A Study of Different Visualizations for Visualizing Differences in Process Models

Manuel Gall[1], Günter Wallner[2], Simone Kriglstein[3], Stefanie Rinderle-Ma[1]

[1] University of Vienna, Faculty of Computer Science, Austria
{manuel.gall, stefanie.rinderle-ma}@univie.ac.at
[2] University of Applied Arts Vienna, Institute Art & Technology, Austria
guenter.wallner@uni-ak.ac.at
[3] Technical University of Vienna, Institute for Design & Assessment of Technology, Austria
simone.kriglstein@tuwien.ac.at

Abstract. Finding differences between two processes can be a complex, time consuming, and expensive task. Our work is based on the difference graph approach which calculates the differences between two process models and – if available – their instances. In this paper we evaluate different possibilities for visualizing these differences. For this purpose we have selected some common visual properties such as color, shape, and size and evaluated these different visualizations with 31 participants through an online survey. Our results show that color coding and symbols were the preferred methods of the participants for depicting differences in a graph visualization.

Keywords: Process Differences, Difference Graph, Visualization, Process Model, Instance Flow

1 Introduction

Processes are an indispensable part of today's businesses. Be it for visualizing processes for communication or optimization purposes, companies constantly use processes to gain additional business intelligence. Using process mining algorithms on data collection during process execution allows, for example, detection of bottlenecks, problems, and violations (cf. [2]). Conformance checking [18] is a process mining technique which assesses if an event log of a process deviates from the process model. Unfortunately conformance checking focuses only on detection of deviations between event logs and process models. However, the analysis of differences and commonalities between process models and, optionally, their

Published in ER 2015 Workshops Proceedings, pp 99-108, 2015
Springer link: http://link.springer.com/chapter/10.1007/978-3-319-25747-1_10

instance traffic – which shows how the instances have progressed through the model – are also of interest for different use cases in business process management (cf. [3]). This includes finding deviations between two process models which have been generated through process mining techniques. Consider, for example, a company which executes the same task at two different locations. Unfortunately in one location the execution takes twice as long. Comparing the process models of these locations can reveal why the execution takes longer. When comparing two process models one can also gain additional information if these processes can be merged. Finding deviations allows an analyst to determine where problems may occur. Another example is to assess one process at different points in time, for instance, to evaluate how one process has evolved from one year to the next. In addition to the analysis of process models, Kriglstein et al. [13] point out that the analysis of instances and how they have progressed through the model (i.e., instance traffic) based on execution logs or simulation data can give interesting insights into the distribution of instances between different process models.

In this regard, visualizing data is a very important task to enhance the users understanding of the data. For example, Kriglstein et al. [13] presented a visualization approach to visualize differences and commonalities between process models and their instances by means of color coding. The input models themselves can either be generated through process mining or manually. However, the approach presented by Kriglstein et al. [13] is lacking in regard to the evaluation in order to identify how to best visualize the differences to support an effective interpretation of them. In this paper, we address this issue by investigating which visualizations of differences suit the interpretation of differences in process models and instance traffic best. To answer this question we evaluated nine different visualizations which are based on different visual properties (gathered through a literature review) with company employees and students via an online survey.

The remainder of this paper is structured as follows. Related work is shortly discussed in the next section. Section 3 then gives an overview of the difference graph model and its calculation. This is followed by a short introduction of the investigated visualizations in Section 4. The evaluation itself is then covered in Section 5. The limitations of the study are discussed in Section 6. Section 7 concludes the paper and shortly discusses possible directions for future work.

2 Related Work

To represent differences in a graph, various approaches like difference map (e.g., [6]), animation (e.g., [8]), and small multiples (e.g., [4]) were developed in the last years. With regard to business processes, there exist several approaches that focus on the analysis of differences and similarities between process models (see, e.g., [1, 12–14, 19, 20]). Often changes are directly visualized within the process model via color coding, for example, by coloring new activities in green or activities that were removed in the newer version in red (cf. [13]). However, evaluations about how suitable the suggested approaches for visualizing differences within process models are, are often missing. An example for an evaluation is the user study from Kabicher et al. [12] in which the authors evaluated the visual properties color, brightness, and size for the change operations *add* and *delete* (which correspond to our markings NEW and DELETED) in order to identify which visual property the participants preferred for visualizing changes in process models. In contrast to our study the visual properties were preassigned to the two change operations. For example, they used orange to represent deleted elements and green to highlight added elements. Furthermore, we consider six additional visualizations (e.g., shapes, symbols) and three further markings (UNCHANGED, INCREASED, and DECREASED).

Overloading existing process modeling languages leads to a more complex visualization which, in turn, can lead to an increase in cognitive load. For example, Moody and Hillegersberg [16] investigated the cognitive effectiveness of UML based on five principles. Similarly, Genon et al. [10] investigated the cognitive effectiveness of the visual notation of BPMN 2.0. Both studies showed that the cognitive effectiveness can be improved within those languages and agree that involving the users within the development of these languages is a key aspect.

3 Difference Graph Model

The difference graph concept presented by Kriglstein et al. [13] consists of two parts: difference model and instance traffic. A process model is considered to be a directed connected graph consisting of a set of nodes and a set of edges. Optionally, nodes and edges may have weights assigned which represent the instance traffic. Instance traffic measures how often a specific activity has been executed. As an example, Figure 1 shows two versions of a process model with their instance traffic. When visually comparing these two variants one can observe that, for example, B has been deleted from the right model. However, with increasing size of the model manually finding the differences and commonalities can become increasingly cumbersome and time-consuming. For this purpose the difference model was introduced by Kriglstein et al. [13]. The difference model is calculated by subtracting two process models, in our example referred to as PM_1 and PM_2. Calculating $PM_2 - PM_1$ results in a difference model with markings associated with its nodes and edges (for an in-depth discussion see [13]). The following list shows the different markings and describes in which case these markings are assigned:

NEW: a node/edge is marked as NEW if the node/edge was added to PM_2.

UNCHANGED: a node/edge is marked as UNCHANGED if it appears in both input models (and has the same weights in case instance traffic is available).

DELETED: a node/edge is marked as DELETED if the node/edge was removed from PM_2.

In case instance traffic is available, two further markings exist:

INCREASED: a node/edge is marked as INCREASED when the weight has increased from PM_1 to PM_2.

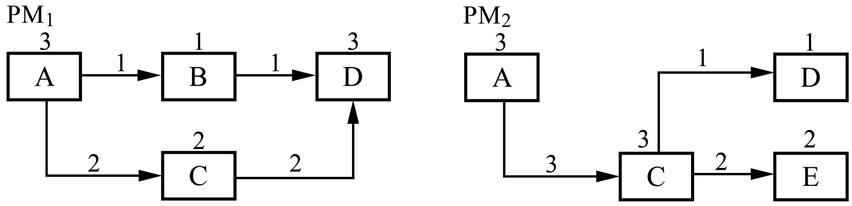

Fig. 1. Two variants of a process model with their instance traffic.

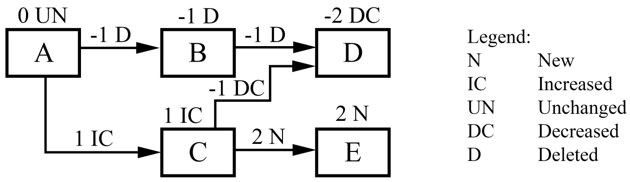

Fig. 2. Calculating the differences between the process models and their instance traffic from Figure 1 ($PM_2 - PM_1$) leads to this difference graph. Above each node and edge the calculated weights and assigned markings are shown.

Decreased: a node/edge is marked as Decreased when the weight has decreased in PM_2.

Figure 2 shows the resulting graph when the left model in Figure 1 is subtracted from the right model. For example, Node B was deleted from PM_2 and is therefore marked as deleted while the weight of node C has increased from 2 to 3 and thus C has been marked as Increased.

4 Visualizations

The main focus of this paper is how to best visualize such a difference graph in order to promote an effective interpretation by users. For this purpose we used different visual properties (see, e.g., [5, 7, 11, 15] to mention but a few) for the visual encoding of the five markings: New, Unchanged, Deleted, Increased, and Decreased.

In order to identify relevant visual properties which can be used for the different markings we conducted a literature review to analyze different visualizations that were used in the last years to depict differences between process models. In summary 31 papers were found. Collected literature as well as categorization can be found on our website [9]. We analyzed these papers to identify different potential visualizations and their corresponding visual properties. Based on these findings nine different visualizations (summarized in Figure 3) were created based on different visual properties. Each visualization depicted the markings in a different way, for example, through different colors, shapes, or symbols. For example, the visualization *Color* uses the colors green, blue, black, orange, and red to distinguish between the different markings.

5 Evaluation

The goal of the evaluation was to identify which of the nine visualizations (cf. Figure 3) – *Brightness*, *Font Size*, *Line Width*, *Edge Pattern*, *Symbols*, *Background Color*, *Color*, *Edge Ending*, and *Node Shape* – gathered through our literature review suit the interpretation of differences in process models and their instance traffic best.

Design. For the evaluation we used an online questionnaire in order to be more flexible and to reach a broader community than it would be possible with face-to-face interviews (e.g., to attract not only local participants). Each question was developed in accordance with the *Ten Commandments* by Porst [17]. The questionnaire was divided into three groups:

Introduction: The concept of the difference graph as well as the meaning of the weighted edges were introduced by means of an example. Furthermore, participants were asked if they already have experience with graphs.

Visualizations: This group included questions regarding the nine different visualizations. Each of the visualizations was presented on a single page which contained three questions related to the visualization. One question required the participants to rate the expressiveness of the visualization on a 5 point rating scale from *very good* (5) to *poor* (1). The second question asked the participants to assign each visual element to the marking for which they think the element is best suited for. An additional open-ended question allowed participants to share their opinions about the visualization type.

Final and Demographic Questions: We asked the participants to rank the visualizations according to their expressiveness. In addition, participants were asked if they prefer to use the same encoding for nodes and edges or not. In addition to demographic questions (e.g., about age, gender, and employment), the participants were also asked if they have already worked with business processes or not.

A two-stage pretest to validate the design of the survey was conducted to ensure the questionnaire is understandable and the time required to complete the survey is adequate. In the first pretest a discussion with two people took place to asses the wording of the questionnaire. After changing the survey according to the user feedback a second pretest with five participants was conducted. Participants were encouraged to ask questions and to provide feedback for possible improvements. All their questions and comments were noted and analyzed and based on their feedback the questionnaire was improved.

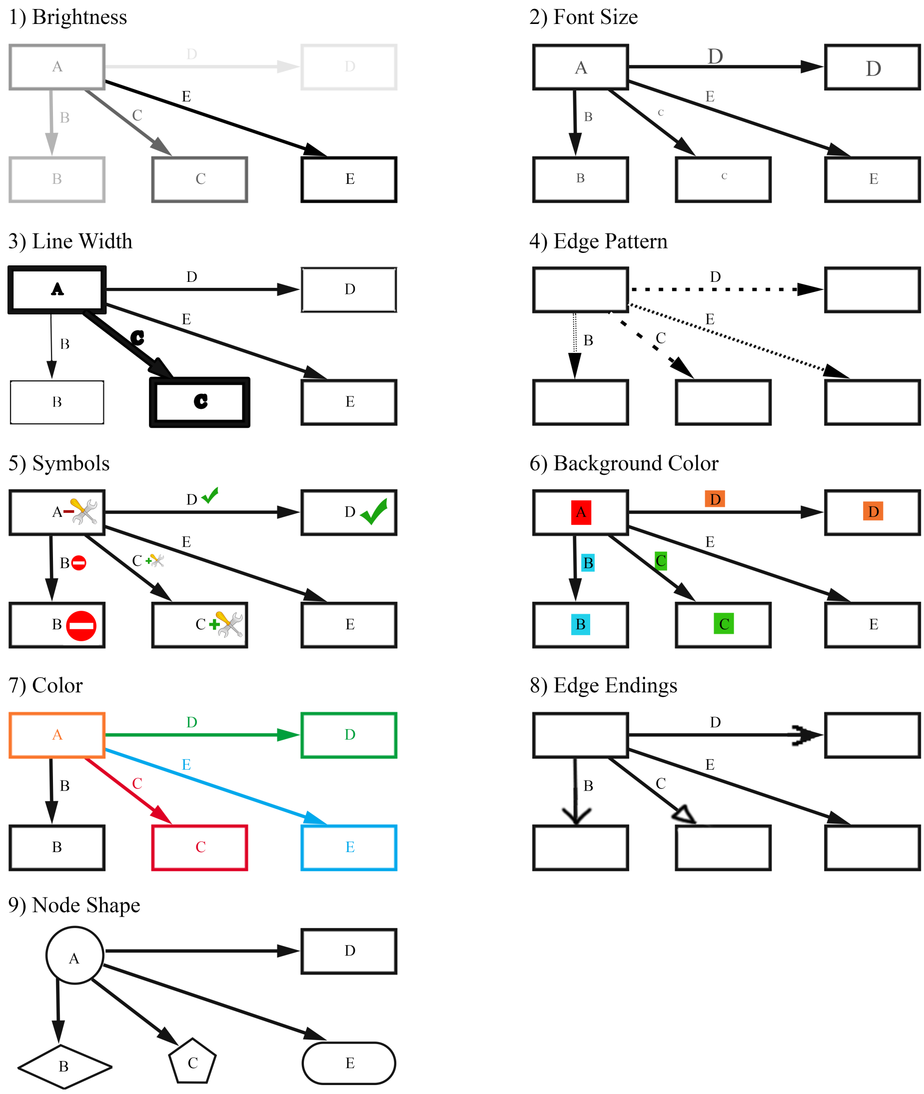

Fig. 3. The nine visualizations to present the difference graph of two process models and their instance traffic. The design of the visualizations is based on the findings from our literature review.

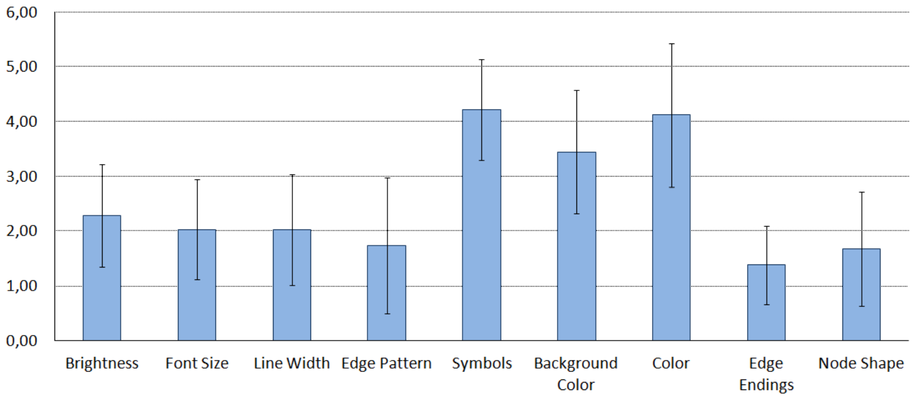

Fig. 4. Average expressiveness rating across all participants for each of the nine visualizations (error bars represent standard deviation).

Sample. The survey was run from June to August 2014. To attract participants, the link to the survey was distributed by e-mail to company employees and students. Furthermore, the link was posted on a private Facebook page initiated from and addressing business informatics students. In the end, we received 31 complete responses. Of the 31 participants, 24 were male (77.4%), 6 were female (19.4%) and one participant did not specify the gender. The participants were aged between 19 and 46 years ($M = 28.25$, $SD = 7.43$). Three participants did not report their age. 19 participants (61.3%) worked with graphs before and 12 participants (38.7%) did not. With regard to process modeling, 13 participants (41.9%) had already worked with process models while 17 (54.8%) had not. One participant did not answer the question.

5.1 Results and Discussion

Figure 4 shows the average expressiveness rating (5 is best) across all participants for the nine evaluated visualization types. We then assessed if experience with graphs and process models influenced the rating of the expressiveness of the different visualizations using two separate one-way MANOVAs. In the former case, no statistically significant difference in the rating of the visualizations could be observed ($F(9, 21) = 2.224, p = .063$). However, there was a statistically significant difference in the rating of *Font Size* ($F(1, 28) = 10.845, p = .003$) and *Line Width* ($F(1, 28) = 5.449, p = .027$) if prior experience with process models was taken into account. Participants who worked with process models before rated these two visualizations considerably lower than people who did not. Specifically, $M = 1.46, SD = .66$ compared to $M = 2.35, SD = .79$ in case of *Font Size* and $M = 1.54, SD = .66$ compared to $M = 2.35, SD = 1.11$ for the *Line Width* visualization.

After the participants had seen all nine visualizations they had to rank them from best (9) to worst (1). Analysis of the ranking with a Friedman test revealed a rank-ordered preference for the visualizations, $\chi^2(8) = 155.84, p < .001$, with *Color* being the most highly ranked visualization with a mean rank of 7.97, followed by *Symbols* (7.50) and *Background Color* (6.87). *Node Shape*, *Edge Endings*, and *Edge Pattern* where the lowest ranked with a mean rank of 2.65, 2.58, and 2.39 respectively. *Brightness*, *Line Width*, and *Font Size* received moderate mean rankings of 5.77, 4.79, and 4.48. In general this ranking reflects the average expressiveness rating of the individual visualizations. Due to space limitations we will focus on the three top and lowest ranked visualizations in the following. Post hoc analysis with Wilcoxon signed-rank tests and a Bonferroni corrected p-value of 0.0014 between all pairs of visualizations showed no significant differences in the ranking of the *Color* and *Symbols* visualization ($Z = -1.66, p = .098$) as well as between the *Background Color* and *Symbols* visualization ($Z = -1.77, p = .077$). Differences in the ranking of the three lowest ranked visualizations were also statistically insignificant. It is noticeable that visualizations which only depicted changes of either nodes or edges were ranked lowest. This is, however, in line with the participants preference (77.4%) to use the same visualization of changes for both, nodes and edges.

To assess how much participants agree on which visual element should represent which marking we used Krippendorffs' alpha (α) as a measure of inter-rater agreement. The highest agreement was found for the *Symbols* visualization with $\alpha = .742$ followed by *Color* and *Background Color* with almost identical values of .575 and .57 respectively. The similar result between *Color* and *Background Color* is not surprising as the same colors have been used in both visualizations. Again, *Node Shape* ($\alpha = .095$), *Edge Endings* ($\alpha = .053$), and *Edge Pattern* ($\alpha = .001$) scored lowest. Of course, these results are influenced by the actual choice of, for example, symbols and colors. A different set of colors or symbols is likely to have led to a different outcome. However, it also shows that our set of symbols was quite well chosen to represent the five markings. The poor performance of the latter three visualizations is also in line with the qualitative feedback by some of the participants who considered the edge endings to be *not suitable* or *not very meaningful and hard to discern in large graphs* (2 participants). Similarly, line patterns were also perceived as *not meaningful* (1 participant).

In summary these results suggest that symbolic visualizations of the markings as well as color coding of edges and nodes are best suited to visualize differences in process models. However, further evaluations should assess which encodings should be used and how these visualizations scale with increasing size of the process models. We suspect, that in large process models it might be challenging to relate symbols to their corresponding edges. Figure 5 shows the example from Figure 2 by using color coding or symbols.

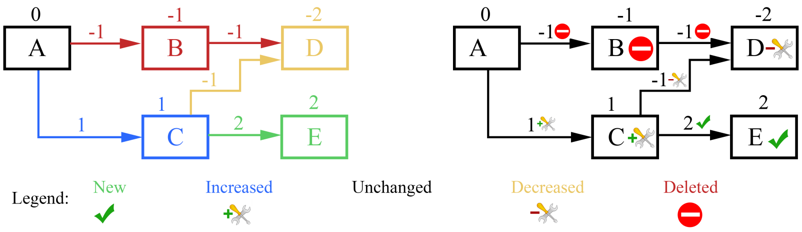

Fig. 5. *Color* and *Symbols* visualization of our example from Figure 2. Encodings of markings have been chosen according to the participants preference.

6 Limitations

It should be noted that the survey only investigated one specific encoding scheme (e.g., color scheme, set of shapes) for each visualization. The concrete choices of symbols or colors, however, may have an important influence on the result as pointed out above. The results of the survey should thus be rather viewed as an indication of which encodings are promising and which should be investigated in more detail in the future. In addition, the survey mainly assessed the expressiveness and the participants' preferences of the investigated visualizations. Subsequent studies should thus also assess these visualizations in terms of correctness of interpretation. Such studies may also take into account the participant's background in certain business process notations and how these background influences the perception of the utilized difference encodings which has been outside the scope of this study.

7 Conclusion

In this paper we investigated nine different visualizations (based on different visual properties) in order to assess which visualization suits the interpretation of differences in process models best. The findings of our user study show that color coding or symbolic visualizations are very promising for visualizing differences between two process models and their instance traffic. In order to support the users' intuitive understanding we suggest to use a legend describing the encodings. The results presented here have contributed to the implementation of a prototype implementation of the difference graph concept utilizing color coding and symbolic encoding for the ProM framework. An installation guide for the plug-in as well as survey results are available on our website [9].

A possible direction for future work is to conduct further studies to find specific colors and symbols for encoding and to investigate if overloading of nodes and edges is reasonable for different process modeling languages (e.g., overloading *BPMN* with shapes could influence the different semantic meaning). In addition, we were only concerned with visualizing the differences between two models and, optionally, their instance traffic. Therefore another interesting topic for future work would be the visualization of process evolution, for example, to visualize how a process evolves over the period of one year.

Acknowledgments. Simone Kriglstein was supported by CVAST (funded by the Austrian Federal Ministry of Science, Research, and Economy in the exceptional Laura Bassi Centres of Excellence initiative, project nr: 822746).

References

1. van der Aalst, W.M.P.: Business alignment: Using process mining as a tool for delta analysis and conformance testing. Requir. Eng. 10(3), 198–211 (2005)
2. van der Aalst, W.M.P.: Process mining: discovery, conformance and enhancement of business processes. Springer Science & Business Media (2011)
3. van der Aalst, W.M.P.: A decade of business process management conferences: personal reflections on a developing discipline. In: Proc. of the 10th Int. Conf. on Business Process Management. pp. 1–16. Springer (2012)
4. Albrecht, M., Estrella-Balderrama, A., Geyer, M., Gutwenger, C., Klein, K., Kohlbacher, O., Schulz, M.: Visually comparing a set of graphs. In: Graph Drawing with Applications to Bioinformatics and Social Sciences. No. 08191 in Dagstuhl Seminar Proc. (2008)
5. Andrienko, N., Andrienko, G.: Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach. Springer (2005)
6. Archambault, D.: Structural differences between two graphs through hierarchies. In: Proc. of Graphics Interface. pp. 87–94. Canadian Inf. Processing Society (2009)
7. Bertin, J.: Semiology of graphics : diagrams, networks, maps / Jacques Bertin ; translated by William J. Berg. University of Wisconsin Press (1983)
8. Erten, C., Harding, P.J., Kobourov, Stephen, G., Wampler, K., Yee, G.: GraphAEL: graph animations with evolving layouts. In: Graph Drawing, LNCS, vol. 2912, pp. 98–110. Springer (2004)
9. Gall, M., Rinderle-Ma, S.: Differencegraph (2015), http://gruppe.wst.univie.ac.at/projects/diffgraph/
10. Genon, N., Heymans, P., Amyot, D.: Analysing the cognitive effectiveness of the BPMN 2.0 visual notation. In: Software Language Engineering, LNCS, vol. 6563, pp. 377–396. Springer (2011)
11. Green, M.: Toward a perceptual science of multidimensional data visualization: Bertin and beyond. ERGO/GERO Human Factors Science (1998)
12. Kabicher, S., Kriglstein, S., Rinderle-Ma, S.: Visual change tracking for business process models. In: Proc. of the 30th Int. Conf. on Conceptual Modeling. pp. 504–513. Springer (2011)
13. Kriglstein, S., Wallner, G., Rinderle-Ma, S.: A visualization approach for difference analysis of process models and instance traffic. In: Business Process Management, LNCS, vol. 8094, pp. 219–26. Springer (2013)
14. Küster, J.M., Gerth, C., Förster, A., Engels, G.: Detecting and resolving process model differences in the absence of a change log. In: Proc. of the 6th Int. Conf. on Business Process Management. pp. 244–260. Springer (2008)
15. Mackinlay, J.: Automating the design of graphical presentations of relational information. ACM Trans. Graph. 5(2), 110–141 (1986)

16. Moody, D., van Hillegersberg, J.: Evaluating the visual syntax of UML: An analysis of the cognitive effectiveness of the UML family of diagrams. In: Software Language Engineering, LNCS, vol. 5452, pp. 16–34. Springer (2009)
17. Porst, R.: Fragebogen. Ein Arbeitsbuch 3 (2011)
18. Rozinat, A., van der Aalst, W.: Conformance checking of processes based on monitoring real behavior. Information Systems 33(1), 64 – 95 (2008)
19. Soto, M., Münch, J.: Process model difference analysis for supporting process evolution. In: Software Process Improvement, LNCS, vol. 4257, pp. 123–134. Springer (2006)
20. Wang, Z., Wen, L., Wang, J., Wang, S.: TAGER: Transition-labeled graph edit distance similarity measure on process models. In: On the Move to Meaningful Internet Systems: OTM 2014 Conferences, LNCS, vol. 8841, pp. 184–201. Springer (2014)