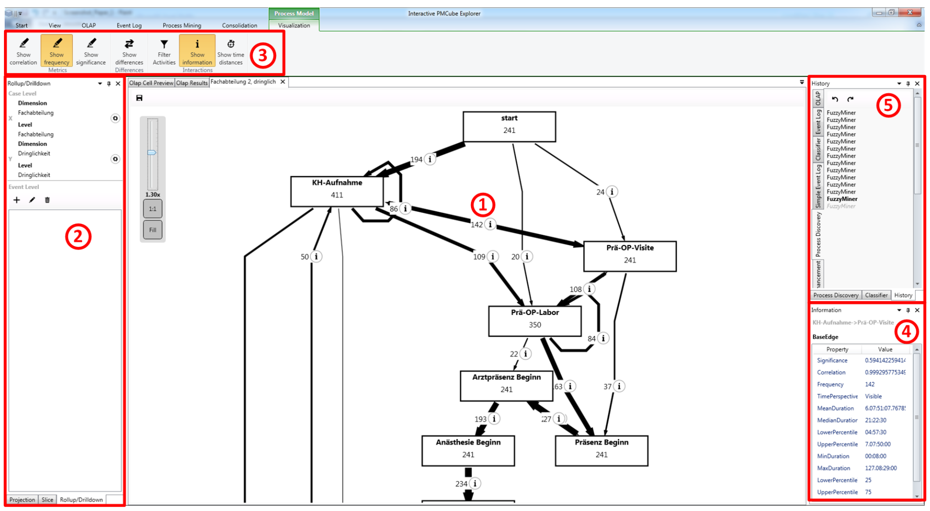

Fig. 2. Screenshot of the Interactive PMCube Explorer tool

Figure 2 shows a screenshot of the Interactive PMCube Explorer tool. The process mining results – which form the major subject of the analysis – are presented at the center of the application (cf. item 1). As MPM typically creates multiple sublogs and process models, the results are organized in tabs in order to easily switch between different results. The result overview (showing statistics about each cell), the process model matrix (presenting all mined models in a grid), and the preview for OLAP queries (indicating the resulting cells and the estimated data distribution) are also presented in tabs.

All other parts of the user interface (e.g., dialogs for the configuration of mining algorithms or OLAP queries) are arranged around the mining results. Using a docking system allows the user to customize the user interface (e.g. hiding parts or changing their location on screen). The example in Figure 2 shows the configuration of OLAP queries (cf. item 2) and the available visualization options (cf. item 3) for the currently shown process model (cf. item 1). Additional

[1] `http://uol.de/pmcubeexplorer`

information for the selected edge are shown in an additional view (cf. item 4). The history view (cf. item 5) shows the sequence of performed operations separately for each operation level. For each level, an undo/redo of operations is available.

4 Demo Scenario

In the demo scenario, we give a walk-through of an example analysis using the Interactive PMCube Explorer tool. A screen-cast of our demo is available on the web[2]. It demonstrates the key features of the tool with a special focus on its interactivity. The intention is to give an impression of the concepts like the operator framework, the change propagation, the visualization operations and the direct interaction with process models. Therefore, the OLAP queries and analysis steps of the demonstration are selected with focus on the different features. For the demo scenario, we use the data set of the BPI Challenge 2017[3].

As target audience, we are addressing researchers as well as practitioners with an interest in multidimensional process mining. Besides, this demo might also be interesting for the process mining community in general as the operator framework can also be incorporated in traditional process mining tools which may also benefit from a seamless and intuitive interaction with the user.

References

1. Alfredo Bolt and Wil M.P. van der Aalst. Multidimensional Process Mining Using Process Cubes. In K. Gaaloul et al., editors, *Enterprise, Business-Process and Information Systems Modeling*, volume 214 of *Lecture Notes in Business Information Processing*, pages 102–116. Springer International Publishing, 2015.
2. Matteo Golfarelli and Stefano Rizzi. *Data Warehouse Design: Modern Principles and Methodologies*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 2009.
3. IEEE Task Force on Process Mining. Process Mining Manifesto. In F. Daniel, K. Barkaoui, and S. Dustdar, editors, *Business Process Management Workshops (1)*, volume 99 of *LNBIP*, pages 169–194. Springer, 2011.
4. Wil M. P. van der Aalst. Process Cubes: Slicing, Dicing, Rolling Up and Drilling Down Event Data for Process Mining. In M. Song, M. Wynn, and J. Liu, editors, *Asia Pacific Business Process Management*, volume 159 of *Lecture Notes in Business Information Processing*, pages 1–22. Springer International Publishing, 2013.
5. Thomas Vogelgesang and H.-Jürgen Appelrath. Multidimensional Process Mining with PMCube Explorer. In F. Daniel and S. Zugal, editors, *BPM Demo 2015*, volume 1418 of *CEUR Workshop Proceedings*, pages 90–94. CEUR-WS.org, 2015.
6. Thomas Vogelgesang and H.-Jürgen Appelrath. PMCube: A Data-Warehouse-Based Approach for Multidimensional Process Mining. In M. Reichert and H. Reijers, editors, *BPM Workshops 2015*, volume 256 of *LNBIP*, pages 167–178. Springer, 2015.
7. Thomas Vogelgesang, Stefanie Rinderle-Ma, and H.-Jürgen Appelrath. A Framework for Interactive Multidimensional Process Mining. In M. Dumas and M. Fantinato, editors, *BPM Workshops 2016*, pages 23–35. Springer, 2017.

[2] `https://youtu.be/5ItNtByDetQ`
[3] `http://dx.doi.org/10.4121/uuid:5f3067df-f10b-45da-b98b-86ae4c7a310b`