Title	A sequence analytics approach for detecting handoff patterns in workflows: an exploratory case study on the Volvo IT incident management process
Author(s)	Kumar, Akhill; Thummadi, B. Veeresh
Editor(s)	Maedche, Alexander vom Brocke, Jan Hevner, Alan
Publication date	2017
Original citation	Kumar, A., Storey, V., Woo, C. 2017. 'Domain Design Principles for Managing Complexity in Conceptual Modeling'. In: Maedche, A., vom Brocke, J., Hevner, A. (eds.) Designing the Digital Transformation: DESRIST 2017 Research in Progress Proceedings of the 12th International Conference on Design Science Research in Information Systems and Technology. Karlsruhe, Germany. 30 May - 1 Jun. Karslruhe: Karlsruher Institut für Technologie (KIT), pp. 1-10
Type of publication	Conference item
Link to publisher's version	https://publikationen.bibliothek.kit.edu/1000069452 http://desrist2017.kit.edu/ Access to the full text of the published version may require a subscription.
Rights	**©2017, The Author(s). This document is licensed under the Creative Commons Attribution – Share Alike 4.0 International License (CC BY-SA 4.0): https://creativecommons.org/licenses/by-sa/4.0/deed.en** https://creativecommons.org/licenses/by-sa/4.0/deed.en
Item downloaded from	http://hdl.handle.net/10468/4437

A Sequence Analytics Approach for Detecting Handoff Patterns in Workflows: An Exploratory Case Study on the Volvo IT Incident Management Process

Akhil Kumar, B Veeresh Thummadi

Pennsylvania State University, University Park, PA, 16802. USA
{akhilkumar, veereshthummadi}@psu.edu

Abstract. In this study, we analyze the activity logs of fully resolved incident management tickets in the Volvo IT department to understand the handoff patterns i.e., how actors pass work from one to another using a sequence analytic approach (a method for studying activity patterns from event log sequences). A generic actor pattern here describes the sequence in which actors participate in the resolution of an incident. We classify actor handoff patterns as straight, loop and ping-pong. Then we analyze the patterns by frequency and duration to draw insights about how actor patterns affect the incident resolution time. The results are quite surprising. In particular, we find that certain loop and ping-pong patterns outperform straight patterns even though more steps are involved in them. Our results have implications for resource allocation in organizations. They suggest that handoff patterns should be another factor to be considered while allocating work to actors along with position, role, experience, skill, preferences, etc.

Keywords: Workflows · Routines · Handoffs · Sequence analytics · Actor patterns · Pattern variety

1 Introduction

Any business or healthcare process can be viewed as a series of handoffs between task actors (or workers) who perform successive tasks until the process instance is completed. After an actor completes her task she hands off the process instance or case to the next actor. Such behavior is observed in various kinds of application areas ranging from medicine and software development to insurance claim processing. Some handoffs also occur in a *ping-pong pattern* such that an actor A hands off a task to actor B only to have it returned later, either after some work is done or just untouched. This leads to an actor handoff pattern represented by the sequence A-B-A, A-B-C-D-A, etc. Such ping-pong behavior arises from an alternating pattern in which the same actor appears more than once.

Most of the research in the Business Process Management (BPM) and workflow literature has been devoted to the discovery of process models, conformance checking and process enhancement. This research assumes that historical process execution logs of completed process instances are available for analysis. Thus, say, we have a log like:

T1 – T4 – T5
T1 – T3 – T5 – T4
T1 – T2 - T4 – T5
T1 – T2
T1 – T5

By applying a process mining algorithm [18] we may discover a process model that always starts with task T1; next tasks T2 and T3 appear in a choice (or alternative) structure and are followed by a parallel structure of T4 and T5 that can appear in any order. Moreover, it is also possible to skip the T2-T3 or T4 -T5 substructure, but only one skip is allowed not both. The discovery of process models in this manner is useful for it helps us to understand the control flow of a process. The drawback with nearly all process mining approaches though is that when considerable variety is inherently present in a process, capturing it in excruciating detail leads to a model that becomes unreadable. The model becomes overrun with so many connectors and edges to accommodate all the flow paths that it looks like spaghetti. In turn, it becomes very hard to decipher and this diminishes its real value.

Researchers in the area of *organizational routines* define routines as "repetitive, recognizable patterns of interdependent actions, carried out by multiple actors."[4] They accept that variety exists in real world processes and have developed approaches to quantify routine variation, by posing questions like: Does the process in Unit A have more variety than the one in Unit B?[11] Moreover, researchers in organizational routines ascribe greater agency and tacit knowledge to actors of various tasks in terms of their interpretations of how the task should be done rather than treating actors as fully interchangeable. Their focus is on interdependencies among actions, people and technology in contrast to the control flow, data flow and resource perspectives of BPM.

Our approach is in part inspired by previous work in the context of routines [11,12], but our work fits into the broader area of work distribution and resource allocation in BPM (e.g. [6, 14]). To this end, we conduct a preliminary exploratory study, and pose new questions for understanding handoffs patterns in workflow data. Thus we ask: what are the patterns of interaction among generic actors (as opposed to actions) in a large real log and how can they inform us? A pattern like 1-2-1-3-1 shows the order of involvement of actors 1, 2 and 3 in the completion of a process instance through four handoffs among themselves. In this pattern, one might place multiple interpretations. We can interpret this pattern as actor 1 dividing some work between actors 2 and 3 and finally integrating the two pieces of work for resolving the ticket. Alternatively, this pattern could arise when 2 was unable to complete the work sent by 1 and returned it, thus 1 had to instead turn to 3 to perform it. The first interpretation suggests a productive way of completing a work instance while the second is counter-productive. By examining the patterns in more detail in conjunction with the duration of various instances that conform to that pattern we expect to be able to design better work allocation methods.

Our study was made possible by access to a large log from the incident resolution process at the Volvo automotive company. This data set is public and hence the results can be verified. By correlating the most frequent types of actor sequence patterns in this dataset with the duration for resolving the incident we were able to gain many useful insights about the significance of actor patterns. By analyzing this data, we hope to address questions like: What are the common actor patterns found in resolving tickets? Are some patterns better than others and why? How does actor pattern variety affect resolution time of the tickets? In this way, we can shed more light on the resource perspective in a business process. This perspective has implications for assignment of resources to a process in an efficient way.

This paper is organized as follows. Section 2 gives some preliminaries about handoffs and sequence analytics. Then, Section 3 describes our log data and the main results of our analysis. Next, Section 4 offers a discussion and mentions some related work. The paper concludes with Section 5.

2 Sequence Analytics

A handoff is a transfer of work from one actor to another. Research in healthcare has shown that communication breakdowns among medical professionals can lead to adverse effects on surgical patients [10]. These breakdowns result from poor handoffs involving verbal communications and ambiguities about responsibilities. There is ample evidence to suggest that the nature of social interactions and interdependencies among participants (or resources) who collaborate on a routine or a process does have an impact on the outcome and performance of the process in terms of quality, failure rate, etc. [7].

Sequence analytics refers to the concept of analyzing the sequences of actions or elements to detect similarities and differences across the sequences [11]. For example, in biology this concept is used to detect evolutionary patterns, rate of mutation and any genetic modifications that occurred in time. This concept was later adopted by sociologists and more recently in information systems to detect socio-material entanglement in work processes [5]. An *action* or *task sequence pattern* is a series of possible orderings of related tasks to complete a process or a workflow. Some examples are:

T1-T2-T3-T4
T1-T3-T4-T1

Similarly, it is possible to also consider actor sequences. An actor sequence would define sequences of specific actors such as: A1-A2-A3-A4, or A3-A2-A4-A1, etc. Each sequence denotes the order in which various actors perform tasks to complete a workflow or routine. In contrast to these two notions, in this paper we are interested in studying actor patterns. By a pattern, we mean an ordering in which *generic actors* perform tasks to complete a workflow or a routine. Thus an actor pattern like 1-2-3-4 means that some actor 1, handed over the work to actor 2 who in turn passed it along to 3 and so on. We call this a *straight pattern*. Another pattern is 1-2-3-1 is a *loop pattern* where the work is returned *at the end* to the same actor who started it. Yet another pattern may be 1-2-1-2-1-3. This is a *ping-pong* pattern since the actors alternate with one another. It is important to note here that 1, 2, 3 are generic placeholders for actors, and not specific names of actors. See Fig. 1 for examples of these patterns.

Our goal is to analyze such generic patterns to determine the kinds of patterns appear most frequently and also to understand if some patterns are better than others in terms of incident resolution times.

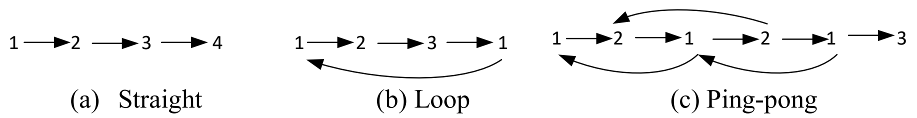

Fig. 1. Different kinds of generic actor patterns

3 A Case study of Volvo IT Incident Management

3.1 Data Set Description and Analysis

This process log data is publicly available and was initially released as a part of the Business Process Intelligence (BPI) Challenge in 2013 [1]. The dataset contains a log of incidents or cases to be resolved. Each incident has a unique serial number. Typically, there are many log records for each incident to reflect any status or owner change. A log record captures the status of the incident and includes information like serial number, date-timestamp, status, sub-status, impact, product, country, owner, support team, and organizational line (see the partial log shown in Table 1). There are 7554 cases or incidents and 65553 events or records in the log for an average of 8.7 log records per incident. The period of this data set extends from the end of March 2010 until middle of May 2012.

Table 1. A snapshot of the incident log

SR Number	Date	Status	Impact	Product	Country	Owner Name
1-364285768	2010-03-31	Accepted	Medium	PROD582	France	Frederic
1-364285768	2010-03-31	Accepted	Medium	PROD582	France	Frederic
1-364285768	2010-03-31	Queued	Medium	PROD582	France	Frederic
1-364285768	2010-04-06	Accepted	Medium	PROD582	France	Anne Claire

The owner attribute in the log record denotes the actual actor who performs a task. When two successive (in time sequence) log records have different owners it indicates a handoff of work from the previous owner to the new one. If two successive records have the same owner, it means that there is a status change and not a handoff. We were only interested in the incidents where at least one handoff occurred.

The dataset was loaded in a MySQL database for the analysis. We first removed the log records for the owner 'Seibel' because this is the information system, and not a human owner. Our focus was on studying the effect of handoffs among human actors only. After all, a resource allocation algorithm can only select a specific human from a set of alternatives. Then we removed all records for incidents where only one human owner was involved and also for incidents that were not resolved. This left us with 4375 incidents - 1755 tickets of low impact, 2413 of medium impact, 204 of high impact and 3 of major impact. Next, we wrote MySQL queries to determine the duration, number of handoffs, number of owners, and handoff pattern for each incident.

Table 2. Incidents by impact level between April 1 and May 15, 2012

Impact level	Number of Incidents
Low	1670
Medium	2204
High	187
Major	3

By plotting resolution time for incidents against time it was observed that resolution times declined with time. This was attributed to a learning curve effect in the initial period that stabilized later on. Moreover, the largest number of incidents were also concentrated in the last part of the dataset, within a short period from April 1 until May 15 of the year 2012. There were 4064 incidents during this period out which the largest number by impact were 2204 medium impact incidents (see Table 2). By concentrating on this period, we were able to eliminate any learning curve effect by removing just a small fraction of the total number of incidents. For our analysis we decided to focus only on the medium and low impact incidents.

Next, we created generic or abstract patterns from the sequences of handoffs for each incident in the log in the following manner. For example, the ticket #1-523391859 has 8 events recorded in the log and contained series of operations in resolving the case. The incident went through multiple hands, 'Elaine-Elaine-Elaine-Elaine-Elaine-Rafael-Rafael-Siebel', before finally getting resolved in Siebel. As we were interested in the abstract handoff patterns among actors, we removed consecutive repetitions with the same owner name (and also owner Siebel) resulting in, 'Elaine-Rafael'. By coding first actor as '1' and second actor as '2' and so on, we were able to generate patterns to convert the sequence of owners' names to numbers 1, 2, 3, ... These numbers give the order in which an actor appears in the incident resolution process. For each incident we generated the actor pattern along with other information like frequency, average duration, etc. These results are shown in Tables 3 and 4 for medium and low impact incidents, respectively.

3.2 Analysis of Results

Tables 3(a) and (b) show the top 10 most frequent actor patterns that appear in the data set for low and medium impact incidents, respectively. The top 5 actor patterns account for about 70% and the top 10 account for about 80% of the incidents in both tables. The last column of Tables 3(a) and (b) shows the rank by duration time of the various patterns. We have excluded from both tables patterns that had a frequency of less than roughly 1% of the total number of incidents in their category. From these patterns one can easily determine the number of unique owners that took part in resolving the corresponding incident and also the number of handoffs.

Note that 9 out of 10 patterns are common to both tables. Further, the top 3 patterns are identical, and 4 out of the top 5 are common as well. This suggests that similar patterns are used to resolve incidents of both low and medium impact.

Table 3. Top 10 most frequent actor patterns for **Low** and **Medium** Impact Incidents

(a) Low impact tickets

Actor Pattern	Freq-uency	Average Duration	Rank
1-2	577	10.2	1
1-2-3	317	10.8	3
1-2-3-4	177	14.3	7
1-2-3-4-5	83	15.0	8
1-2-3-4-5-6	48	15.5	9
1-2-3-1	33	10.5	2
1-2-1	31	10.9	4
1-2-3-4-5-6-7	27	17.8	10
1-2-3-2	25	13.0	5
1-2-3-4-1	19	13.8	6

(b) Medium impact tickets

Actor Pattern	Freq-uency	Average Duration	Rank
1-2	747	9.3	5
1-2-3	492	10.4	6
1-2-3-4	195	13.8	8
1-2-1	92	9.0	3
1-2-3-4-5	90	12.9	7
1-2-3-1	78	7.6	1
1-2-3-4-5-6	35	14.3	9
1-2-3-4-1	30	8.0	2
1-2-3-2	19	16.2	10
1-2-3-4-5-1	18	9.2	4

One interesting effect found in Table 3(b) is that a pattern like 1-2 is more frequent than a similar pattern 1-2-1, but the latter takes smaller duration. Similarly, we find that pattern 1-2-3 is more frequent than 1-2-3-1 that has a shorter duration by 27%. We find that many such loop patterns have a shorter duration than their straight pattern counterparts. This raises the question, why does an instance with one additional handoff take a shorter duration than without it? Further investigation showed that in many incidents there was a large lag time between the last two log entries of “Completed-Resolved” and “Completed-Closed”. The last step was performed by the system. In many cases with the straight pattern it increased to 8 days but was lower in the loop pattern. We also examined all 106 incidents for `Prod424'and show the results for the top-10 patterns in Table 4. Notice from the last column for pattern type that only two of the 10 patterns are straight patterns though they account for 80% incidents, while the other patterns for 20% incidents.

Table 4. Top 10 smallest duration actor patterns for **Medium** Impact Incidents for **Prod424**

	Pattern	Frequency	Average Duration	Pattern type
1.	1-2-3-4-2	1	6.0	Ping-pong
2.	1-2	24	6.8	Straight
3.	1-2-1-2-3	1	8.0	Ping-pong
4.	1-2-3-1-2-4-5-6	1	8.0	Ping-pong
5.	1-2-1	5	8.8	Loop
6.	1-2-3-4-5-2-5	1	9.0	Ping-pong
7.	1-2-3-2	1	9.0	Ping-pong
8.	1-2-3-1	2	9.0	Loop
9.	1-2-3-4-2-5-6-1	1	9.0	Ping-pong
10.	1-2-3	26	9.6	Straight

3.3 Understanding Factors that Affect Duration

To gain a better understanding of the factors that affect the duration of an incident, we made an ordinary least regression (OLS) model in R to predict *Duration* using number of owners (*Owners*) and number of handoffs (*Handoffs*) as independent variables. Since there is a correlation of 0.90 between *Handoffs* and *Owners* we introduced a new variable, *Ping* = *Handoffs* – *Owners* + *1*. The correlation between *Owners* and *Ping* is 0.55. The results in Fig. 2 show that there is a significant relationship between Owners and Duration at the 1% level. After trying several models, this model produced better results than models with a single or two variables. More importantly, by trying single variable models with Owners and Handoffs we found that Owners is a better predictor of Duration than Handoffs with a higher R-squared and coefficient values.

```
Call:
lm(formula = Duration ~ Owners + ping + Owners:ping, data = cdata)
Residuals:
   Min      1Q  Median      3Q     Max
-15.922  -4.968  -1.256   2.839  40.744
Coefficients:
              Estimate   Std. Error t value Pr(>|t|)
(Intercept)    6.77316    0.40440  16.749  <2e-16 ***
Owners         1.24127    0.11658  10.647  <2e-16 ***
Ping           0.42191    0.23728   1.778  0.0755 .
Owners:ping   -0.04813    0.02889  -1.666  0.0959 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 7.794 on 2200 degrees of freedom
Multiple R-squared:  0.07659,   Adjusted R-squared:  0.07533
F-statistic: 60.82 on 3 and 2200 DF,  p-value: < 2.2e-16
```

Fig. 2. Output of OLS regression

4 Discussion and Related Work

Several interesting initial results emerge from our study. By categorizing the actor patterns into straight, loop and ping-pong, we found that straight patterns are often dominated by loop and ping-pong patterns with the same (or even more) number of owners. The OLS models confirmed that handoffs have a smaller effect on duration.

The superior performance of loop and ping-pong patterns suggests that rather than one actor holding onto an incident, frequent exchanges among actors are better. A second feature of loop, and also ping-pong patterns, is that a single actor may take ownership of the incident and monitor it. This accelerates its progress leading to faster resolution. In fact, the actor who appears more than once may be playing the role of a "coordinator" to facilitate the smooth transfer of work among others. Evidence of this is also found in the work of Liu, et al. [8] who have constructed social network analysis to develop an enhanced organizational model. In their model, various actors play social roles like team leader, coordinator, etc. arguably leading to superior team formations.

Although the loop and ping-pong patterns perform better, yet the straight patterns are predominant. This suggests that for relatively easy incidents straight patterns are perhaps the best. What our results also indicate is that resource allocation algorithms should be designed to take handoff patterns into consideration. In terms of pattern variety, out of 106 cases pertaining to Prod424, there were 32 actor patterns. Among these, 28 had a frequency of 5 or less (24 patterns only 1). The longest pattern had 18 handoffs and the smallest 1. This illustrates the diversity of patterns that are used to resolve incidents.

Typically, there are many decisive factors that are used to determine what actor would be assigned to perform a task in a process. These include among others the role or position of the actor, experience, skills, etc. These factors should naturally be taken into consideration. There is a large body of work on resource allocation in BPM (see, e.g. [6,14]) that we cannot review for space reasons. However, one part of this work relates to resource assignment languages [2], resource preference models [3], etc. There is also some valuable work in the area of organizational mining that helps to understand resource assignment patterns and relates to how the involvement of resources influences the control flow of a process [9,15,16]. Related work has also looked into affordance networks that combine actors, actions and artifacts [13]. Event interval analysis [17] is also very relevant in the context of actor patterns to understand the nature of handoff intervals.

5 Conclusions

Our work here is complementary to a large body of work already published on human resource allocation in business processes. In this paper we presented an approach for analysis of actor pattern sequences that can help to improve our understanding of the resource perspective of a process and give fresh insights into work design practices. This empirical study was conducted in the context of a rich data set from an incident resolution process. We found that the straight, ping-pong and loop patterns were predominant in the dataset we analyzed and that often the straight pattern was dominated by the other two patterns.

In future work, we would like to examine some other data sets to see if the patterns found here are of a general nature and whether there are other kinds of frequent patterns that are found in them as well. Moreover, it would be nice to study how handoff patterns can be more tightly integrated into resource allocation methods. A resource allocation algorithm should be able to learn to distinguish good handoff patterns from bad ones and then promote those patterns. Finally, it would be helpful to analyze the role of variety of patterns to see whether more variety is conducive or detrimental to better performance.

References

1. BPI Challenge, 2013. http://www.win.tue.nl/bpi/doku.php?id=2013:challenge
2. Cabanillas, Cristina, Manuel Resinas, and Antonio Ruiz-Cortés: Defining and analysing resource assignments in business processes with RAL. ICSOC. Springer Berlin Heidelberg (2011)
3. Cabanillas, Cristina, et al.: Priority-based human resource allocation in business processes. International Conference on Service-Oriented Computing. Springer Berlin Heidelberg (2013)
4. Feldman, M.S. and B.T. Pentland: Reconceptualizing Organizational Routines as a Source of Flexibility and Change. Administrative Science Quarterly. 48(1): p. 94-121 (2003)

5. Gaskin, J., Berente, N., Lyytinen, K., and Yoo, Y.: Toward Generalizable Sociomaterial Inquiry: A Computational Approach for 'Zooming In & Out' of Sociomaterial Routines. MIS Quarterly, 38(3) (2014)
6. A. Kumar, W. M. P. van der Aaslt and E. M. W. Verbeek: Dynamic work distribution in workflow management systems: How to balance quality and performance. J. Manage. Inf. Syst. Vol. 18 No. 3 pp. 157-193 Winter 2002.
7. A. Lindberg, N. Berente, J. Gaskin, K. Lyytinen (2016): Coordinating Interdependencies in Online Communities: A Study of an Open Source Software Project. Information Systems Research. Published online in Articles in Advance (2016)
8. Liu, R., Agarwal, S., Sindhgatta, R. R., Lee, J.: Accelerating collaboration in task assignment using a socially enhanced resource model, BPM 2013: 251-258, Springer (2013)
9. Ly, L.T., Rinderle, S., Dadam, P., Reichert, M.: Mining staff assignment rules from event-based data. In: BPM Workshops, pp. 177–190 (2006)
10. Mazzocco, K., et al.: Surgical team behaviors and patient outcomes, The American Journal of Surgery. 197(5), 678–685 (2009)
11. Pentland, B.T.: Sequential Variety in Work Processes: Organization Science 14(5):528-540, 2003
12. Pentland, B.T., T. Haerem, and D.W. Hillison: Using workflow data to explore the structure of an organizational routine. Organizational routines: Advancing empirical research (2009)
13. Pentland, Brian, Jan Recker, and George Wyner. A thermometer for interdependence: Exploring patterns of interdependence using networks of affordances, manuscript (2015)
14. N. Russell, A. H. M. ter Hofstede, D. Edmond and W. M. P. van der Aalst. 2005: Workflow data patterns: identification, representation and tool support. In: Proceedings of the 24th international conference on Conceptual Modeling (ER'05), L. Delcambre, C. Kop, H. C. Mayr, J. Mylopoulos and O. Pastor (Eds.). Springer-Verlag, Berlin, Heidelberg, 353-368.
15. Schönig, S., Cabanillas, C., Jablonski, S., Mendling, J.: A framework for efficiently mining the organisational perspective of business processes. Decision Support Systems 89: 87-97 (2016)
16. Schönig, S., Cabanillas, C., Di Ciccio, C., Jablonski, S., Mendling, J.: Mining team compositions for collaborative work in business processes. Software & Systems Modeling: 1-19. (2015)
17. Suriadi Suriadi, Chun Ouyang, Wil M.P. van der Aalst, and Arthur H.M. ter Hofstede: Event interval analysis. Decis. Support Syst. 79, C, 77-98 (2015)
18. Van der Aalst, W.M.P.: Discovery, Conformance and Enhancement of Business Processes. Springer, 2016