Process Mining Case Study

Winter Term 2021/2022

Group 3

Report

Process Mining on the VINST Incident and Problem Management System

(By VOLVO IT Belgium)

|  |
| --- |
| Editor:  Name: Deepanshu Mehta  Course: Data Science and Artificial Intelligence |

Content

[1 Introduction 1](#_Toc95309332)

[2 Descriptive Analysis of the Data Set 2](#_Toc95309333)

[2.1 Incidents 2](#_Toc95309334)

[2.2 Open Problems 6](#_Toc95309335)

[2.3 Closed Problems 9](#_Toc95309336)

[3 Questions](#_Toc95309337)

[3.1 Push to Front 13](#_Toc95309338)

[3.2 Ping Pong Behavior](#_Toc95309339)

[3.3 Wait User abuse 21](#_Toc95309340)

[3.3.1 What does the “Wait-User” status mean and what is it good for? 21](#_Toc95309341)

[3.3.2 How should the “wait-user” status be used? 21](#_Toc95309342)

[3.3.3 How can the “wait-user” status be abused? 21](#_Toc95309343)

[3.3.4 Log analysis for the “wait-user” status 22](#_Toc95309344)

[3.4 Process Differences per Organization](#_Toc95309345)

[3.4.1 Incident Management 25](#_Toc95309346)

[3.4.2 Problem Management 26](#_Toc95309347)

[4 Conclusion](#_Toc95309348)

[5 Literature IV](#_Toc95309349)

6 Reference .......................................................................................................34

List of Figures

[Figure 1: Activities in Incident Handling Log 3](#_Toc95310372)

[Figure 2: Spaghetti-Like diagram 3](#_Toc95310373)

[Figure 3: Typical Incident Handling process model 4](#_Toc95310374)

[Figure 4: Involved org lines 5](#_Toc95310375)

[Figure 5: Open Problems Process Model 7](#_Toc95310376)

[Figure 6: Open Problems Involved Org line 8](#_Toc95310377)

[Figure 7: Open Problems Organization country 8](#_Toc95310378)

[Figure 8: Closed Problems Process Model 9](#_Toc95310379)

[Figure 9: Closed Problems Activity 10](#_Toc95310380)

[Figure 10: Closed Problems Involved Org line 3 10](#_Toc95310381)

[Figure 11: Closed Problems Organization country 12](#_Toc95310382)

[Figure 12: Organizations involved 13](#_Toc95310383)

[Figure 13: Organization A2 Process model 13](#_Toc95310384)

[Figure 14: Python script used to detect and analyze Ping Pong behavior 19](#_Toc95310385)

List of Tables

[Table 1: Log file details 2](#_Toc95310386)

[Table 2: Comparison of the Open and Closed Problem event logs 12](#_Toc95310387)

# Introduction

With organizations like VOLVO, the data frameworks have become key players for storing away the daily generated data as log. This process of storing data and finding useful insights is known as Process mining. Process Mining includes various tools and techniques which help in discovering, monitoring, and improving real processes by using the extracted insights from event logs. Process Mining comes with a package full of advantages to business owners from increasing efficiency to effectiveness and from reducing the cost to human efforts. This report is a using the Process Mining techniques [5] and tools to support VOLVO Belgium IT. We characterize and analyze the process for managing incidents and problems. This work summarizes the findings that we have discovered related to the analysis of the VOLVO Belgium IT data set and how the existing process can be improvised. The data set has three event logs from an incident and problem handling system called VINST: Incident Logs and Problem (Open and Closed) Logs [3]. The goal of this project is to provide answers to a few queries that the business owner has concerning the incidents and problems: push-to-front, Ping-Pong behavior, wait-user status, and process conformance [2].

# Descriptive Analysis of the Data Set

Every organization's incident and problem management procedure are critical to its success. We studied the event log file of VINST incident and problem management system [3] to undertake a deep study of various efficiency and performance parameters using the dataset generated by VINST, and to find some significant insights which can be a lead for further improvements. In this section, we use descriptive approaches to analyze the various data sets and present a summary of our most key insights. To do this, we employed a combination of process mining and data mining techniques, as well as technologies such as Disco and Python (PM4py). The given dataset is classified into

1. Incidents
2. Problems: Open and Closed

|  |  |  |
| --- | --- | --- |
| Event Log | Events | Unique Cases |
| Incidents | 63533 | 7554 |
| Open problems | 2351 | 819 |
| Closed problems | 6660 | 1487 |
| Total | 74544 | 9860 |

Table 1: Log file details

## Incidents

The events took place between the 31st of March, 2010 at 16:59:42 and the 23rd of May, 2012 at 01:22:25. From April 16th to May 19th, 2012, 81 percent of all events were logged. 705 distinct products are linked to activities. There are 13 unique activities and more than 80% of all includes- Accepted+In Progress (46.14%), Queued+Awaiting Assignment (17.62%), Completed+Resolved (9.33%), and Completed+Closed (8.72%).

Ein Bild, das Tisch enthält.

Automatisch generierte Beschreibung

Figure 1: Activities in Incident Handling Log

A process instance lasts up to 2 years and 41 days. The Incident process is unstructured when we look at the entire dataset of Incidents, as seen in Figure below. The numerous transitions between activities result in a spaghetti-like figure.

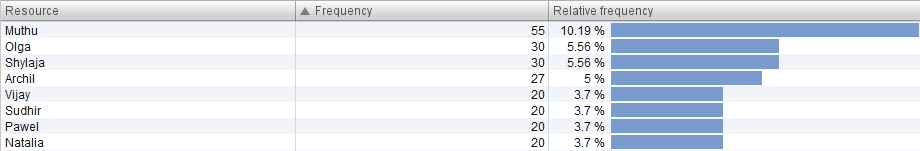


Figure 2: Spaghetti-Like diagram

We discovered that the majority of Incident processes begin with an Accept+In Progress activity when we reduced the activities. Completed+Closed completes the majority of process instances. The Incident Handling procedure is never started with four activities: Unmatched+Unmatched, Completed+Closed, Completed+Cancelled, Assigned+Wait-customer The most common behavior is depicted in the diagram below.

Note: Only the most performed activities and transitions are displayed. The numbers attributed to activities and transitions show how many times they appeared throughout the process.

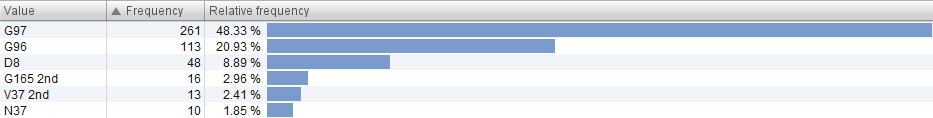


Figure 3: Typical Incident Handling process model

Some attributes can only take a predefined set of values. These are:

|  |  |  |
| --- | --- | --- |
| Problem Status | Problem Latest Impact | Problem Sub Status: |
| Queued (17.62%)  Accepted (61.22%)  Completed (21.16%)  Unmatched (0.01%) | major (0.07%)  high (4.13%)  medium (53.26%)  low (42.54%) | assigned (4.92%)  awaiting assignment (17.62%)  cancelled (0%)  closed (8.72%)  in progress (46.14%)  wait (2.34%)  unmatched (0.01%)  resolved (9.33%)  wait-user (6.43%)  in-call (3.11%)  wait-implementation (0.75%)  wait-vendor (0.48%)  wait-customer (0.15%) |

*Table 2: Attributes of Incident Logs*

The Incident event log is divided into24 organizations, 594 STs and 605 function divisions.

* There are three organizations especially active in the event log: Org line C (41189 activities, 64,38%), Org line A2 (12508, 19,09%), Org line B (4623, 7,05%).
* From 598 STs, the four of them execute in total almost 30% of work: G97 (Org line C, 7466 activities, 11,39%), G96 (Org line, 5999, 9,15%), S42 (Org line, 4382, 6,69%), G230 (Org line, 1661, 2,53%).
* The three most active function divisions include: V3\_2 (30950, 47,23%), A2\_1 (9.977, 15.22%), and E\_10 (4.527, 9.41%).

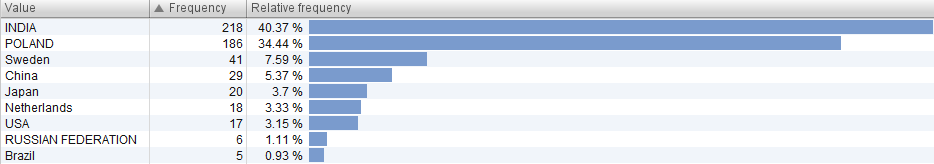


Figure 4: Involved org lines

The figure below depicts the geographical distribution of activities.

* In the 1st line, 46.042 (70.26 percent) of activities are carried out, primarily in Sweden(se), Poland(pl), India(in), and Brazil(be).
* Sweden, India, Poland, and Brazil are the most active countries in the 2nd line, conducting 16.541 (25.24 percent) activities.
* In the third line, 2.911 activities (4.44 percent) are carried out, mostly in Sweden, Poland, France, and India.

Finally, France oversees the "2nd 3rd" support line, which includes 39 activities (0.06 percent). Though we have are going to explore the geographical information only of Incident logs, we highly recommend for future work to give weightage to these insights. The insights from geographical information can help in deducing the cause of the problem more precisely.

 *Figure 5: Organization countries*

## Open Problems

In total, there are 2,351 cases in the closed problem dataset, with 819 unique features. The first case begins at 10:00:36 on November 7, 2006, and the very last case ends at 12:19:56 on June 15, 2012. There are five distinct activities going on at the same time:

* Accepted-In-Progress (49.09%): is the first activity in the majority of Open Problem Dataset processes.
* Completed-Closed (16.46%): completes the majority of process instances.
* Queued-Awaiting Assignment (16.29%)
* Accepted-Assigned (9.15%)
* Accepted-Waiting (9.15%) (9.02 percent)

The most common behavior is depicted in the diagram below.

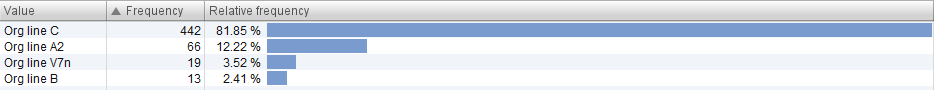


Figure 5: Open Problems Process Model

Some attributes can only take a predefined set of values. These are:

|  |  |  |
| --- | --- | --- |
| Problem Sub Status: | Problem Status | Problem Latest Impact |
| assigned (9.15%)  awaiting assignment (16.29%)  closed (16.46%)  in progress (46.04%)  wait (9.02%) | Queued (16.29%)  Accepted (67.25%)  Completed (16.46%) | major (7.02%)  high (27.39%)  medium (53.08%)  low (12.51%) |

*Table 3: Attributes of Open Problem Logs*

VOLVO IT is divided into24 organizations, 594 STs and 605 function divisions. There are three organizations especially active in the event log: Org line C (1126 activities, 47.89%), Org line A2 (612, 26.03%), Org line G4 (332, 14.12%). These three lines covers 80 percent of the activities occuring over 6 years and therefore are the major lines for generating data.



Figure 6: Open Problems Involved Org line

The geographical distribution of activities is presented in Figure below.

* 72.61% of all activities are performed in the 1st line, mainly in Sweden(se), the ed States(us), and India(in).
* The most active countries in the 2nd line are Sweden, India, Poland, and France, performing 25.24% of all activities.

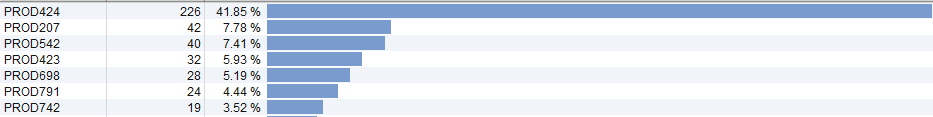


Figure 7: Open Problems Organization country

## Closed Problems

In total, there are 6660 examples in the closed problem dataset, with 1487 individual instances. The dataset spans approximately six years, beginning on January 11, 2006 at 15:49:42 and ending on May 31, 2012 at 23:49:06. With 32.62 percent of all cases, the most common version only has two activities: "Accepted-In-Progress" and "Completed-Closed." When we look at the lengths of individual instances more closely, we can observe that some extreme outliers have a significant impact on the median case duration. One case, for example, took 6 years and 87 days to finish, while another took 4 years and 163 days. In comparison, around 52 percent of all cases are resolved in less than 68 days. This dataset contains a subset of all problems that are completed and closed at some time, i.e., problems with a "completed+closed" ending.

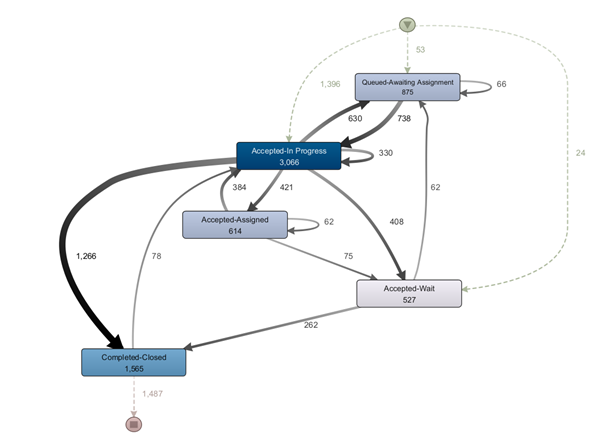


Figure 8: Closed Problems Process Model

In total, the dataset “bpi\_challenge\_2013\_closed\_problem” contains seven different activities in this dataset:

* Accepted-In Progress (46.04%)
* Completed-Closed (23.5%): completes the majority of process instances.
* Queued-Awaiting Assignment (13.14%): is the first activity in the majority of Open Problem Dataset processes.
* Accepted-Assigned (9.22%)
* Accepted-Wait (7.91%)
* Unmatched-Unmatched (0.15%)
* Completed-Cancelled (0.05%)

In comparison to the Instance Handling process, the Problem Handling process is significantly more structured, while it is still not lasagna-like in nature. The complete behavior recorded in the closed event log is shown in the form of a process model in the picture above.

Ein Bild, das Tisch enthält.

Automatisch generierte Beschreibung

Figure 9: Closed Problems Activity

VOLVO IT is divided into 15 organizations. There are four organizations especially active in the event log: C (2702 activities, 40.57%), A2 (1755, 25.52%), G3 (1164, 17.48%), and G4 (608, 9.13%).

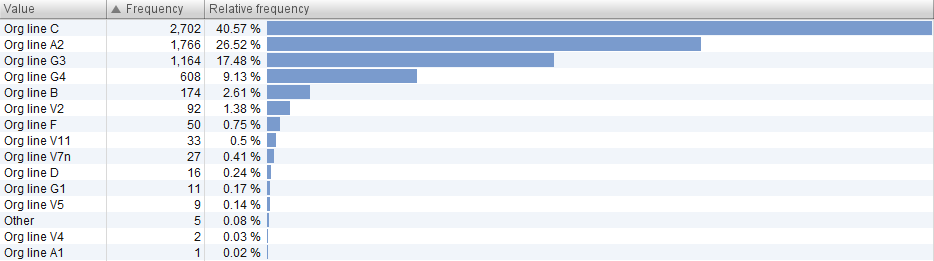


Figure 10: Closed Problems Involved Org line 3

Some attributes can only take a predefined set of values. These are:

|  |  |  |
| --- | --- | --- |
| Problem Sub Status: | Problem Status | Problem Latest Impact |
| assigned (9.22%)  awaiting assignment (13.14%)  cancelled (0.05%)  closed (23.5%)  in progress (46.04%)  wait (7.91%)  unmatched (0.15%) | Queued (13.14%)  Accepted (63.17%)  Completed (23.54%)  Unmatched (0.15%) | major (8.63%)  high (20.47%)  medium (57.31%)  low (13.59%) |

*Table 4: Attributes of Open Problem Logs*

The graph below depicts the geographical distribution of activities. In the first line, 1.754 (72.61 percent) operations are carried out, primarily in Sweden, the United States, and India.

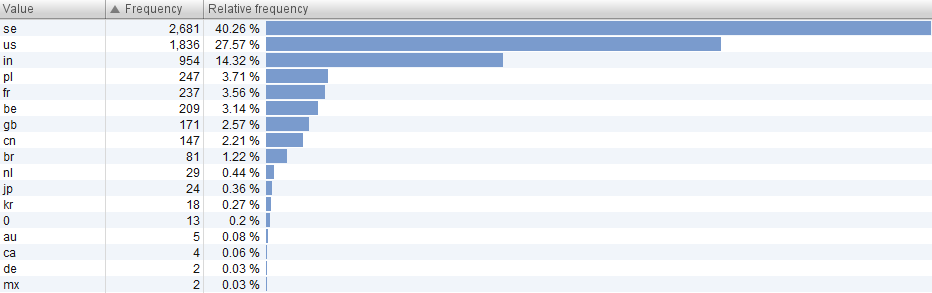


Figure 11: Closed Problems Organization country

The overall finding is that, while the log sizes differ, the registered usual behaviors in both Open and Closed processes are similar. The table below shows a comparison of the Open and Close Problem Handling processes.

|  |  |  |  |
| --- | --- | --- | --- |
| Attributes | Closed Problems | Open Problems |  |
| Maximum time-span | 356 days 22 hours | 4 years 300 days | |
| Minimum time-span | 1min 22 sec | 0 millisec |  |
| Top support lines | 2 and 3 | 2, 3, and 1 |  |
| Top support teams | G199, G88 | G42, S33 |  |
| Top organizations | G3 | G4 |  |
| Top function division | C 6, and E 10 | E 10, and A2 1 |  |

Table 2: Comparison of the Open and Closed Problem event logs

# Questions

In this subsection, we address the different questions of interests. In particular, we give a brief introduction for each question and present our findings and recommendations for the question.

## Push to Front

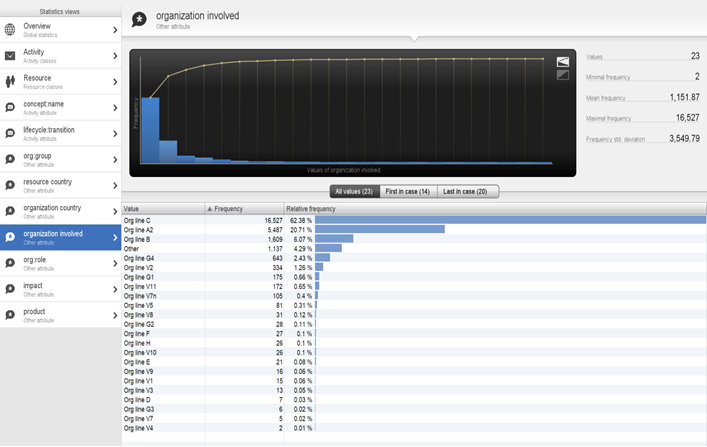


Figure 12: Organizations involved

Out of so many organizations, Organization C and Organization A2 accounts for approx. 83% of the incidents, hence there is more probability for getting push to front cases in these two organizations, hence our reason to consider to them only.

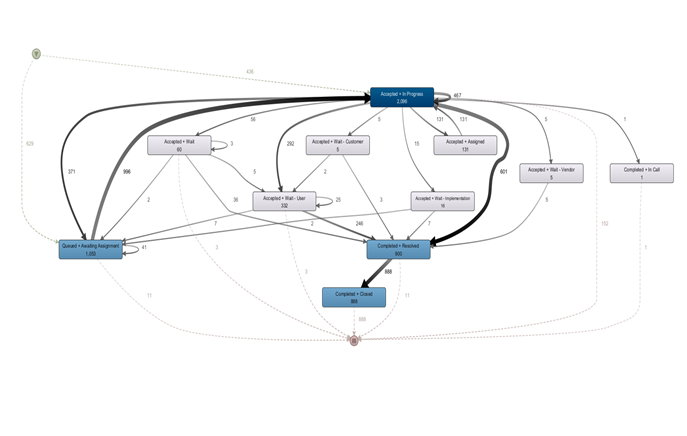
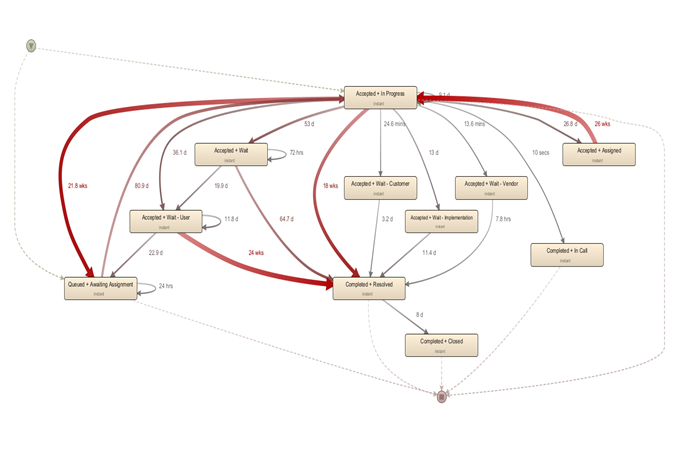
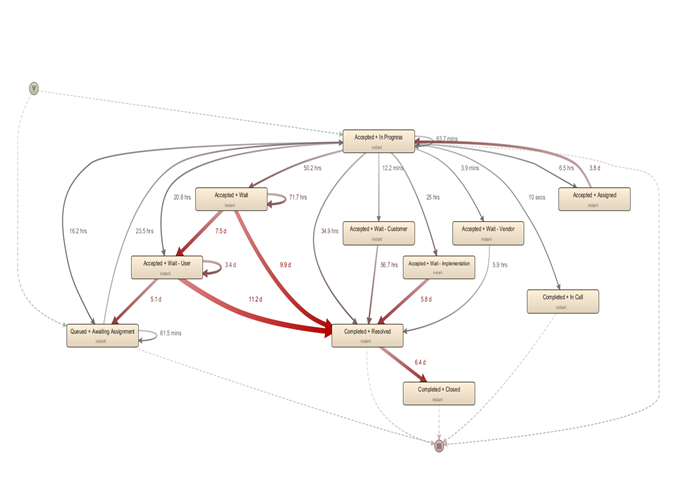


Figure 13: Organization A2 Process model

After filtering the most varying activities (removing most exceptional activities to get the general idea of the situation) and considering the **Organization A2**, it can be seen, there are 629 incidents going directly to the “Queued + Awaiting Assignment” Activity instead of going to “Accepted + In Progress” Activity, hence this implies that a large proportion of incidents i.e. 629 incidents are directly pushed to the 2nd line skipping the acceptance and assignment to the 1st line. 436 incidents have firstly gone to 1st line for their assignment. Hence the incidents are pushed too soon from 1st to 2nd line, but not that often from 2nd to 3rd i.e just 41 incidents.



Bottlenecks: In **Organization A2,** the more the red arrows are, the more is the maximum time to transit from one activity to another. Some of the cases show abnormally high amounts of time consumed, for instance one case consumes 24 weeks to resolve the incident and in one case, 21.8 weeks are taken to push the incident from 1st line to 2nd line. In one of the cases, 1st line takes 18 weeks to resolve the incident.



In **Organization A2,** the above process map shows mean durations, why does it take 6.4 days on average to close the completed and resolved incident, while it should have taken no time.

These suspicious Case IDs should be looked upon.

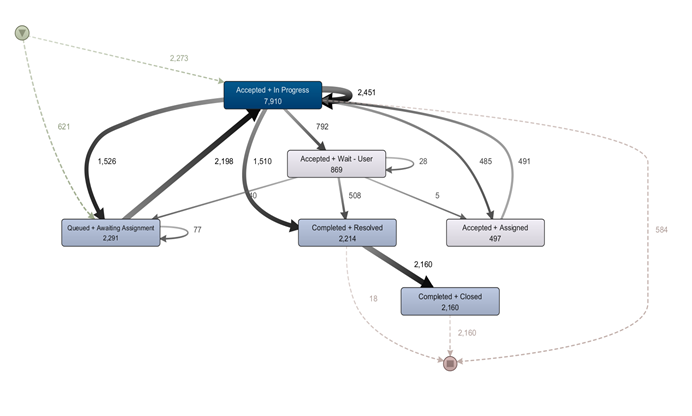
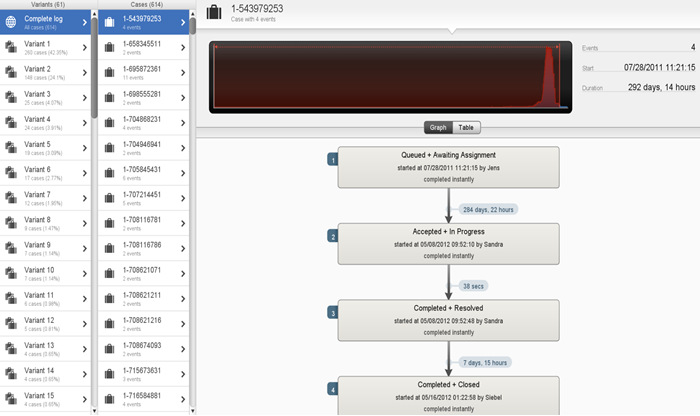
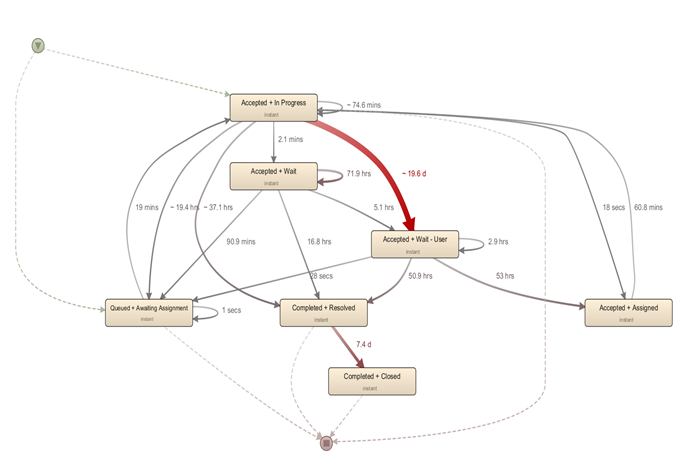


Figure 16: Organization C Process model

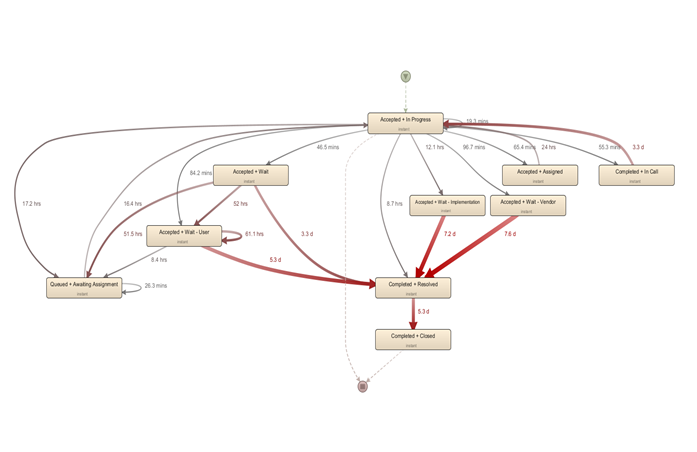
After taking elimination most variant cases and 36% of the Activities and 60% of the paths in the process map, we see that here too the push to front cases have happened i.e 621 incidents are directly shifted to 2nd line and a large number of incidents transferred from 1st line to 2nd line i.e. 1526 incidents and not much incidents are transferred from 2nd line to 3rd line i.e 77 incidents.



After filtering those 621 cases which were directly pushed to the 2nd line, we get to know those cases and analyze them. In the particular shown, the incident was directly pushed to the 2nd line on 07/28/2011 and it took approximately 285 days which is not normal. Also, it took approximately 8 days to close the incident after completion and resolution which is not normal as well.



Bottlenecks: If we consider the median duration, it seems that the users are taking a lot of time trying to solve the incident i.e. around 20 days. The incident should not be allowed to remain with the user for so long.



The process map above shows the mean duration between activities. It can be seen that there is a lot of time being wasted in waiting for the user or for implementation or for the vendor by the 1st line. Also, the same question as in the case of Organization A2 arises, why does it take too long i.e. 5.3 days to close the incident when it is resolved and completed.

## Ping Pong Behavior

Ideally, if a new issue needs to be resolved, one support team can take a look into the issue, solve it and then close it. However, often responsibilities are not clear or the issue cannot be resolved by a single support team. This can lead to situations in which an issue is moved between two support teams repeatedly. This behavior is called “Ping Pong behavior”. Naturally, this can slow down the support process and, thus, should be avoided where possible. Hence, we analyze whether such behavior occurs in the given dataset.

To detect Ping Pong behavior, we use a similar approach as vanden Broucke, Vanthienen and Baesens (vanden Broucke, Vanthienen, & Baesens, 2013) and look for instances where a case is assigned to a support team A, then moved to a support team B and then moved back to support team A at some later point in time. We do this for all of the three datasets that we have. We analyze each dataset with PM4Py. While doing this, we assume that the “org:group” attribute corresponds to the support team. We depict the Python script that we used for our analysis in .



Figure 14: Python script used to detect and analyze Ping Pong behavior

Our analysis for the *incidents* dataset yields that there are 1258 cases that include at least one Ping Pong behavior. Since there are 7554 cases in this dataset, roughly 16.6 % of all cases include a Ping Pong behavior.

Support team “G97” is involved in 123 cases of Ping Pong behavior, support team “D5” is involved in 122 such cases and support team “D8” is involved in 112 cases of Ping Pong behavior. Since “G97” with a relative frequency of 11.39 % of all cases is the support team that is involved in the most cases, it seems reasonable that it is also the team that has the most cases of Ping Pong behavior. However, in contrast, “D5” only occurs in 2.53 % of all cases but is involved in almost as much cases of Ping Pong behavior as “G97” which might be a point that requires further investigation by people that have more insights into the internal processes. The same holds for “D8” which occurs in 2.42 % of all cases but also is involved in 112 cases that involve Ping Pong behavior.

For the *closed problems* dataset, our analysis shows that there are 59 cases that include Ping Pong behavior. Considering a total amount of 1486 cases, approximately 3.9 % of all cases in this dataset include a Ping Pong behavior.

“Org line G3” is involved in 26 cases of Ping Pong behavior and, hence, in almost 50 % of all cases of Ping Pong behavior. “Org line A2” and “Org line C2” are involved in 12 such cases and “Org line B” and “Org line G4” are involved in 6 and 3 such cases, respectively.

For the *open problems* dataset, we see that there are 10 instances of cases that include Ping Pong behavior. These are caused by “Org line C” (60 % with 6 cases), “Org line B” (20 % with 2 cases), “Org line A2” (10 % with 1 case) and “Org line G1” (10 % with 1 case). Due to the fact that this dataset consists of 819 cases, roughly 1.22 % of all cases include a Ping Pong behavior.

We see that there are more than three times as much as cases with Ping Pong behavior in the set of closed problems than in the set of open problems. We cannot provide a definite explanation for this phenomenon, but we expect that this might be due to the fact that, naturally, open problems are generally handled by less support teams so far since they have not yet been resolved.

Similar to another submission (vanden Broucke, Vanthienen, & Baesens, 2013), we compare cases without Ping Pong behavior with cases that include Ping Pong behavior. In the *incidents* dataset, we saw that the mean duration of cases that do not include any Ping Pong behavior is 8.8 days. In contrast, the mean duration of cases that include Ping Pong behavior is 28.8 days which is significantly higher. In the *open problems* dataset and the *closed problems* dataset, we similar differences. Note that these numbers should be treated with care since we do not know anything about the context of the incidents. This should be especially considered since vanden Broucke, Vanthienen and Baesens (vanden Broucke, Vanthienen, & Baesens, 2013) claim that their deeper analysis showed that there is no statistically significant difference in the total life time of incidents with and without Ping Pong behavior.

## Wait User abuse

### What does the “Wait-User” status mean and what is it good for?

According to the documents, "wait user" is a status used to get special customer-specific information needed to resolve the incident. Once the incident is set to this status, the active working time of the incident stops counting and continues only when the status is changed to another one. This can lead to the "wait user" status being abused and used only to save time. However, it does not specify when it is used "correctly" and when it is really used “incorrectly”.

### How should the “wait-user” status be used?

This status is used when user input is required to solve the incident. For example, information about the case or when the user needs to perform an action. As long as one waits for this information or action, the incident cannot be continued. Since it is not specified in more detail what the correct usage is, we try to find out when it was used correctly and when it was used incorrectly and what are the indicators of it with the help of the logs.

### How can the “wait-user” status be abused?

Because the incident cannot continue without user input, the incident must be set to the "Wait User" status, which also pauses the processing time. The exploitation happens when users who are supposed to solve the incidents set their incidents to the "Wait User" status in order to shorten the actual processing time and thus score better in the statistics. The processing time for an incident is always the time it is in the "Accepted - In Process" status, and only stops when it is moved to another status. And by setting the incident to the "Wait User" status, the processing time is stopped.

### Log analysis for the “wait-user” status

We look at the file "bpi\_challenge\_2013\_incidents.xes" with the program Disco. The dataset contains the logs of the incidents. Different filters are applied to filter out all correct applications of the status "Accepted + Wait - User". In the following, we will explain each filter used and see what results the application of the filter brings. This way we can see at the end which statistics result from the wrong application of the status.

Total cases before filters: 7,554

Total events before filters: 65,533

1. Only Cases that contain the “Accepted + Wait – User” Activity

Since we want to examine specifically the status "Wait - User", all cases in which this status does not occur are not relevant. Now we see what start activities we have, “Accepted + Wait - User”, “Accepted + Assigned”, “Accepted + Awaiting Assignment”, “Completed + Resolved” and “Accepted + In Progress”.

Logically it makes no sense to start with “Accepted + Wait - User” without having it first in “Accepted + In Progress”, so we see this already as an abuse. But only 12 cases used this approach, so it could be a mistake as well. Therefore, we will filter them out.

All other start activities except "Accepted + In Progress" also have a too low frequency, therefore they are also filtered out, because they are not relevant. Thus, we see that most cases start with "Accepted + In Progress".

Total cases after 1 filter: 2,495

Total events after 1 filter: 34,625

1. Only Cases that begin with “Accepted + In Progress” and end with “Completed + Cancelled“, “Completed + Closed“, “Completed + In Call“ or “Completed + Resolved“

This way we remove all cases that do not follow the correct scheme, as we only want to check correct cases. We also consider any status with "Completed +" as an end, since for this analysis it is not important whether the Incident was solved or not.

Total cases after 2 filter: 2,060

Total events after 2 filter: 29,003

1. Remove “Unmatched + Unmatched” activity

Cases with the status Unmatched were filtered out because they are not relevant.

Total cases after 3 filter: 2,058

Total events after 3 filter: 28,966

1. Remove User Siebel

We remove the User Siebel, because we assume that it is the system user, and that the system user always uses the status in the right way.

Total cases after 4 filter: 218

Total events after 4 filter: 2,527

1. Filter cases, where time between “Accepted + Wait – User” and “Accepted + In Progress” is more than 1 day and vice versa less than 2 minutes

We assume, that waiting a minimum of 1 day for a user reply is still okay and not seen as abuse. Even if it could be still an abuse of the status, we don’t think that it has a big impact. So now we are left with our final cases. And we assume, that taking only 2 minutes to switch from “Accepted + In Progress” to “Accepted + Wait – User” is too short, to do something productive. We see this behavior as abuse of that status.

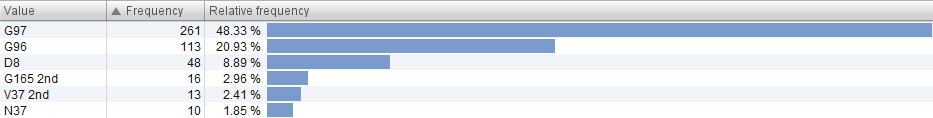
Total cases after 5 filter: 37

Total events after 5 filter: 540

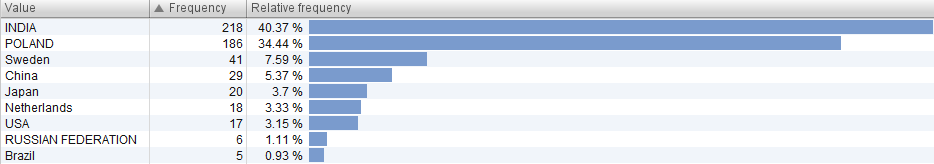
Ein Bild, das Tisch enthält.

Automatisch generierte Beschreibung

So, we see that Muthu, Olga and Shylaja uses the status the most in an abusive way.



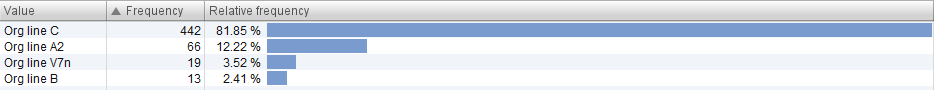
We also see, that the organizations G97 and G96 use that status in an abusive way.



Indian and Polish workers use the status most in an abusive manner with almost 75%.



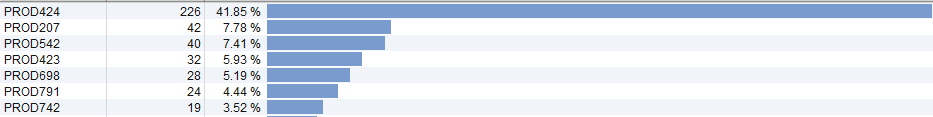
Poland was involved in 95% of the cases.



Org line C has with 81% the most cases where the status gets abused.



The abuse only happens in the Low and Medium impact of the incidents.



And the product PROD424 is mostly involved in an abusive status usage.

So we see, that the status got abused, but only in 1% of the total cases.

## Process Differences per Organization

The Process Owner states that Volvo’s IT organization is split mostly into Organizations A2 and C. As we have seen before, both of these organizations indeed account for more than 83% of all activity in the Incident Management dataset and 65% of all activity in the Closed Problem dataset. Because of this large contribution to the general workflow of their IT department, the Process Owner is therefore especially interested in how these two organizations differ from each other, and how conform they are to the ideal mode of operation. In order to answer this question, we first analyze how both organizations perform individually in the Incident and Problem Management, and then compare them to point out the differences found.

### Incident Management

Looking at the Incident Management first, we see that Line A2 contributes to 23% of all cases, with a median case duration of 9.5 days and an average of 21.1 days. We also see that the most frequent activities used are “Accepted + In Progress”, “Queued + Awaiting Assignment”, “Completed + Resolved” and “Completed + Closed”. It is notable that some cases exhibit unwanted behavior. One of the cases starts with “Completed + Resolved”, while another one starts with “Completed + In Call”. Both of these are not in line with the workflow we would expect, which is that first a case must be accepted, and only then it is possible to close them. Similarly, we see that some cases end with the “Accepted” status. It is especially noteworthy that those cases do not use the “Closed” status at all. Both of these observations suggest that the corresponding Action Owner did not enter the information correctly into the system.

Considering the impact of the incidents associated with Line A2 we notice that the majority of all incidents have an impact of Medium (66.34%). This is followed by Low (26,86%), High (6.68%) and Major (0.12%) impact incidents. This appears normal, as we expect more Low and Medium impact incidents to happen than High or Major ones. Finally, line A2 works with a variety of different products, they do not seem to have a specific range of products associated with them.

Organization Line C is involved in 83% of all cases in the incident dataset. Its median case duration is 7.5 days, while its average is 10.4 days. The most frequent activities are “Accepted + In Progress”, “Queued + Awaiting Assignment”, “Completed + Resolved” and “Completed + Closed”. Here we again find a case which starts with “Completed + Resolved”, as well as cases where the concluding activities are “Accepted + In Progress”, “Accepted + Wait – User” or “Accepted + Wait - Implementation”. Again we assume that most of these conformance problems are the result of incorrect entered information by the Action Owner. Although one case, which length is above the average with 11 days total duration, seems to use the “Wait – User” Sub-Status unreasonably often.

The incident impacts in Line C are a bit more equally distributed between Medium (48.4%) and Low (48.13%) impact incidents, while High and Major impacts occur rather seldom with 3.39% and 0.08% respectively. In the products we can make out that Line C seems to work primarily with “PROD424”, which is associated with 14.7% of all incidents involving Line C, while only 5 products overall occur in more than 2% of all incidents involving Line C.

To conclude our analysis of the Incident Management we first want to point out that considering the frequency of the activities, both Line A2 and Line C seem to generally be in line with the wanted workflow. Only small differences could be found which generally can be attributed to misclassification by the respective Action Owner in a handful of cases. Our analysis also showed, that Line A2 works mostly with Medium impact cases, while Line C equally works with both Medium and Low impact cases. Both organizations are similarly often associated with High and Major impact incidents. Finally, it appears that Line C has a specialization with product 424, which appears significantly more often than other products. While Line A2 does not seem to prefer one specific product. Due to the nature of the data we were not able to further investigate into this specialization of Line C.

### Problem Management

In order to adequately compare both organizations in terms of their problem management workflow, we solely focused on the closed problem dataset. This removes differences which occur due to different progressions in the cases, and thus makes the two organizations more comparable.

Again, starting with Organization Line A2, we find that 30% of all cases have some contribution by Line A2. The median case duration is 17.6 weeks and the average is 35.7 weeks. 44% of activities involving Line A2 are “Accepted-In progress”, 20% are “Completed – Closed” and 12% are “Queued-Awaiting Assignment”. This already suggests much of the time the cases are actively worked on and closed, and fewer time are idle while awaiting the assignment. Further, all cases end on “Completed-Closed”, which is as expected as we are looking at the closed problem dataset. Some activities are labeled “Unmatched-Unmatched”, which we assume to be a misuse by the corresponding Action Owner.

The majority of all problems worked on by Line A2 are of Medium impact (62%), while both High (17%) and Low (14%) impact problems are equally often represented, and Major problems (6%) are seldom solved by Line A2. Thus, it appears as if Line A2 is primarily concerned with Medium impact problems. Products 97 and 98 appear most often in the problems solved by Line A2, with 7.62% and 6.2% respective relative frequency. Products 236 and 753 also appear relatively often with both having each 4% relative frequency.

Organization Line C is involved with 40% of all cases in the closed problem dataset, with a median case duration of 48.7 days and an average case duration of 14.1 weeks. In these cases, the activity “Accepted – In Progress” appears at 45% relative frequency, followed by “Completed – Closed” at 22% relative frequency and “Queued – Awaiting Assignment” at 17% relative frequency. Again, this suggests high productivity. Here as well three cases have the activity “Unmatched – Unmatched” associated with them, which we again think is a misuse by the respective Action Owner.

Organization Line C also works mostly on Medium impact problems, with 44% relative frequency. High impact problems occur at 27% frequency. Interestingly this is followed by Major impact problems at 15% frequency and only then Low impact problems with 14% frequency. This distribution suggests that Line C has a larger focus on Major problems than Line A2, as we would expect Low impact problems occurring more often than Major impact problems. Product 802 appears in 10% of all cases, the following most frequent products are 793 (5%), 597 (4%) and 660 (4%). This gap of 5% occurrence again leads us to believe that Line C has a specialization in product 802.

Comparing both organizations we notice a large gap in their average and medium case time. While the median problem in Line C takes 48.7 days to be closed, the median in Line A2 is at 17.6 weeks (123.2 days). The averages show the same gap with 14.1 weeks in Line C against 35.7 weeks in Line A2. We therefore assume that Line C is faster in solving its assigned problems than Line A2. Besides this time difference both Lines mostly follow the expected workflow of accepting a problem, queuing it and eventually closing it, with both Lines having activities which were unmatched to any status or sub-status attribute. Concerning the impact of the problems solved we see that Line A2 and Line C both primarily work with Medium impact problems, but due to the relative frequency of Major problems compared to Low impact problems in Line C, we further conclude that Line C focuses also on Major impact problems. Finally, both Line focus on different products. Line A2 mostly focuses on products 97, 98, 236 and 753, while Line C primarily works with products 802, 793, 597 and 660.

# Conclusion

The Process mining project is an intriguing way to test the utility of various tools and strategies for digging into real-world log event data. Despite the fact that the datasets were rather limited, we divided our analysis amongst ourselves. In addition, a number of questions arose from four key issues relating to the two procedures. To solve the issues raised by Business Owners, we tried to comprehend the datasets for the queries in order to answer them. Based on our findings, Volvo's application of Incident and Problem processes is positive: challenges are typically limited to certain STs, products, or organizational lines. During analysis, we identified all the products for which the push-to-front mechanism is mostly implemented. As a result of these findings, organizations should be more inclined towards exchanging best practices for push-to-front mechanisms across organizations and function divisions [3]. For the ping pong behavior issue, we discovered that most of the ping pong behavior was caused by function divisions, organizations, and support teams. Ping-pong activity isn't very frequent. We also discovered which products are most influenced by this behavior. These findings aid organization in reducing the overall duration of an incident or a problem. To address the third issue, we looked at who and what was causing the most problems. We investigated the pattern of using the wait-user sub status by the support team, function division, and organization. We also found places where the sub status is being abused. However, only 1 percent of all instances are affected. The major conclusions are that Poland is the country with the most involvement and that only Low and Medium Incidents are impacted. In addition, Org line C is implicated in 81 percent of all cases.

Finally, we addressed the final issue: organizational process conformance. We examined the incident and problem processes of the two organizations lines A2 and C [2]. We discovered that both organizations operate in a manner that is reasonably like our ideal mode of operation. We recommend that you look into the cause of the delay in Line A2's problem management and see if the product focus, as well as Line C's concentration on medium and major problems, are desired. It's also a good idea to emphasize to Action Owners the necessity of utilizing the correct status and sub-status data when working on a case; this will help with subsequent analysis and, as a result, will allow for more exact action recommendations. As a result, we strongly advise that the organization should pay attention to these process mining findings for better performance and decision making [6]. It would be intriguing to implement our analysis and see if our findings are supported by larger VINST event logs.

# Literature

vanden Broucke, S. K., Vanthienen, J., & Baesens, B. (2013). *Third International Business Process Intelligence (BPIC'13): VOLVO IT Belgium VINST.*

6 References

1. Conformance checking of processes based on monitoring real behavior, March (2018)
2. A Business Process Management Guide for Managers and Process Professionals by Paul Harmon (2019)
3. 9th International Workshop on Business Process Intelligence (2013)
4. Summary of BPI Challenge (2013)
5. Process Mining: Discovery, Conformance and Enhancement of Business Processes by Wil M. P. van der Aalst (2012)
6. Discovering social networks from event logs by Wil M. P. van der Aalst (2005)