Experimenting with an OLAP approach for interactive discovery in process mining

Gustavo Pizarro[1], Marcos Sepúlveda[1]

[1] Computer Science Department, School of Engineering,
Pontificia Universidad Católica de Chile, Santiago, Chile
e-mail:marcos@ing.puc.cl

Abstract. Business process analysts must face the task of analyzing, monitoring and promoting improvements to different business processes. Process mining has emerged as a useful tool for analyzing event logs that are registered by information systems. It allows the discovering of process models considering different perspectives (control-flow, organizational, time). However, currently they lack the ability to explore jointly and interactively the different perspectives, which hinder the understanding of what is happening in the organization. This article proposes a novel approach for interactive discovery aimed at providing process analysts with a tool that allow them to explore multiple perspectives at different levels of detail, which is inspired on OLAP interactive concepts. This approach was implemented as a ProM plug-in and tested in an experiment with real users. Its main advantages are the productivity and operability when performing process discovery.

Keywords: Process mining, business process discovery, OLAP.

1 Introduction

During the last decades, the concept of Business Process Management (BPM) has been gradually adopted by organizations worldwide. In many of these organizations, dedicated BPM areas have been created. The key role in these areas is the process analyst, who is responsible for modeling, monitoring and auditing business processes. More recently, the growth of the digital universe has made possible that these activities have an even greater reach, since information systems create event logs that store when the different processes' activities are performed [16, 18]. These logs allow performing a robust analysis about what is actually happening in the organizations through process mining. This discipline enables the automatic analysis of a process using various techniques, including process discovery, conformance checking and process enhancement [18]. Several of them are described in [16] and are sorted out by frequency of use in [5]. In process discovery, a model is created based on an event log without using any a-priori information; see [4, 16, 18, 19, 21], among others.

The current discovery tools allow creating either a control-flow diagram or a social network, but do not allow exploring interactively different process perspectives. We can infer from [2] that understanding the whole process requires an effort that is beyond the capabilities of traditional processes analysts, since the analysis of each per-

spective: control-flow, social or time, should be done separately. In addition, the process analyst must choose among a large number of discovery algorithms, each of them having a different representational bias and a set of parameters that must be configured. It is therefore very difficult to achieve a correct multi-perspective understanding of the process in an efficient way using existing tools.

This article proposes a novel approach for interactive discovery aimed at providing process analysts with a tool that allow them to explore a process in several perspectives at the same time, considering different levels of detail, and using dynamic and interactive filtering. Processes can be visualized considering different scenarios, allowing a comparative analysis on the control-flow, organizational or time perspectives. Considering the desired features, and the usability and quality expected, we concluded that such a concept of discovery must be implemented using an OLAP paradigm, providing navigation operations such as drill down, roll up, slice and dice, and pivot, allowing to explore the different perspectives: from groups performing similar tasks or groups that work together to single performers; from all possible paths to variants grouped in clusters; from traces that last little to traces that last long. Pivot views must be available showing control-flow diagrams or social networks that display the relationships between different performers. This approach was implemented as a plug-in in ProM [20] and tested in an experiment with real users.

2 Related work

In this section related work relevant to the development of this research is discussed.

2.1 The analyst and his interactive exploration task

Within the process area there are different roles. The most important are: process manager, process owner, process architect and process analyst [7]. Different skills and abilities are required to perform each of them, since they have different responsibilities. According to [1], the process analyst role is to support a holistic view of business processes and to offer the ability to quickly transform the organization. [13] notes that the process analyst should support the BPM life cycle that, among other activities, includes the analysis, design, monitoring and control of business processes. This requires demonstrating knowledge to audit processes, to perform gap analysis, and to support transformational changes in business environments.

In process mining, it has not been cataloged what kind of tool is most appropriate for each type of user. However, in the field of business intelligence, a broader field that includes process mining, different types of users have been identified. According to [8], there are two large categories of users: information consumers and information producers. Within the first group are business analysts and casual users. Casual users include executives, managers, employees and external users; they see reports regularly, but they do not calculate numbers or perform detailed trend analysis on a daily basis. Business analysts can also do more sophisticated analysis. Information consumers, in general, use dashboards, covering query and reporting tools, OLAP, spreadsheets, standard reports and output of statistical models. Within the information

producers are report authors that use statistical data mining tools, as well as some business analysts. Some power users can be both information producers and information consumers at the same time. An example of this mixed user is the business analyst. They typically use spreadsheets, pivot tables, simple database queries, and create custom reports. While there is more than one recommended tool for business analysts, most of them are related to dynamic information filtering. Considering the multidimensional approach required in this work and tools that are familiar to business analyst, we decided to implement the OLAP paradigm for processes analysis.

OLAP (Online Analytical Processing) is a solution aimed at speeding up queries to large amounts of data; multidimensional structures are used for this purpose. Some characteristic aspects of this approach are the multidimensional conceptual view, intuitive manipulation of data, and the use of unlimited dimensions and aggregation levels [6].The different dimensions can be explored using different operations, such as drill down, roll up, slice and dice, and pivot. The relevance of its application to process mining has already been mentioned by [15] and [22], but it is still an open challenge to define what the best way to do it is.

2.2 Process Discovery

In recent years there has been an evolution in the algorithms used for process discovery, but there has been little progress in this task from the process analyst viewpoint [18]. Since 2005, several algorithms have been proposed for control-flow discovery, such as Alpha Mining [19], Heuristic Mining [21] and Region Mining [4], among others. These algorithms have different approaches at the algorithmic level and contribute in different use cases. However, from the process analyst's viewpoint, they do not provide a task-oriented interactive exploration of the event log and do not provide a multidimensional understanding of the process. The first tool aimed at a more complex task than just getting a flow is Fuzzy Mining [10]. Despite being innovative, this proposal only addresses the control-flow perspective and requires process analysts to understand parameters that are algorithmic-centric rather than business-centric.

BPMNAnalysis [2] has as its primary motivations reducing the complexity of dealing with hundreds of algorithms in ProM and getting closer to the business language. It provides a multidimensional discovery by combining different existing tools, considering a holistic view of data, including variants, resources and time dimensions. It is a step forward towards a process analyst oriented discovery, but it does not cover the whole process analyst's viewpoint and it has not been tested on business users.

Recently, there have been a few multidimensional proposals covering the process through different dimensions and incorporating the OLAP concept at the algorithm level. This is the case of Multidimensional [15], which explores what tools are extensible to a multidimensional level, developing, for example, process discovery using HeuristicMiner, considering flows with roles or time. While this work was developed from the algorithmic viewpoint, it contributes to promote the multidimensional approach, incorporating conformance metrics to the multidimensional analysis, and an evaluation of the tool using ISO9126. Another recent proposal is the ProcessCube [12, 22], which allows the comparative analysis of different process models. It com-

pares different segments of the process (based on subsets of the event logs) or different organizational levels, incorporating a formal description to multidimensional analysis of the event log. These last two works do not focus on the day to day task of process analysts, as BPMNAnalysis and the present investigation do.

3 Proposal – OLAP Discovery

This paper develops a concept of interactive process discovery for process analysts, combining interactive exploration, dynamic filtering, navigation using OLAP operations, and automatic updates. It is also aimed at improving some aspects of usability with respect to existing tools, maintaining currently available discovery functionalities, and achieving a better understanding about what happens in organizations. The type of user (process analyst) and the task to be accomplished (interactive multi perspective discovery) were considered in the design. Already existing discovery tools were used at the algorithmic level.

The OLAP approach allows navigating the process in a similar way to a pivot table found in spreadsheets, exploring the different dimensions of a process, dynamic filtering of some dimensions leaving constant the others, going from a top level view to a detailed one, pivoting from a control-flow view to an organizational view, and comparing different versions. The proposed tool is called OLAPDiscovery, and it is integrated within the ProM framework. The dimensions considered are: social networks using Social Miner [17], variants and trace clustering using Trace Alignment [3], and time based on event timestamps. Results are displayed using the visualization diagrams of Heuristic Miner (control-flow view) and Social Miner (organizational view).

The process analyst must monitor and control processes to promote improvements and to create reports for business managers. Such users are familiar with the BPM approach and have knowledge about business goals, but they are not necessarily experts in data mining or process mining. Complementing this information with the classification of users in the Business Intelligence area, it can be state that in some cases these users are information consumers and in others are information producers. Improve usability for non-experts have been recognized as one of the ten most important challenges in process mining [18]. Furthermore, a recent study indicates that the major drawbacks of current tools are that they are unintuitive and difficult to understand [5]. The desirable discovery tool must allow changing the different variables that influence the process and see dynamically how those changes affect the control-flow and the organizational views.

The OLAP paradigm is useful to implement interactive discovery. The information about the execution of a process contained in an event log fits perfectly to be implemented in an OLAP cube with at least three dimensions: variants, resources and time:

Variant: A process variant is a unique start-to-end trace, a certain sequence of activities in which the process is executed, which has been recorded in the event log at least once. It may be found several times in the event log, representing cases involving different performers at different times. This dimension is composed by clusters of similar variants and it can be explored top-down, from the whole event log to a single

variant. The dimension is generated using the Trace Alignment plug-in [3]. Trace Alignment groups variants into clusters based on different criteria. It can be used to explore the process in early stages of analysis, and to answer specific questions in later stages. Therefore, it complements the existing process mining techniques that focus on process discovery and conformance checking. It allows studying the backbone of the process, the critical points that share the variants, to see the gaps between what really happened and what should have happened, to detect patterns and discover variants with minor differences compared to the desired behavior, among others.

Resource: In this dimension, we used the Social Miner plug-in, which allows grouping performers using different criteria. We considered three types of groups: people performing similar tasks, people working together, and handover of work. In addition, the user can change the threshold value that defines when two or more resources are grouped together. The user can also drill down or roll up through the different groups.

Time: A simple criterion was used in this dimension: the duration of each case. The user can filter the log between a minimum and a maximum duration.

Results can be displayed either in a control-flow view or an organizational view, being able to pivot between them at any time. For the control-flow view, the result of the Heuristic Miner was chosen because it describes the main observed behavior and can deal with noise [21]. The resulting view of the Social Miner plug-in is used for the organizational view since this view also allows discovering interactions between performers through visual inspection.

Initially the control-flow view is displayed (see Fig. 1). Buttons in the top right allows pivoting to the organizational view, switching from the comparison view to the single analysis (large) view, resetting the analysis, and saving the current view. In the middle right, a tab is displayed for each dimension (see Fig. 2); by clicking on each of them, the user can filter the event log. Fig. 3 shows the two discovery views: the control-flow view and the organizational (resources) view.

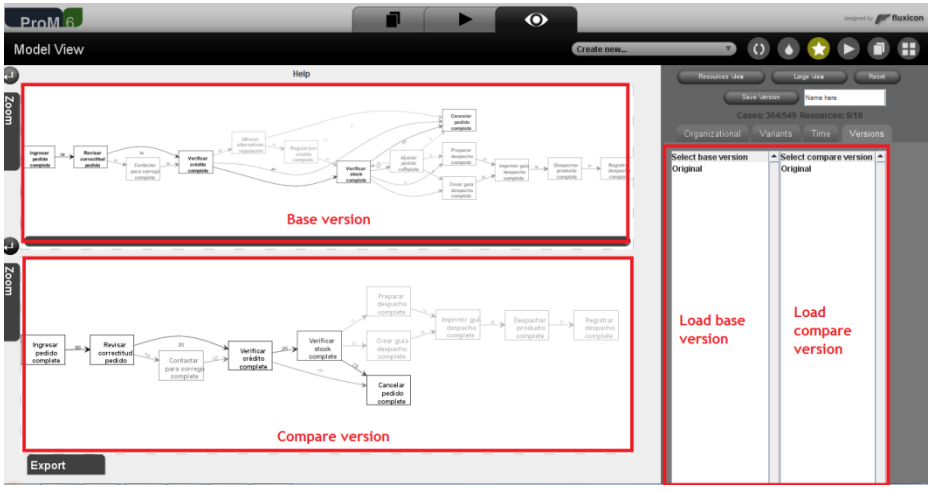

Fig. 1. OLAPDiscovery, comparison view and versions management

The comparison view was included to increase usability. It is possible to view the last view (after different OLAP operations have been applied to the original event log) for a single analysis or a comparison view, where a reference view (at the beginning the one corresponding to the whole event log) can be compared against the last view (see Fig. 1). The user can store a given analysis at any time and later load it for further analysis; it can be loaded in the upper part (as a reference for future comparisons) or at the bottom, for further filtering the recently loaded version.

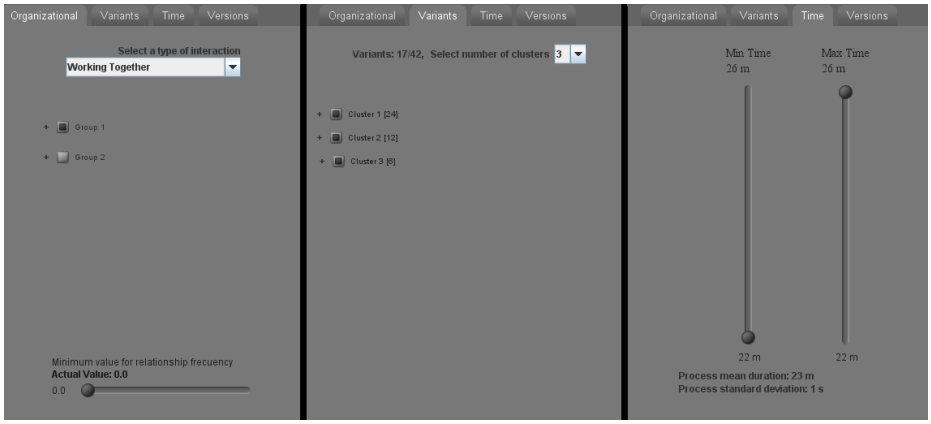

Fig. 2. OLAPDiscovery dimensions

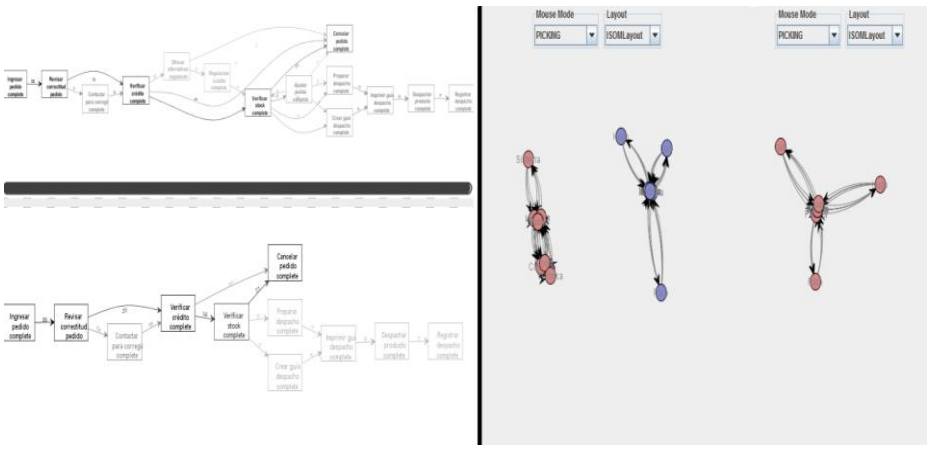

Fig. 3. OLAPDiscovery control-flow and organizational views

4 Experiment

This section describes the experiment, in which OLAPDiscovery, Disco [9] and ProM (without the new functionality provided by OLAPDiscovery) were tested with different users, to analyze what are the perceived advantages of OLAPDiscovery.

4.1 Participants

The sample was composed of 38 people: 16 experts and 22 basic users. Tests were carried out at different times, in groups of 1 to 6 users. Users were students in the last two years from Information Systems or related degrees. We considered as basic users those who have completed a full semester business process modeling course, where student learn BPM concepts and ProM is used in a basic way. On the other hand, we considered as expert users those who have completed a full semester process mining course, where students learn process mining exhaustively based on [16].

4.2 Tasks performed by the users

The tasks performed by the users required them to answer two sections of four questions each. The first section was about unidimensional analysis, one for each dimension of the process behavior. Some questions were statements about the process, and the user had to determine if they were true or false; the rest were multiple-choice questions about basic characteristics of the process. The second section consisted of tasks that combined different dimensions. In this case, users had to answer whether some statements were true or false. For answering the statements, users were required to analyze whether in a given scenario (a group of performers executing some variants in a certain time interval) the process behaves or not in a certain way. The user had to complete each unidimensional task in a maximum of 5 minutes and each multidimensional task in a maximum of 12 minutes. To reduce learning curve effects, each user received an accurate training about how to perform each task before answering any questions. In addition, each question had a companion explanation that described step by step how to perform the analysis. Every user had to answer three different tests (one for each tool) with eight questions each, based on three different event logs. There was no time limit, but they could finish the task if the recommended time was exceeded. Each tool was evaluated in a different order by each user to evaluate whether using one of the tools before the others might influence the results.

4.3 Description of the evaluation

We used a simple software quality assessment model based on ISO9126 [11], a. standard that has already been used for evaluating process mining tools [15]. We combined ISO9126-3 and ISO9126-4 techniques [11]. The following attributes from the functionality and usability features of ISO9126-3 were considered: suitability, understandability, learnability and operability. Other attributes were excluded, because they were less relevant to the approach of our tool. From ISO9126-4, effectiveness, productivity and satisfaction were considered; security was left out because it is not the focus of this research.

Both objective and subjective metrics were considered for measuring software quality. Table 1 shows the objective metrics, which are based on users' performance answering all questions. Table 4 shows the subjective metrics, which measure users' perception after completing all activities. A brainstorming was performed to find out the attributes we wanted to measure, creating an initial list of 80 questions. Based on

the judgment of BPM experts, we selected the 16 most representative. We did not want an extensive questionnaire, since users were required to answer it for each tool.

Table 1. Objective metrics

Attributes	Metrics
Completed tasks	Tasks completed / Total tasks
Effectiveness	Tasks completed successfully / Total tasks
Unidimensional time	Time needed to complete unidimensional tasks.
Multidimensional time	Time needed to complete multidimensional tasks.
Productivity	Effectiveness/Time needed to complete tasks

4.4 Results

In this section, we first focus on some objective metrics: effectiveness and productivity, and then we discuss the results for subjective metrics.

Table 2 shows the average number of unidimensional and multidimensional tasks that were correctly completed by the users using the different tools (note that OLAPDiscovery is abbreviated OLAP-D in all figures and tables).

All tools allow users to complete unidimensional tasks. The number of correct answers in ProM and OLAP-D is similar, showing both tools are well suited to answer unidimensional requirements. With Disco more incorrect answers are obtained because this tool has some limitations for analyzing resources. Users could not complete successfully all multidimensional tasks in ProM, showing the need for a tool that can handle multi-perspective tasks. In average, users could not complete or answer wrongly 1 out of 4 tasks in ProM; most of them basic users. On the other hand, only one user could not complete a multidimensional task using OLAP-D; and two, using Disco. These multidimensional tasks are more challenging and therefore the number of incorrect answers is higher for all tools.

Table 2. Average completed unidimensional and multidimensional tasks.

	Completed unidimensional tasks		Completed multidimensional tasks		
	Correct answers	Incorrect answers	Correct answers	Incorrect answers	Incomplete answers
ProM	4,82	0,18	2,95	0,55	0,50
Disco	4,58	0,42	3,34	0,63	0,03
OLAP-D	4,84	0,16	3,29	0,68	0,03

Not only completing tasks is relevant, but also the time it takes. Table 3 shows the time required to perform unidimensional tasks, multidimensional tasks, and all tasks for the three tools, considering the order in which they were tested by the users; first, second or third. It can be seen the time required to complete both unidimensional and multidimensional tasks is greater in OLAP-D than in Disco and in ProM. The best

case in Disco is not less than the slowest case in OLAP-D. Moreover, the best case in ProM requires more than twice the time required in the slowest case in OLAP-D.

Table 3. Time required for performing different tasks, according to the order of testing.

Tool	Order on the testing	Time in uni-dimensional tasks (minutes)	Time in multi-dimensional tasks (minutes)	Total time (minutes)
OLAP-D	1	4.5	7.8	12.3
OLAP-D	2	2.9	5.9	8.8
OLAP-D	3	2.4	6.4	8.8
Disco	1	5.9	9.6	15.6
Disco	2	4.3	9.4	13.7
Disco	3	4.2	9.5	13.7
ProM	1	12.5	18.5	31.0
ProM	2	11.2	18.0	29.1
ProM	3	10.0	15.9	25.9

Fig. 4 shows the productivity achieved by the users with the different tools. It displays the relationship between effectiveness (tasks completed correctly / total tasks) and the total time required for performing all tasks; it also considers the order in which they were tested by the users. It can be seen that the productivity achieved with OLAP-D is greater; the tasks were performed with greater effectiveness in less time. This figure also shows that with the current sample of users, it is not possible to conclude that the learning curve is smaller with any of the three tools.

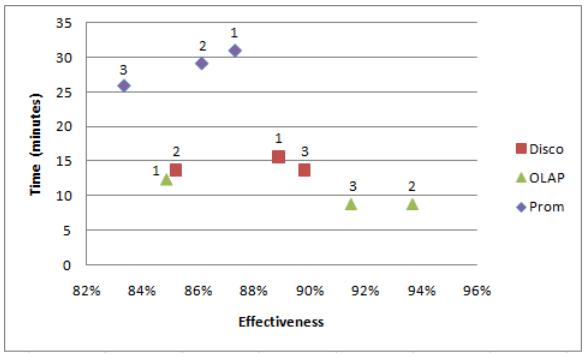

Fig. 4. Productivity represented as the relationship between effectiveness and total time.

Table 4 summarizes the results of the subjective comparison between OLAP-D and the other two tools, ProM and Disco. It is shown the non-parametric Mann-Whitney U test that compares two samples to evaluate if they are the same; both the Mann-Whitney U test and the corresponding significance are displayed. The Mann-Whitney U test [14] was used since distributions were not normal; this method is widely used to study ordinal variables such as Likert scale statements.

Table 4. Subjective metrics - statistic tests for comparing the different tools

Attribute - Metric	*OLAP-D vs Disco*		*OLAP-D vs ProM*	
	Mann–Whitney U	Significance	Mann–Whitney U	Significance
Effectiveness - I can understand a process as a whole (with all the characteristics that distinguish it).	611,0	0,214	357,0	*< 0,001*
Effectiveness - I can understand a process accurately (and clearly).	682,5	0,666	476,5	*0,001*
Productivity - The steps required to perform an analysis are adequate.	636,0	0,315	225,5	*< 0,001*
Productivity - The time required to perform an analysis is reasonable.	511,0	*0,017*	110,0	*< 0,001*
Satisfaction - I would recommend this software to my colleagues.	635,5	0,323	217,5	*< 0,001*
Satisfaction - I would use this software to analyze processes.	630,5	0,286	268,5	*< 0,001*
Suitability - The available parameters are suitable.	689,0	0,700	481,5	*0,007*
Suitability - The software is useful for the analysis of my processes.	689,0	0,710	470,5	*0,006*
Suitability - The organization of menus is logical.	705,0	0,851	354,0	*< 0,001*
Understandability - The features of this software are easy to understand.	639,0	0,348	157,0	*< 0,001*
Understandability - The information generated is easy to understand.	658,0	0,453	316,0	*< 0,001*
Learnability - It is easy to remember how to do things.	555,0	0,050	250,0	*< 0,001*
Learnability - The tool is simple to use.	647,5	0,390	124,0	*< 0,001*
Operability - It is easy to act based on the information generated.	522,5	0,022	272,0	*< 0,001*
Operability - It is easy to combine an analysis of all aspects of a process.	409,0	*0,001*	174,0	*< 0,001*
Operability - It is easy to move from one feature to another one.	502,0	*0,013*	273,5	*< 0,001*

When comparing OLAP-D and ProM, the difference in most statements is quite broad. OLAP-D has their mode values between “strongly agree” and “agree” for all statements while ProM has their mode values in “agree” in only seven cases. For all metrics, the significance of the Mann-Whitney U test is less than 0.05, showing OLAP -D is different (better) in all statements compared to ProM.

When comparing Disco and OLAP-D, quite similar results were obtained. However, in some statements OLAP-D was considered to be better than Disco. It is the case of the statement “The time used for the analysis is reasonable” (Productivity) and all statements about Operability. Therefore, according to the users’ viewpoint, OLAP-D is as good as Disco, but it is considered to be more productive (which is consistent with the objective metrics discussed above) and having a higher operability.

When results are split by type of users (basics vs. experts), the results show expert users appreciate ProM better than basic users, but still like Disco and OLAP-D better. Expert users evaluate OLAP-D better in Operability and Productivity compared to Disco. On the other hand, basis users do not see any significant difference between OLAP-D and Disco. This might be explained because while OLAP-D offers more advanced functionalities, Disco is friendlier for basic users.

4.5 Experiment limitations

There are two main limitations of the experiment. First, the tasks designed to measure multidimensional analysis only cover the most common scenarios considered in the development of OLAPDiscovery; ProM and Disco provide additional features that are useful in other scenarios. Second, the size of the sample of users is small.

5 Conclusions

This article introduces OLAPDiscovery, a novel approach for interactive process discovery aimed at providing process analysts with a tool that allow them to explore process perspectives at different levels of detail. This approach was implemented as a ProM plug-in and it was tested in an experiment with real users.

The experiment shows that the dynamic, multi-perspective and interactive discovery provided by OLAPDiscovery offers a better productivity, measures as the relationship between effectiveness and time, in both unidimensional and multidimensional tasks compared to the usage of ProM or Disco. The key is the dynamic interactivity offered to the user, and the ability to filter and automatically get results both in the control-flow and the organizational perspectives. The analysis of subjective metrics allows us to conclude that users think this multi-perspective interactive process discovery is more operational and productive than the one provide with existing tools. It is interesting that for basic users OLAP-D is as good as Disco because both tools have different strengths; while Disco provides a nicer user interface, OLAP-D provides features that allow them to solve multidimensional tasks in a simpler way. The aim of this research is better reflected in expert users; they think OLAP-D provides a better productivity and operability compared to Disco.

This research shows that it is relevant to assess the functionality, usability and quality in use of any process mining tools from the users' viewpoint, since not only the algorithm is relevant, but also how to provide a satisfactory experience for the users.

In the future we visualize several research opportunities. First, we have used specific control-flow and organizational algorithms, but it could be possible to enable users to select their own favorite discovery algorithms. Second, we have only explored some process perspectives; a promising and straight-forward enhancement is to extend the OLAP approach to other data perspectives. Third, the OLAP approach could be extended for conformance analysis, allowing users to verify conformance in a subset of the original event log against a subset of the reference model, proving a better understanding of eventual discrepancies between the observed and the expected behavior.

References

1. Bandara, W., Chand, D. R., Chircu, A. M., Hintringer, S., Karagiannis, D., Recker, J. C., van Rensburg, A., Usoff, C., & Welke, R. J. (2010). Business process management education in academia: Status, challenges, and recommendations. Communications of the Association for Information Systems, 27, 743-776.
2. Bayraktar, İ. (2011). The Business Value of Process Mining Bringing It All Together, Eindhoven university of technology, Eindhoven University of Technology, Eindhoven.
3. Bose, R. J. C., & van der Aalst, W. (2010). Trace alignment in process mining: opportunities for process diagnostics. In Business Process Management, 227-242.
4. Carmona, J., Cortadella, J., & Kishinevsky, M. (2008). A region-based algorithm for discovering Petri nets from event logs. In Business Process Management, 358-373.
5. Claes, J., & Poels, G. (2013). Process mining and the ProM framework: an exploratory survey. In Business Process Management Workshops, 187-198.
6. Codd, E. F., Codd, S. B., & Salley, C. T. (1993). Providing OLAP (on-line analytical processing) to user-analysts: An IT mandate. Codd and Date, 32.
7. Doebeli, G., Fisher, R., Gapp, R., & Sanzogni, L. (2011). Using BPM governance to align systems and practice. Business Process Management Journal, 17(2), 184-202.
8. Eckerson, W. W. (2010). Performance dashboards: measuring, monitoring, and managing your business. Wiley.
9. Fluxicon Process Laboratories, Inc. [Download]: Disco version 1.5.
10. Günther, C. W., & Van Der Aalst, W. M. (2007). Fuzzy mining–adaptive process simplification based on multi-perspective metrics. In Business Process Management, 328-343.
11. International Organization for Standardization (2001). ISO 9126: Software Engineering – Product quality,Geneva, Switzerland.
12. Mamaliga, T. (2013). Realizing a Process Cube Allowing for the Comparison of Event Data (Master's thesis, Eindhoven University of Technology, Eindhoven).
13. Mathiesen, P., Bandara, W., Delavari, H., Harmon, P., & Brennan, K. (2011). A comparative analysis of business analysis (BA) and business process management (BPM) capabilities. ECIS 2011 Proceedings.
14. Newbold, P., Carlson, W., Thorne, B. (2008). Statistics for Business and Economics. Pearson.
15. Ribeiro, J.T.S. (2013), Multidimensional process discovery, Eindhoven University of Technology, Eindhoven.
16. Van der Aalst, W. M. (2011). Process mining: discovery, conformance and enhancement of business processes. Springer.
17. Van der Aalst, W. M., Reijers, H. A., & Song, M. (2005). Discovering social networks from event logs. Computer Supported Cooperative Work (CSCW),14(6), 549-593.
18. Van der Aalst, W., Adriansyah, A., de Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., ... & Pontieri, L. (2012). Process mining manifesto. In Business process management workshops, 169-194.
19. Van der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering, 16(9), 1128-1142.
20. Van Dongen, B. F., de Medeiros, A. K. A., Verbeek, H. M. W., Weijters, A. J. M. M., & van der Aalst, W. M. (2005). The ProM framework: A new era in process mining tool support. In Applications and Theory of Petri Nets 2005, 444-454.
21. Weijters, A. J. M. M., van der Aalst, W. M., & De Medeiros, A. A. (2006). Process mining with the heuristics miner-algorithm. Technische Universiteit Eindhoven, Tech. Rep. 166.
22. Van der Aalst, W. M. (2013). Process Cubes: Slicing, Dicing, Rolling Up and Drilling Down Event Data for Process Mining. In AsiaPacific Business Process Management, 1-22.