Business Processing Intelligence Challenge (BPIC): Case study: Analizing Volvo information with Process Mining

Jonathan Hevia[1], Cecilia Saint-Pierre[2]

Pontificia Universidad Católica de Chile,
Vicuña Mackena 4860, Macul, Santiago - Chile.
Código Postal 7820436

Abstract.

The present document presents the analysis performed over the available data of the incident and problems management in the Volvo company in the context of the BPI Challenge 2013. In this work, we tried to give response to the client concerns and provide analysis based on the data, with proposals to improve the performance of the processes in the company.

Specifically we focus on four questions, which are: Push to front problems, Ping-Pong behavior, misusage of the substatus wait-user and process conformity in the organizational lines A2 and C.

Thus, this paper attempts to identify the impact of these failures and organizing process so that in the future Volvo can correct and thus, provide a better service to their customers.

Keywords: Process mining, process analysis, BPM, BPI, Volvo, VINST

1 Introduction

The analysis of the process "problem solving" of Volvo was performed using the tools, DISCO, ProM 5.2 and ProM 6.2, where the focus was make different analysis and organizational flows as exhaustive as possible, to conclude which the problems are and who are responsible for these.

The document is divided in three parts mainly aimed to understand the problem, perform the analyzes that lead us to solve the four main questions of the Head of Company Processes and final part related to the conclusions of the work performed.

Also, we include two appendices were developed in detail the analysis of every case, supporting sections 2.2 and 2.3, these appendices contain diagrams and more detailed explanations about the work developed in the above sections.

[1] Engineering student in computer science.
[2] PhD. Student in Computer Science.

2 Proposed questions

2.1 Push to front

According the process description, we expect the different lines of customers' attention work as a stepwise platform, it means, every case is attended by the first line (mainly service desks) and only an small percentage of that are scaled to the second level. In the same way, the second level should resolve the most of the cases received, scaling just the ones where the operators are not capable to solve.

The criteria to scale a case in the model, can be gave by the complexity of the solution required, being the most specialized people placed in the last level.

The "push to front" mechanism is the way that the second and third level return to the lower level which cases that having a simple solution that could be gave for the lower level, were scaled unnecessarily.

We do not expect to find flows in this way, i.e. cases being delegated from the third line to the second one, or from the second to the first line (pushed to front). This behavior means the first line (or second) is not capable to solve simple cases which are delegated to the second level, but are returned by them to the first line because of this simple nature.

We are trying to identify this situations in the data and to identify that, we will look the frequency and time with the cases are pushed from the first to the second and third line.

2.1.1 Questions

a) For what products is the push to front mechanism most used and where not?
b) The product information is available in the 'corrected structure' field?
c) Where in the organization is the push to front process most implemented (field = involved organization), specifically if we compare the Org line A2 with the Org line C
d) What functions are most in line with the push to front process?

2.1.2 Data selected and log construction

Table 1. Field in the log construction

Fields available in the csv archive	Field identified in the data log
SR Number	Case_Id
Change Date+Time	Timestamp
Status - ST level[3]	Activity

[3] "ST level" is a field created from the field ' *Involved ST'*, extracting from the text on the field the explicit reference to the attention level (2nd, 3rd). In the cases without explicit reference, we considered it was from the first level of attention (*'1st'*).

Involved ST Function Div	Not used
Involved Org line 3	Not used
Involved ST	Not used
SR Latest Impact	Not used
Product	Not used
Country	Not used
Owner Country	Not used
Owner first name	Resource

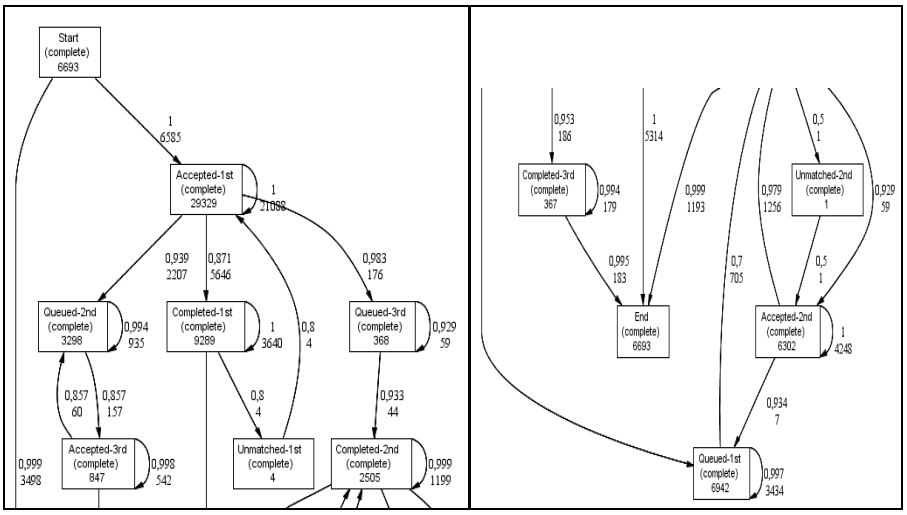

Figure 1. Model obtained using the Heuristic Mining Algorithm

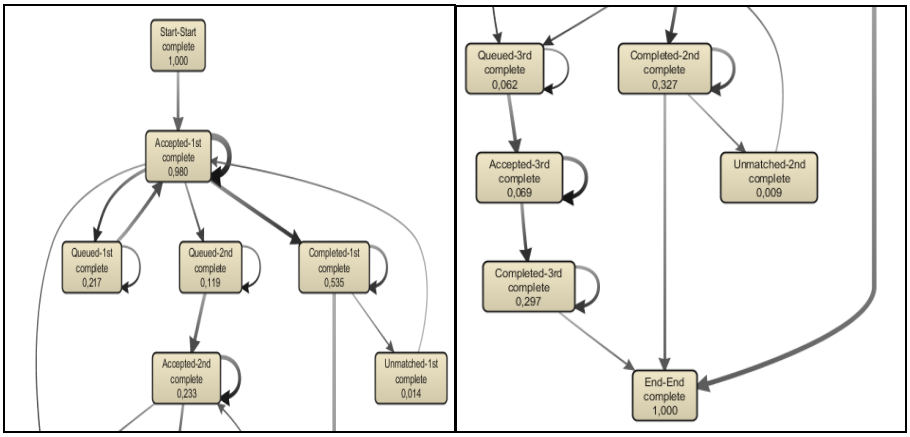

Figure 2. Model obtained using the Fuzzy Mining Algorithm

2.1.3 Data analysis

We analyzed the model obtained and noticed that the cases' behavior is as we expected, i.e. the cases are scaled from first to second and from second to third line and is not observed flows in the other way.

Therefore, the algorithms used to discover the underlying process (Heuristic Miner and Fuzzy Miner) shows important differences about the causal relation between the activities (or status) *'Queued'* y *'Accepted*'.

In the first model discovered, we can see the status *'Queued-2nd*' represents the waiting queue to scale cases to the third line, and the activities *'Queued-1st'* and *'Queued-3rd'* are waiting queues inside the second line. That behavior was not expected, since we expect the status *'Queue'* was the waiting queue for the status *'Accepted'* in every level (for instance, we expect the status *'Queued-2nd'* was the waiting queue for the status *'Accepted-2nd'*).

The second model discovered using Fuzzy miner algorithm do not show this causal problem between *'Queue'* and *'Accepted'* status, instead, every *'Queue'* status was the waiting queue to the respective *'Accepted'* status. In this model we noticed that some cases start the process directly in the status *'Accepted-1st'* skipping the status *'Queue-1st'*.

Even though we did not find flows where the case was pushed back to the lower level of attention, we detected an unwanted behavior in cases which the scale procedure was disrespected scaling from the first line, directly to the third line, skipping the second one. We believe this behavior is highly undesirable because it saturates the third level with less complex cases that could be resolved at a lower level.

In terms of costs, this finding is relevant because the third line, being the one which solves the most complex cases, is also the most costly in human resources, which is making improper use to refer them low complexity cases.

The reasons of this behavior could be because of a wrong execution of the process, preserving bad practices without a significant cause or because the process design. Whatever the cause, this finding reveal a unnecessary cost to the company which should be reviewed in a deeper way.

2.2 *Ping-Pong* Behavior

This behavior is related to the solve capacity of the support teams (ST). Is the action to refer a case from one ST to another and receive it back many times (as a ping-pong ball) without a solution for the client, stretching the solving times of the cases.

We expected to find in the data, evidence of not so many support teams in every case, understanding that in the ideal process model the problems should be solve quick and with interference of not too many support teams.

We analyzed the available data looking for flows between support teams where we could see this undesirable behavior and tried to related this cases with long times of total life of the cases identified.

2.2.1 Questions

a) what is the frequency of the ping-pong behavior among the full set of available cases?
b) Which are the teams involved in this behavior and where are they located?
c) What are the functions, organization, responsible for most of the ping pong?
d) What products are most affected by it?

2.2.2 Data selected and log construction

The process model discovered was obtained from the incident available data. The fields used to the event log construction is shown in Table 2.

Table 2. Fields used in the log construction

Fields available in the csv archive	Field identified in the data log
SR Number	Case_Id
Change Date+Time	Timestamp
Status - ST level[4]	Activity
Involved ST Function Div	Not used
Involved Org line 3	Not used
Involved ST	Resource
SR Latest Impact	Not used
Product	Not used
Country	Other
Owner Country	Not used
Owner first name	Not used

[4] *'ST level'* is a field created from the field ' *Involved ST'*, extracting from the text on the field the explicit reference to the attention level (2nd, 3rd). In the cases without explicit reference, we considered it was from the first level of attention (*'1st'*).

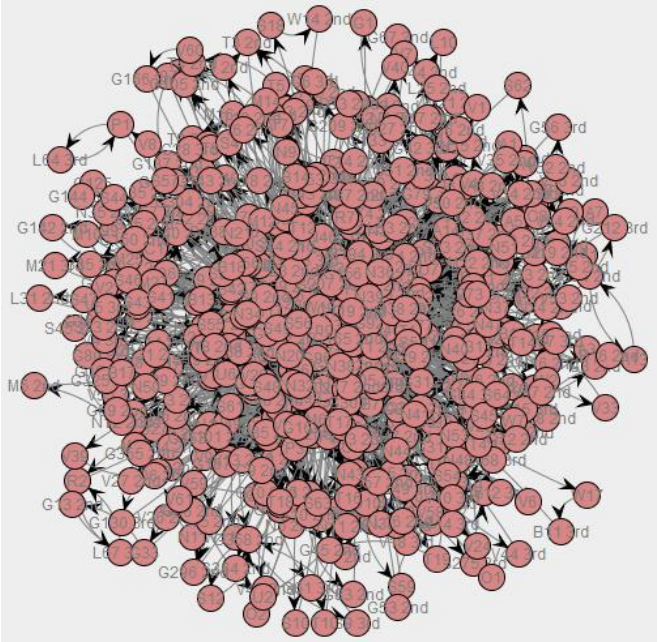

Figure 3. Model obtained using the algorithm Handover Of Work with the 566 teams.

As it can be observed in Figure 1, is impossible make an analysis of the model considering the 566 teams at the same time. Because of this, we decided to segment the log by countries in first instance.

The available data contain cases from 22 countries (considering as *'country'* = 0 the data without this field information). Table 3 shows the percentage of participation by country.

Table 3. Country participation in the total events

Value	*Frequency*	*Relative frequency*
Sweden (se)	18575	31,36%
Poland (pl)	14960	25,25%
India (in)	5939	10,03%
Belgium (be)	5696	9,62%
United States (us)	4773	8,06%
France (fr)	2607	4,40%
Brazil (br)	2540	4,29%
Netherlands (nl)	1277	2,16%
China (cn)	1156	1,95%
Korea (kr)	484	0,82%
Russia (ru)	346	0,58%

Great Britain (gb)	242	0,41%
Australia (au)	184	0,31%
0	159	0,27%
Japan (jp)	137	0,23%
Germany (de)	55	0,09%
Malaysia (my)	47	0,08%
Thailand (th)	24	0,04%
Chile (cl)	16	0,03%
Turkey (tr)	8	0,01%
Peru (pe)	8	0,01%
Canada (ca)	6	0,01%

2.2.3 Data analysis

Further details of this data analysis could be found in Appendix A.

In the analysis of the data filtered by country we have the following findings. In general, countries with so much STs difficult the analysis, this because the behavior is similar to the global model, with a great network of nodes and connection between this (see Figure 3).

We have three kinds of countries:

i. Countries with so much STs: in this category we find Sweden, Poland, India Belgium and United States. For analysis this countries is necessary filter more this logs. Is important note that the behavior of the filter logs will be similar to the next kind of countries.
ii. Countries with less of 100 STs: in this category we find France, Brazil, Netherlands, China, Korea, Russia, Great Britain, Australia, 0, Japan and Germany. Here we have countries with Ping Pong behavior in the most of its STs (Brazil, Netherlands, China, Great Britain and Germany) and another countries with less Ping Pong behavior in its STs (France, Korea, Russia, Australia, 0 and Japan)
iii. Countries with less of 5 STs: in this category we find Malaysia, Thailand, Chile, Turkey, Peru and Canada. In this case just Thailand presents Ping Pong behavior between its two STs.

2.2.4 Answering the proposed questions

From the data analysis we could conclude that in general the presence of Ping Pong behavior is high. This is evidence of a great problem for Volvo due to that the Ping Pong behavior slows the process contradicting a law of the organization, the efficiency in the time of response.

In general, there are so much STs that participant in the Ping Pong behavior. For this reason is primordial revising the functionality of all the process.

2.3 Wait user

The substatus 'wait-user' is how to reflect in the system, a process delay that is not liability with the company, but the customer.

The employees could put stop to the time counter in the incident resolution, changing manually the substatus of the case to 'wait-user'. There are guidelines about not to use this substatus, unless someone is really waiting for an end-user, but is known some action owners are breaking this guideline.

We tried to find evidence in the available data about this misusage of the substatus looking for situations where are two-way causalities between another substatus and wait-user.

2.3.1 Questions

a) Who is making most use of this substatus (*action owner*)?
b) What is the behavior per support team, function, organization etc?
c) Is it possible to detect misusage per location?

2.3.2 Data selected and log construction

To answer the proposed questions we use the available incident data, because the problem data have no information about the substatus of the status wait.

To create the event log, we consider as activity the concatenation of status and substatus, being the activity duration zero, and the waiting time between activities A and B, the time while the case was in the status-substatus of activity A.

We include the substatus in the activity definition because we are trying to answer specifically about one of it.

Finally, we noticed that the process were symmetric in the different levels of customer attention, i.e. we have the same mix status-substatus in every level. Because of that, and include in the activity definition an indicator of the level. Thus, the activity was defined as the concatenation of status, substatus and level.

Table 4. Field in the log construction

Fields available in the csv archive	Field identified in the data log
SR Number	Case_Id
Change Date+Time	Timestamp
Status - Substatus - ST level[5]	Activity
Involved ST Function Div	Not used
Involved Org line 3	Other
Involved ST	Not used
SR Latest Impact	Other
Product	Other
Country	Other

[5] *'ST level'* is a field created from the field ' *Involved ST*', extracting from the text on the field the explicit reference to the attention level (2nd, 3rd). In the cases without explicit reference, we considered it was from the first level of attention (*'1st'*).

Owner Country	Other
Owner first name	Resource

With the event log created, we imported the file in the software DISCO. The log was filtered by start/end events, keeping just the cases where the first activity were Accepted or Queued in the first. As end activity we just considered the status complete no matter which substatus or line was.

We also filter 39 events where the attention line given was second and third simultaneously. This cases where deleted from the event log.

Using those filters we got an event log with an 88% of the original cases, a 90% of the events, and 1269 resources (*owner first name*).

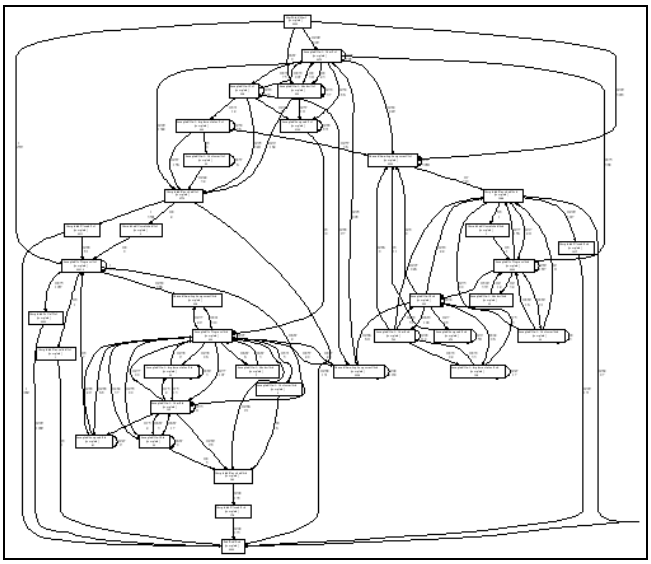

Figure 4. Process model obtained using Heuristic Mining. The figure shows that using the complete log, is hard to handle the resultant model, so we decided to divide it in order to can make a deeper analysis.

We filtered the log by cases where the wait-user status were used and we found that from the 1269 resources in the original log, only 525 use this substatus on any of the attention levels. This filter keep just a 30% of the total cases from 17 countries of the original 32 in the log.

Table 5. Frequency of the different countries in the use of the substatus wait-user. Is interesting to see that Sweden and Poland concentrate an 81% of the total log.

Value	*Absolute frequency*	*Case frequency*	*Relative frequency*
Sweden (se)	967	722	29,43%
Poland (pl)	903	618	27,49%
India (in)	431	273	13,12%
Belgium (be)	363	161	11,05%

United States (us)	220	201	6,70%
France (fr)	111	98	3,38%
Brazil (br)	102	82	3,11%
China (cn)	83	45	2,53%
Netherlands (nl)	58	25	1,77%
Russia (ru)	11	8	0,33%
Great Britain (gb)	9	4	0,27%
Germany (de)	8	2	0,24%
Australia (au)	7	6	0,21%
Japan (jp)	6	5	0,18%
Malaysia (my)	3	2	0,09%
Korea (kr)	2	2	0,06%
Thailand (th)	1	1	0,03%

2.3.3 Data analysis

More detailed information about the analysis can be found in Apendix B.

Thailand

There are only 6 incidents registered in the available data. The wait-user is used just once, and the total life of the case was 5 days. It seem that there is not a misusage of the status.

Korea (kr)

The available data of Korea's Incidents has 22 cases, which are solved by the first attention line, without participation of the second and third lines.

The substatus *wait-user* is used two times in two different cases and we do not detected misuse of it.

Malaysia

There was 6 incidents in the available data, and the substatus where used in 2 of this 6 cases. We detected misusage in 2 of the 3 times the substatus where used (1 case). Considering the occurrence by cases we conclude that there are misusage un a 16,6% of the cases registered.

Japan

It has 13 cases in the available data. In this cases, the substatus was used in 5 cases in the first level, 4 of it were considered misusage of the substatus. Considering that, we conclude that there are misusage in a 30% of the cases.

Australia

It has 24 cases and the substatus wait-user is used in 6 of those. Every case was solved in the first line of attention and there are no evidence of misusage of the substatus.

Germany

It has 2 cases with a long life time (average: 115 days). The substatus are used several times, as pause or not. We consider there are misusage in a 100% of the cases.

Great Britain

It has 21 cases and the substatus wait-user is used in 4 of these. We found evidence that the substatus has misusage in the 4 cases, representing a 100% of the cases.

Russia

It has 45 cases and the substatus was used in 8 cases in first and second levels. We did not find evidence of misusage in the available data.

Netherlands

It has 57 cases, and substatus was used in 25 of these. In 14 of these 24 we found misuse, related to one *action_owner* (Olga).

China

China data has 96 cases and the substatus was used in 45 of these. We detected that the substatus was misused in 8 cases, which represent an 8% of the total cases and 18% of the uses of the substatus.

Brazil

Available data of Brazil has 289 cases, and the substatus is used in 82 cases of the total in first and second levels. We found evidence of misusage of the substatus in 25 cases, it represent an 8% of the total cases.

France

It has 253 cases, and the substatus was used in 98 of these. After the analysis we conclude that in 37 cases the substatus was misused, it represent a 15% of the total cases.

United States

US has 737 cases. The substatus was used in the 3 levels of attention, in 201 cases. After the analysis we could conclude that there were 30 cases which present a misuse of the substatus. It represent a 18% of the total cases.

Belgium

It has 452 cases and the substatus is used in first and second level, in 161 cases. We found cases with misuse of the substatus in first level (61 cases) and second level (8 different cases), adding 69 cases which represent 15% of the total.

India

Available data of India has 402 cases, and the substatus was used in 273 of these. we found evidence of misuse in the 3 level of attention, 57 cases in first level, 63 cases in second level and 3 cases in the third. These 123 cases represent a 31% of the total cases.

Poland

Poland has 1.725 cases in the available data, and the substatus was used in every level in a total of 618 cases. We detect a serious case of misuse in the first level, with 425 of 587 cases (72% of the use were a misuse). In the second level we found only 3 different cases with misusage of the substatus and in the third level were no misusage. The total cases with misusage (428) represent a 25% of the total cases and 69% of the cases which the substatus was used.

Sweden

The most large available data, it has 2.483 cases. The substatus was used in 523 cases in the three level of attention. we found evidence of 227 cases with misusage of the substatus, that represent a 9% of the total cases.

2.3.4 Answer to the propose question

Using the available data, we could find evidence of the misuse of the substatus wait-user in several countries.

Table 6. Result of the data analysis

Value	*Total cases*	*Cases which use substatus wait-user*	*Cases with misuse of the substatus*	*Percentage of misuse of the total cases*	*Percentage of misuse of the case which use the substatus*
Sweden (se)	2.483	722	227	9%	31%
Poland (pl)	1.725	618	428	25%	69%
India (in)	402	149	123	31%	45%
Belgium (be)	452	161	69	15%	43%
United States (us)	737	201	30	4%	18%
France (fr)	253	98	37	15%	38%
Brazil (br)	298	82	25	8%	30%
China (cn)	96	45	8	8%	18%
Netherlands (nl)	57	25	14	25%	56%
Russia (ru)	45	8	0	0%	0%
Great Britain (gb)	21	4	4	19%	100%
Germany (de)	2	2	2	100%	100%
Australia (au)	24	6	0	0%	0%
Japan (jp)	13	6	4	31%	67%
Malaysia (my)	6	2	1	17%	50%
Korea (kr)	22	2	0	0%	0%
Thailand (th)	6	1	0	0%	0%
TOTAL	**6.642**	**2132**	**972**	**15%**	**46%**

Looking at the total results, we expect that, at least, the countries have a percentage of misuse around the average of 15 %, but we can see that in 7 countries the misuse is higher than the average.

At the same time, we can see the percentage of use of the substatus in the cases and notice important difference, the result in the total data is 32% (2132/6642), but in case of China, Netherlands and Japan, this amount is higher than 40%.

We conclude that there are a misuse of the substatus, and recommend to make a deeper analysis in Poland, India, Netherlands and Japan, which are the countries with the most high percentage of wrong utilization of the substatus.

2.4 Process conformity per organization

Volvo IT organization is spread in organizations, where the most important by size are *Org line A2* y *Org line C*.

2.4.1 Proposed questions

It would be interesting to see how conform or how much in line every organization is with the incident and problem management processes.

2.4.2 Data selected and log construction

In order to answer the proposed questions, we considered both incident and problem process. We used two different log filtered by organizational area (A2 and C).

Incident event log
We considered the same filters used in *2.3.2*. i.e.

- Delete cases with attention line '2nd-3rd'. because it has no logical sense.
- Keep only cases with first activity in the first attention line.
- Keep only cases with end activity with status complete, no matter the attention line.

Using these filters we got a log with 6.693 cases and 59252 events.

We could notice that Volvo IT organization is spread in 24 organizations, where A2 and C together have a participation of 84% in the total log.

Table 7. Organizations participation in the complete incidents log

Value	*Frequency*	*Relative frequency*
Org line C	40159	67,78%
Org line A2	9542	16,10%
Org line B	4422	7,46%
Other	2249	3,80%
Org line V7n	879	1,48%
Org line V2	571	0,96%
Org line V11	446	0,75%
Org line G2	186	0,31%
Org line G4	156	0,26%
Org line G1	143	0,24%
Org line E	110	0,19%
Org line V7	79	0,13%
Org line V5	74	0,12%
Org line V8	52	0,09%
Org line V10	38	0,06%
Org line V3	29	0,05%
Org line F	26	0,04%

Org line H	26	0,04%
Org line D	21	0,04%
Org line V1	15	0,03%
Org line I	10	0,02%
Org line V9	10	0,02%
Org line G3	6	0,01%
Org line V	3	0,01%

We filtered this log by organizations, creating two logs for A2 and C organization as it was required.

Problems log

We considered both closed and open problem logs. The data was merged in just one event log with 1.841 cases in 578 different variants and 9.011 events.

In the same way that in the former case, we identified more organizations than the two required in the question, but Org line C and Org line A2 where the most popular, accounting a 68% of the total log.

Table 8. Organizations participation in the complete problems log

Value	*Frequency*	*Relative frequency*
Org line C	3828	42,48%
Org line A2	2378	26,39%
Org line G3	1288	14,29%
Org line G4	940	10,43%
Org line B	270	3%
Org line V2	109	1,21%
Org line F	68	0,75%
Org line V11	35	0,39%
Org line D	32	0,36%
Org line V7n	27	0,30%
Org line G1	18	0,20%
Org line V5	10	0,11%
Other	5	0,06%
Org line V4	2	0,02%
Org line A1	1	0,01%

We could notice in the process diagram obtained from DISCO, that the different attention lines were separated without flows from one level to another. For this reason, we did not use the start/end points filter, considering every variant in the log.

This event log was filtered by organization, creating two different log to *Org line A2* and *Org line C*.

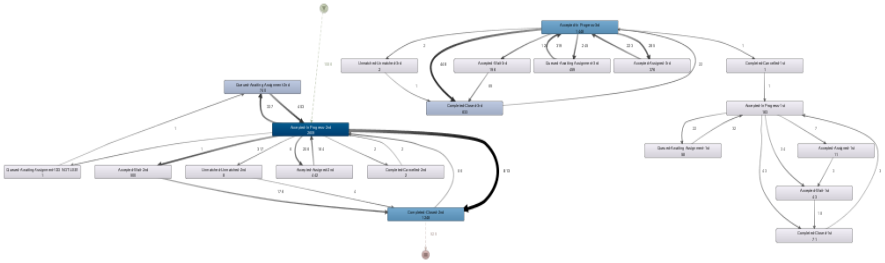

Figure 5. Process diagram obtained from DISCO. Even though the diagram is too small, we can see as every level is disconnected from the others. The least used, in the right of the diagram, is the first level.

2.4.3 Data analysis

Incidents

The Org line A2 has 975 cases, with 453 variants and 7568 events. The resources involved are 153. The event log refers to incidents of 205 different products.

The events in the timeline of the log are concentrated at the end of the period as is shown in Figure 6.

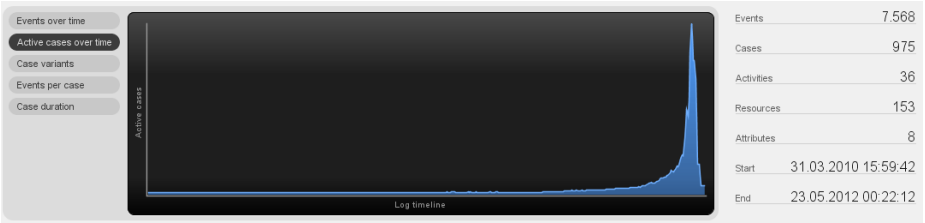

Figure 6. Active cases overtime in Org line A2.

Table 9. Frequency of the activity in Org line A2 event log.

Value	*Frequency*	*Relative frequency*
Accepted/In Progress/1st	1912	25,40%
Queued/Awaiting Assignment/1st	1090	14,48%
Accepted/In Progress/2nd	1006	13,36%
Queued/Awaiting Assignment/2nd	881	11,70%
Accepted/Wait - User/1st	314	4,17%
Completed/Resolved/1st	276	3,67%
Completed/Closed/1st	266	3,53%
Completed/Resolved/2nd	265	3,52%
Completed/Closed/2nd	249	3,31%
Accepted/In Progress/3rd	243	3,23%
Queued/Awaiting Assignment/3rd	161	2,14%

Accepted/Wait - User/2nd	147	1,95%
Accepted/Assigned/1st	117	1,55%
Accepted/Assigned/2nd	112	1,49%
Accepted/Wait/2nd	76	1,01%
Completed/Resolved/3rd	63	0,84%
Completed/Closed/3rd	59	0,78%
Accepted/Wait/1st	59	0,78%
Accepted/Wait - User/3rd	56	0,74%
Accepted/Assigned/3rd	33	0,44%
Completed/In Call/1st	25	0,33%
Accepted/Wait/3rd	24	0,32%
Accepted/Wait - Implementation/3rd	23	0,31%
Accepted/Wait - Implementation/1st	20	0,27%
Accepted/Wait - Implementation/2nd	15	0,20%
Accepted/Wait - Customer/2nd	14	0,19%
Accepted/Wait - Vendor/1st	7	0,09%
Accepted/Wait - Customer/1st	7	0,09%
Accepted/Wait - Vendor/2nd	6	0,08%
Accepted/Wait - Vendor/3rd	2	0,03%
Accepted/Wait - Customer/3rd	1	0,01%

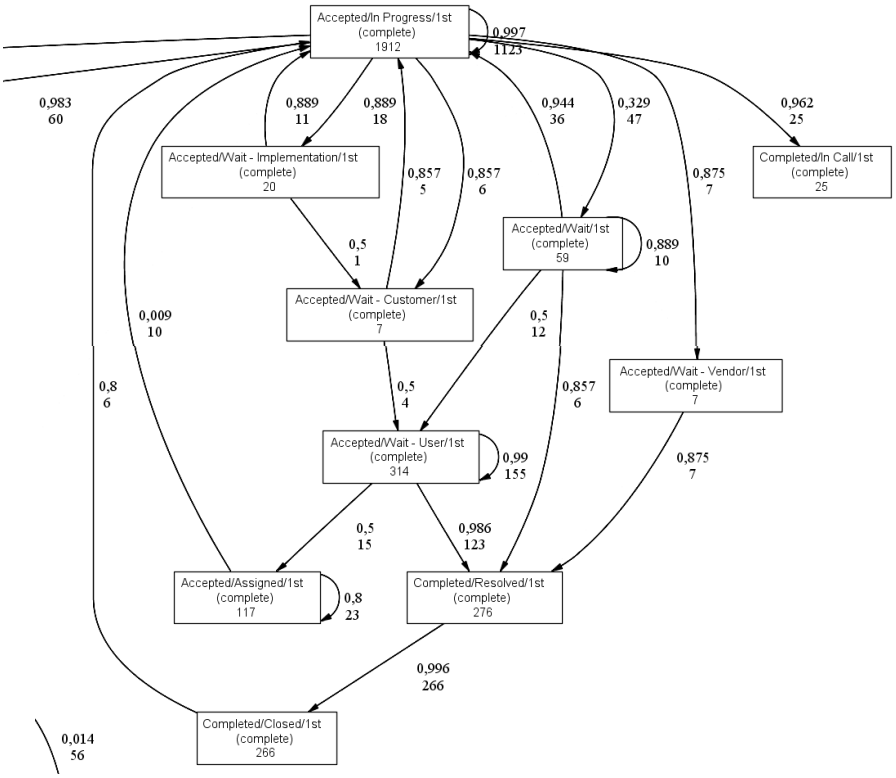

Figure 7. Process flow of the first line in the Org line A2.

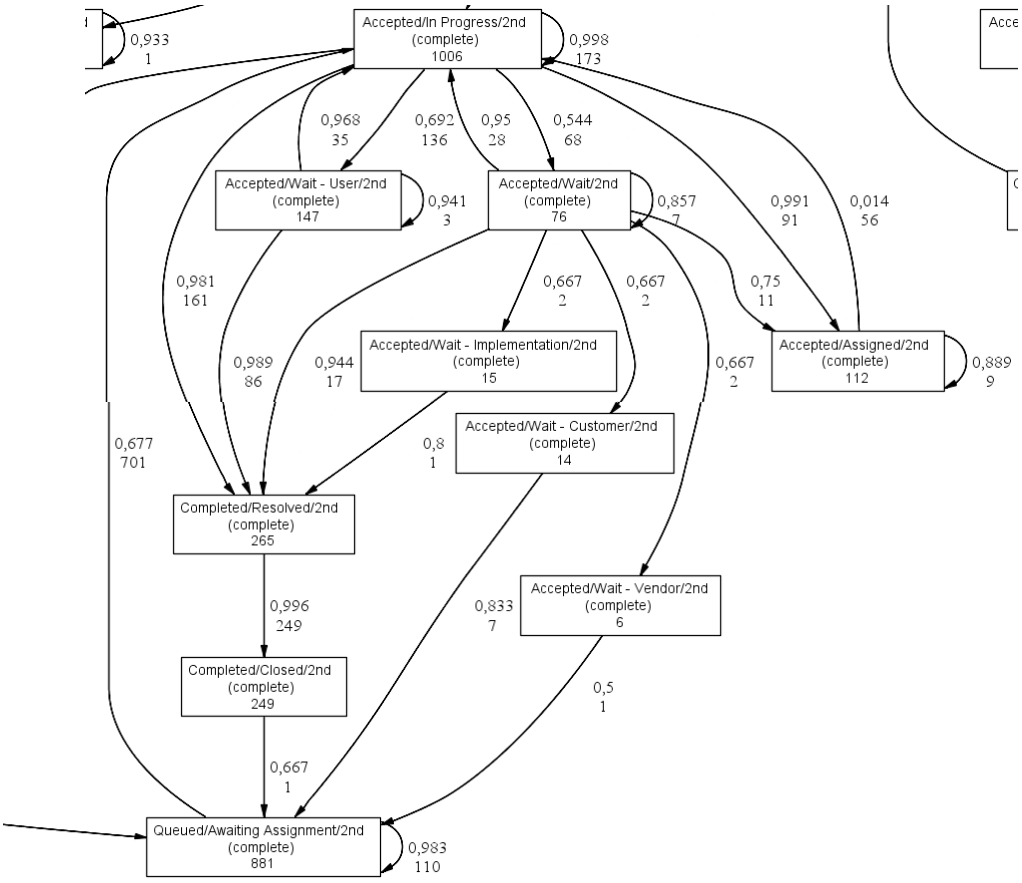

Figure 8. Process flow of the second attention line in the Org line A2.

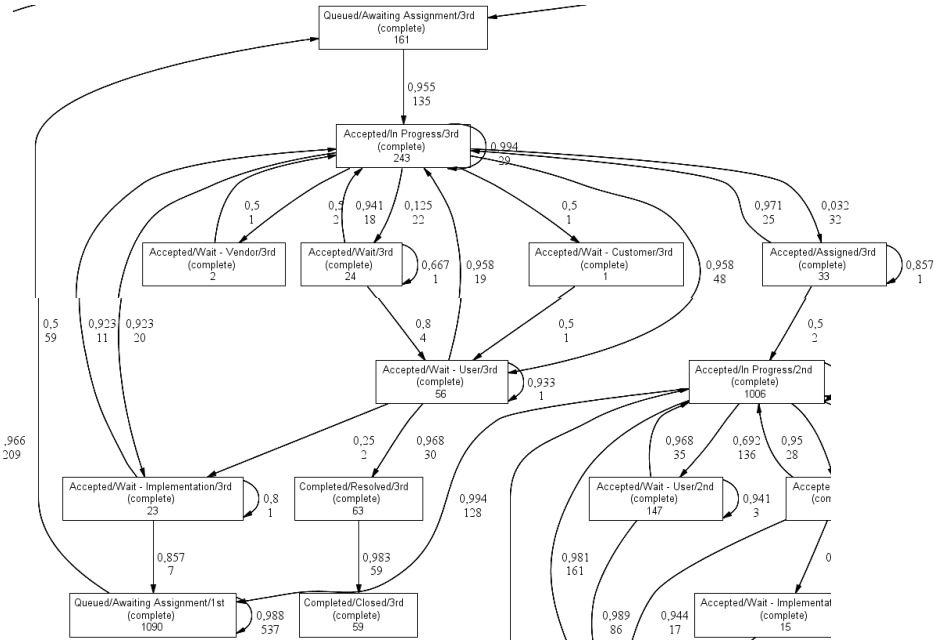

Figure 9. Process flow of the third attention line in the Org line A2.

The **Org line C** has 539 cases, with 272 variants and 4107 events. The resources involved are 27. The event log refers to incidents of 123 different products.

The events in the timeline of the log are concentrated at the end of the period as is shown in Figure 10.

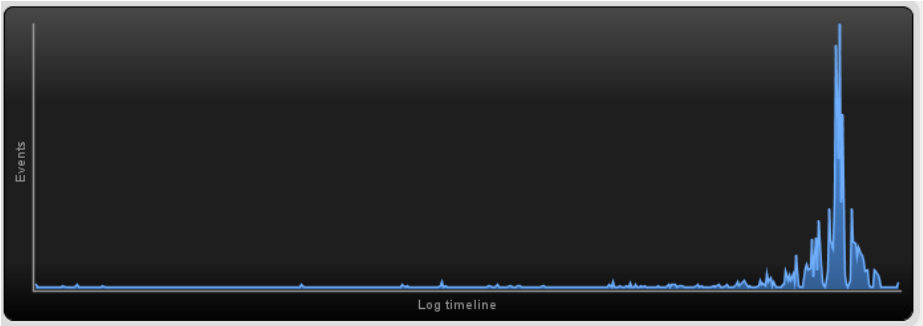

Figure 10 Active cases overtime in Org line C.

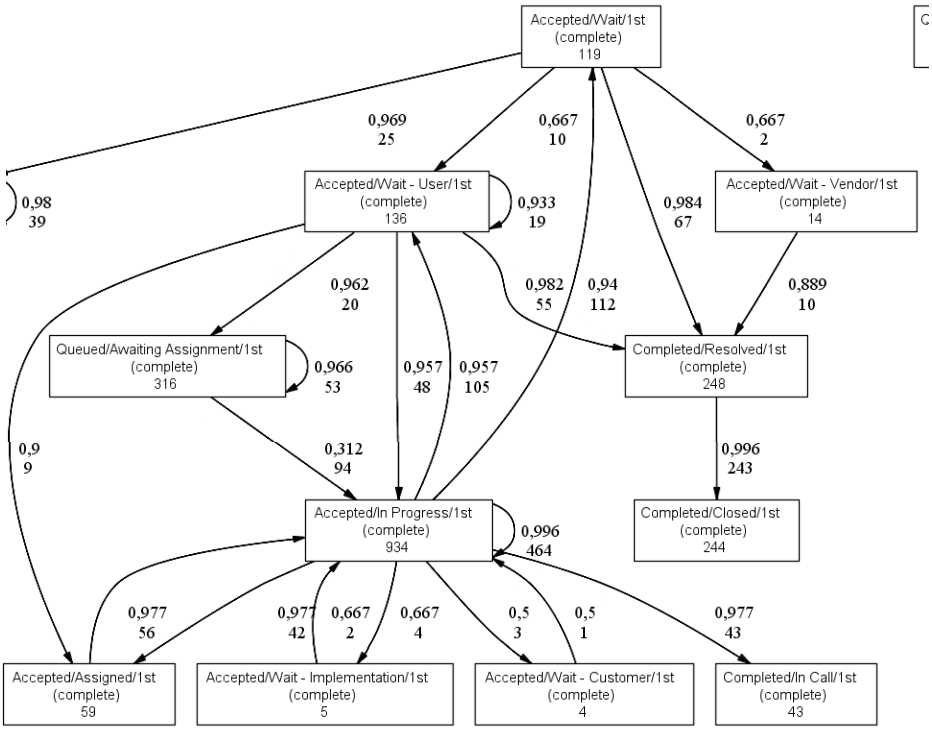

Figure 11. Process flow of the first attention line in the Org line C.

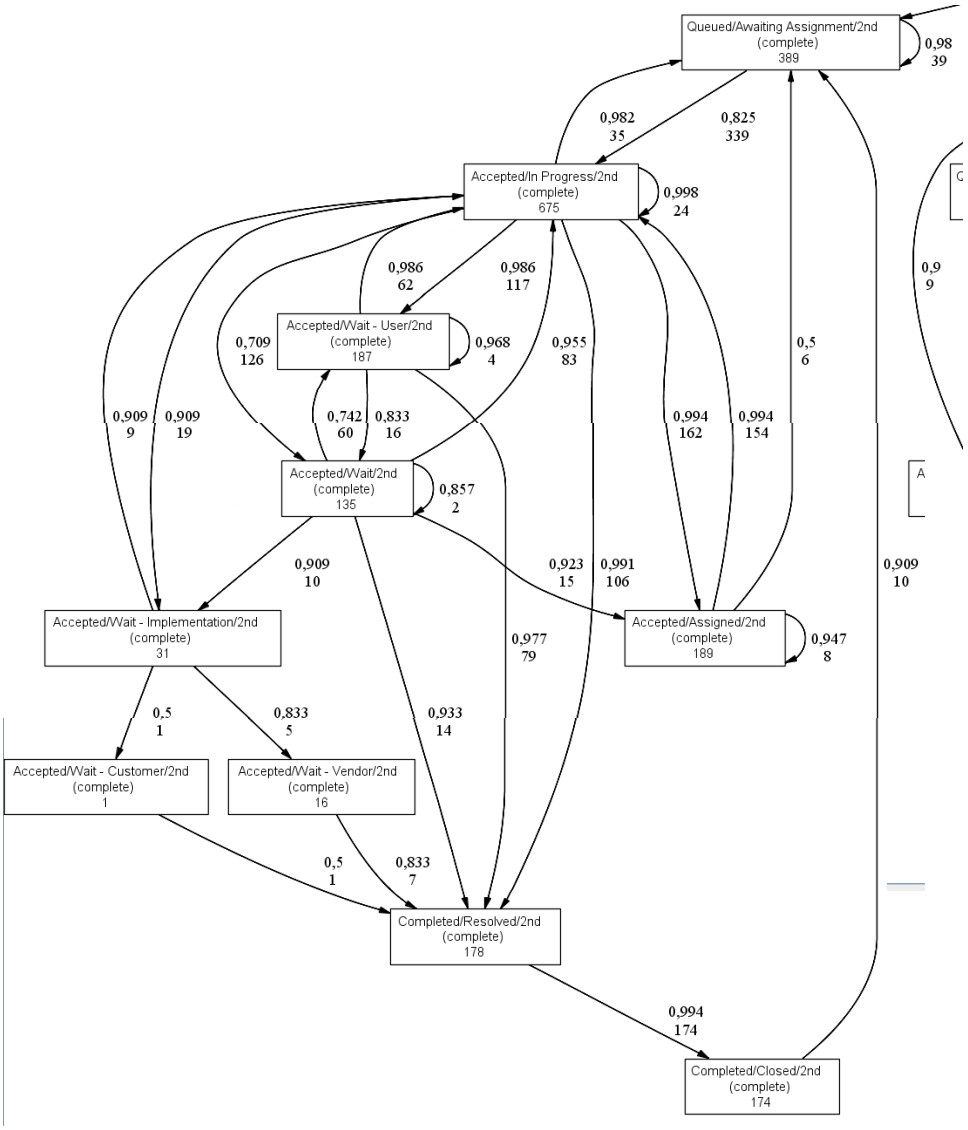

Figure 12. Process flow of the second attention line in the Org line C.

We notice that the size of the log for A2 line is bigger than the log of C line. Due that, the diagram of the process are more complex in the first case.

In A2 line, we have 975 cases, and 401. are entering directly to the second level. In this case, 125 cases are scaled to the third level, it represents a 13% of the total cases.

In C line, we have 539 cases, and 321 enter to the first line in the *accepted/in_progress/1st* activity, that represent a 60%, the other 40% enter to the second line directly. Just two cases enter to the third line.

In both cases, A2 and C, we noticed that the cases are not going in the expected flow, entering directly to the second level.

Problems

In a first review of the log, we could see the Org line C is bigger than org line A in quantity of cases (762 and 598 respectively). The states observed are similar (17 and 18).

Figure 13 and Figure 14 show an overview of the logs.

Figure 13. Overview of the A2 event log.

Figure 14. Overview of the A2 event log.

We use the heuristic miner algorithm in order to obtain the subjacent process of the log. For simplicity of the analysis, we divide the log in the 3 attention lines, considering there were no flows between lines.

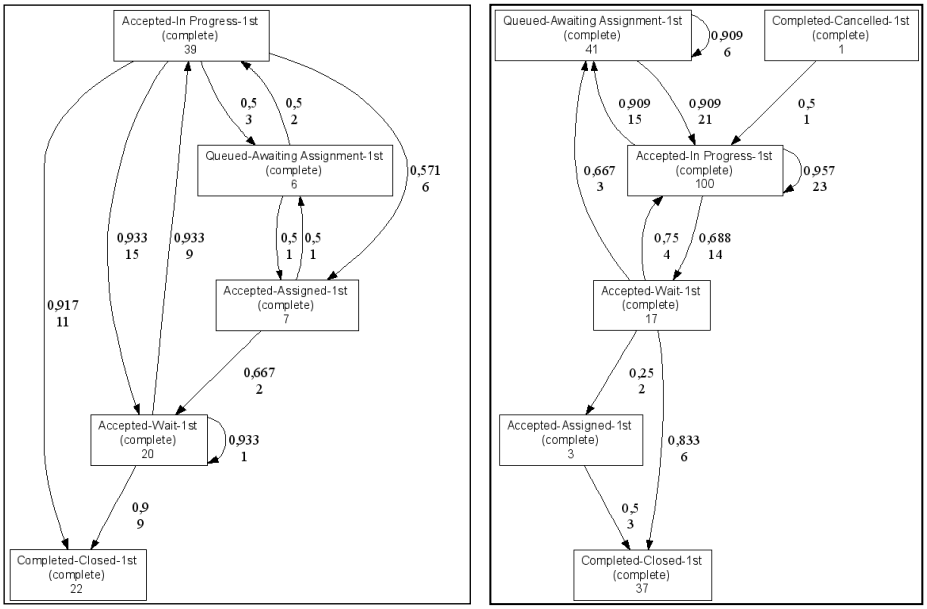

Figure 15. Process model using Heuristic Miner algorithm to the first line (the smaller one). To the left, the diagram to the process for Org line A2, to the other hand, the process for Org line C.

Comparing first line in both line C and line A2, we first can see the size difference, where A2 has 28 cases and 94 events, and C has 69 cases and 199 events.

About the process execution, both organization present the same activities (with the exception of the activity *completed/cancelled* which have just one occurrence).

In the organization line A2, the process discovered make sense. The first activity is *accepted/in_progress* (in an 85% of the cases), followed by *accepted/assigned* if the case were assigned or *queued/awaiting_asignment* if not. After that, the case could change status to *accepted/wait* or go back to former a status. At the end, the case goes to *completed/closed* activity.

From the waiting status, half of the cases go back to the status *in_progress*. It is interesting to see the non linearity of the process, with two way flows between activities (status).

In the organization C, as in A2, the first activity is *accepted/in_progress* in most of the cases (75%). This activity is followed by *queued/awaiting_assignment* and coming back to *accepted/in_progress* without an assignment. A case also could go from the first activity to *accepted/wait* directly.

From *accepted/wait* a case could be assigned or closed, but also could go back to the status *queued/awaiting_assignment*, this behavior seem like there is a unnecessary loop. In the variant where the case goes from *accepted/wait* to *accepted/assigned* is not clear why is used this substatus between the queue and the assignment, there being no direct flow these activities.

Finally, the case is closed with the status *completed/closed*.

Appendix A: Details from point 2.2.3 (*Ping-Pong* Behavior)

The first five countries present problems for the correct analysis due to the large quantity of nodes and arches, for this reason is necessary create a partition of these countries. At any rate, the analysis will be similar to the next group of countries.

1) *'country' = 'SE' and 'country' = 'se'* (Sweden)

Sweden has 316 STs, 307 of this belong to a big cluster. We find 8 clusters in total. This country has the most participation in the log with 32% of the cases and 28% of the events.

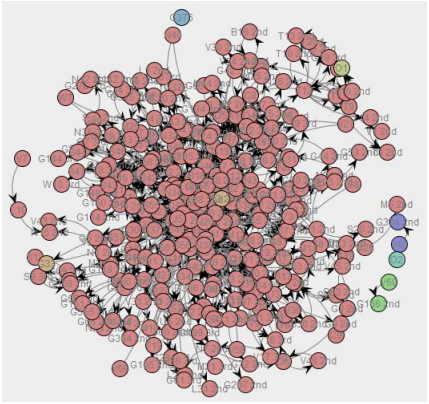

Figure 16. Process model obtained using HoW for Sweden.

2) *'country' = 'pl'* (Poland)

Poland has 163 STs. In this country we find 4 clusters. Poland has a 22% of the cases and 22% of the events.

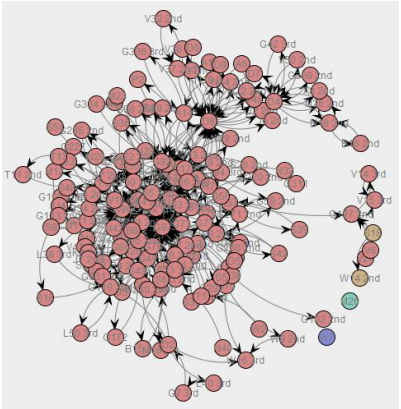

Figure 11. Process model obtained using HoW for Poland.

3) *'country'* = *'in'* (India)

In this country we find 158 STs classified in 2 clusters. India has a 5% of the cases and 9% of the events.

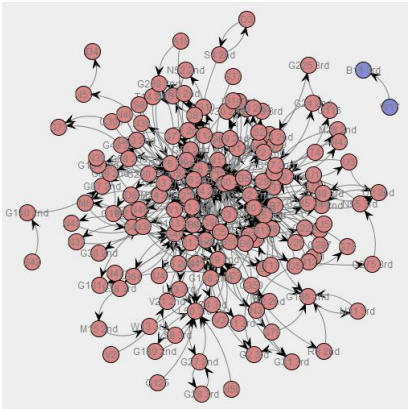

Figure 12. Process model obtained using HoW for India.

4) *'country'* = *'be'* (Belgium)

Belgium has 123 STs. In this country we find 2 clusters. This country has 5% of the cases and 8% of the events.

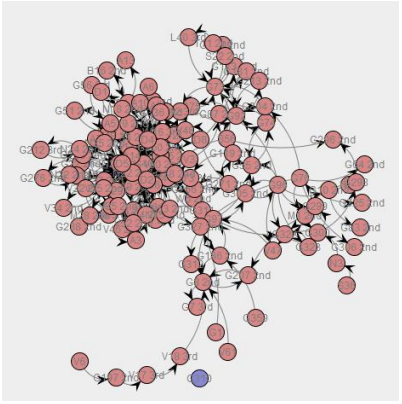

Figure 13. Process model obtained using HoW for Belgium.

5) *'country'* = *'us'* (United States)

In this country we find 110 STs. United States has 2 clusters. This country has 9% of the cases and 7% of the events.

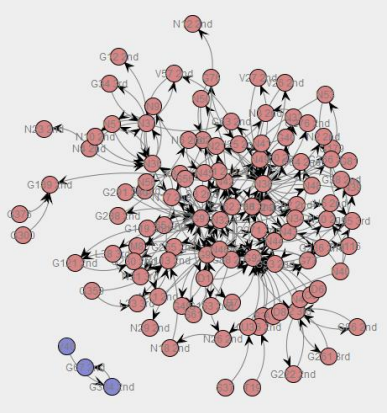

Figure 14. Process model obtained using HoW for United States.

The next group of countries has the correct size for analyze if there is presence of Ping Pong behavior. The size is less to 100 STs.

6) *'country'* = *'fr'* (France)

France has 82 STs. In this country we find various clusters and if visible the Ping Pong behavior in various nodes. This country has 3% of the cases and 3% of the events.

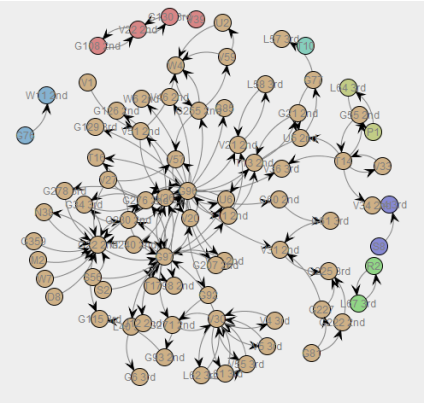

Figure 15. Process model obtained using HoW for France.

7) *'country'* = *'br'* (Brazil)

Brazil has 30 STs. In this country we find 2 clusters and the Ping Pong behavior is present in so much nodes. This country has 3% of the cases and 3% of the events.

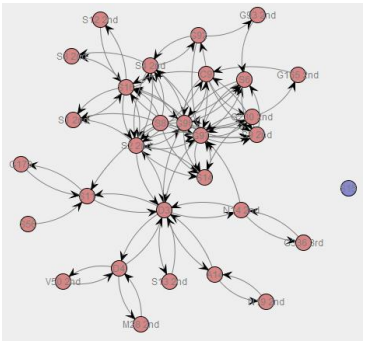

Figure 16. Process model obtained using HoW for Brazil.

8) *'country'* = *'nl'* (Netherlands)

Netherlands has 14 STs. In this country the Ping Pong behavior is present in so much nodes.

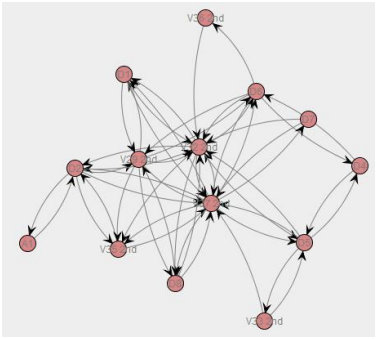

Figure 17. Process model obtained using HoW for Netherlands.

9) *'country'* = *'cn'* (China)

China has 26 STs. This is a special case where we find 2 clusters of similar size. The Ping Pong behavior is present in most of the nodes.

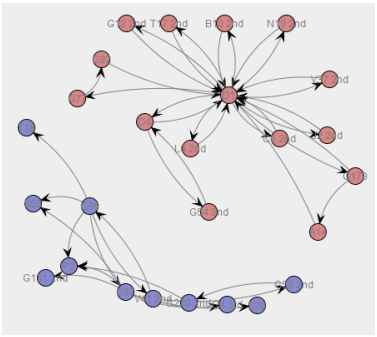

Figure 18. Process model obtained using HoW for China.

10) *'country'* = *'kr'* (Korea)

Korea has 19 STs. The Ping Pong behavior is minimally present.

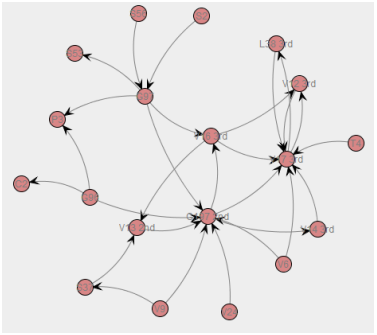

Figure 19. Process model obtained using HoW for Korea.

11) *'country'* = *'ru'* (Russia)

Russia has 9 STs divided in 2 clusters. The Ping Pong behavior is present just in one cluster conformed for 2 STs.

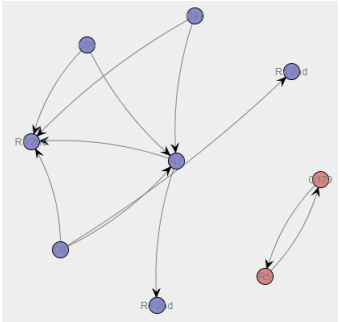

Figure 17. Process model obtained using HoW for Russia.

12) *'country'* = *'gb'* (Great Britain)

Great Britain has 23 STs divided in 2 clusters. The Ping Pong behavior is present in both cluster strongly.

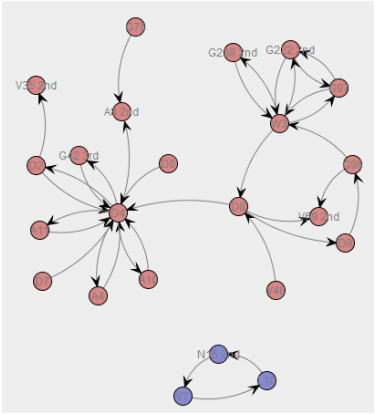

Figure 18. Process model obtained using HoW for Great Britain.

13) *'country'* = *'au'* (Australia)

Australia has 8 STs divided in 2 clusters. The Ping Pong behavior is minimally present just in one cluster.

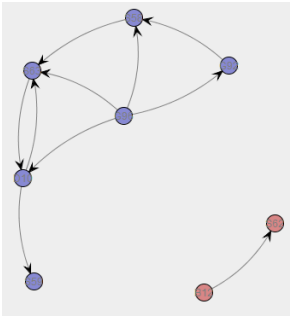

Figure 19. Process model obtained using HoW for Australia.

14) (0)

0 has 25 STs divided in various clusters. The Ping Pong behavior is present just in one cluster but of important form.

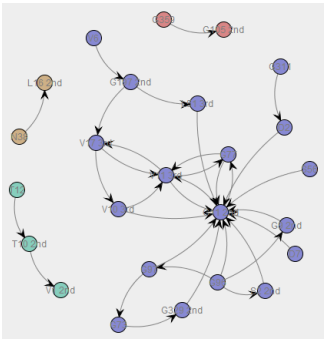

Figure 20. Process model obtained using HoW for 0.

15) *'country'* = *'jp'* (Japan)

Japan has 11 STs divided in 2 clusters. The Ping Pong behavior is present minimally just in one cluster.

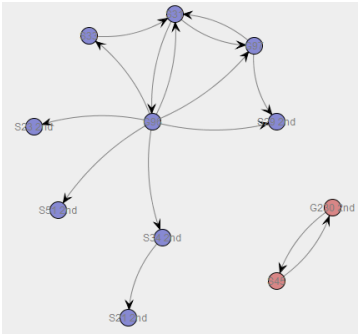

Figure 21. Process model obtained using HoW for Japan.

16) *'country'* = *'de'* (Germany)

Germany has 5 STs. The Ping Pong behavior is present in all STs.

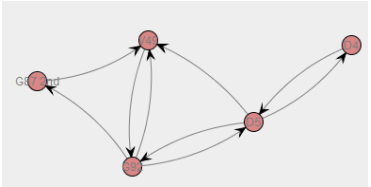

Figure 22. Process model obtained using HoW for Germany.

The last group of countries just has less of 5 STs. In this group just Thailand presents Ping Pong behavior.

17) *'country'* = *'my'* (Malaysia)

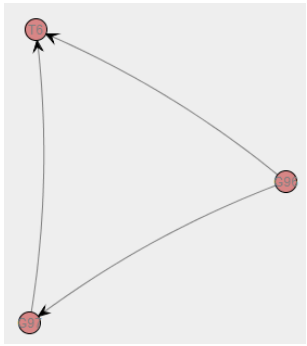

Figure 23. Process model obtained using HoW for Malaysia.

18) *'country'* = *'th'* (Thailand)

Figure 24. Process model obtained using HoW for Thailand.

19) *'country'* = *'cl'* (Chile)

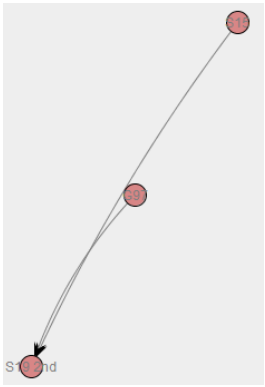

Figure 25. Process model obtained using HoW for Chile.

20) *'country'* = *'tr'* (Turkey)

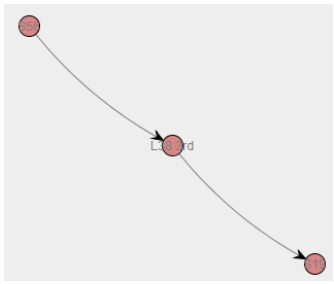

Figure 26. Process model obtained using HoW for Turkey.

21) *'country'* = *'pe'* (Peru)

Just has one ST.

22) *'country'* = *'ca'* (Canada)

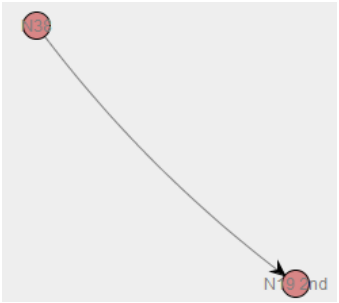

Figure 27. Process model obtained using HoW for Canada.

Appendix B: Details from point 2.3.3 (substatus *wait-user* analysis)

1) *'country'* = *'th'* (Thailand)

There are just 6 cases in Thailand available information and the substatus *'wait-user'* was used just once. The case detected was open for 5 days and the *'owner activity'* was *'Thachuda'*.

The diagram obtained using Heuristic Miner algorithm and Alpha Miner algorithm are show in Figure 28Figure 29.

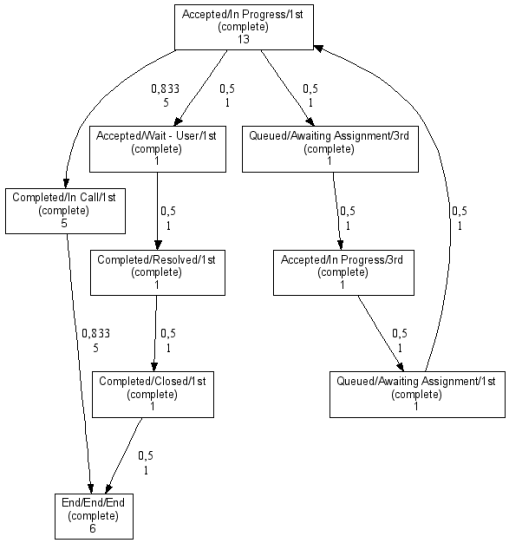

Figure 28. Process model obtained using Heuristic Miner Algorithm

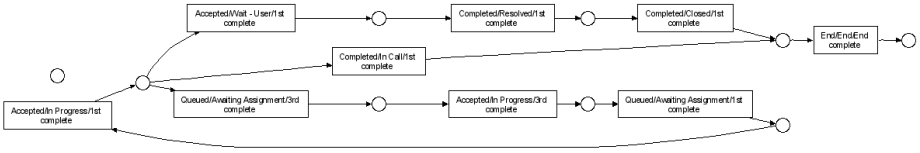

Figure 29. Process model obtained using Alpha Miner Algorithm

2) *'country' = 'kr'* (Korea)

The available data of Korea's Incidents has 22 cases with 7 variants and 140 events. The cases are solved by the first attention line, without participation of the second and third lines.

The substatus *wait-user* is used two times in two different cases, both cases between *Accepted/In_progress* and *Completed/Resolved* status.

We do not detect a misuse of the substatus *wait-user*.

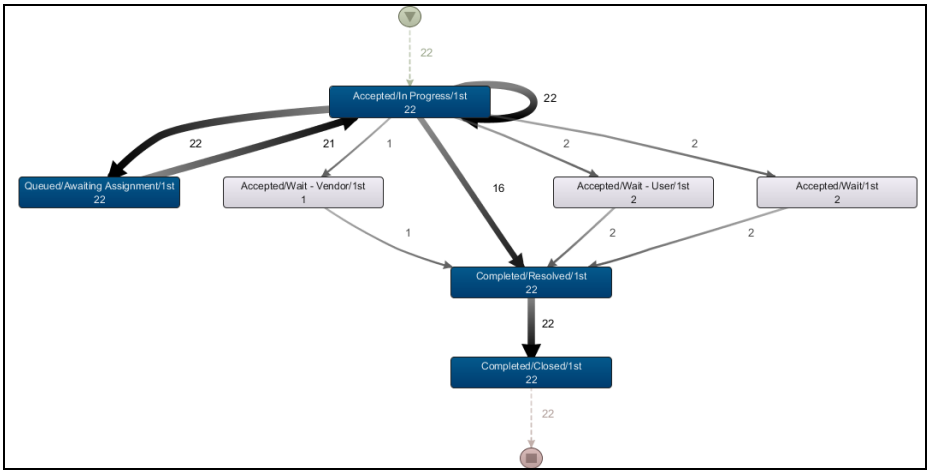

Figure 30. Process model obtained from Korea's incident log using DISCO. Numbers indicate the case frequency of the activities.

3) *'country' = 'my'* Malaysia

Malaysia has 6 cases with 47 events. the process diagram obtained with DISCO is shown in Figure 31 and Figure 32.

We can see the cases are solved in the first level without participation of the second and third ones. This was expectable because the in the ideal world, most of the cases should be solved in the first level.

The model obtained in disco is similar to a flower model, but is the same obtained using Heuristic Miner in ProM 5.2 (Figure 33).

In particular, the wait-user substatus is used 3 times in two cases, by the users Muthu (2 times) and Alex (once), both in Org line C. In the first case detected (Muthu), the substatus was used between two occurrences of the first activity. We understand this behavior as a misusage of the substatus.

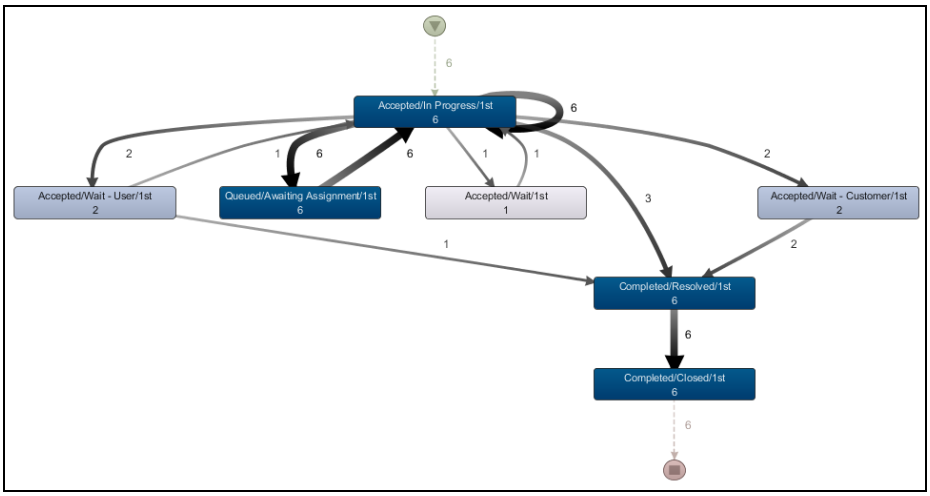

Figure 31. Process model obtained from DISCO, showing the case frequency of the nodes.

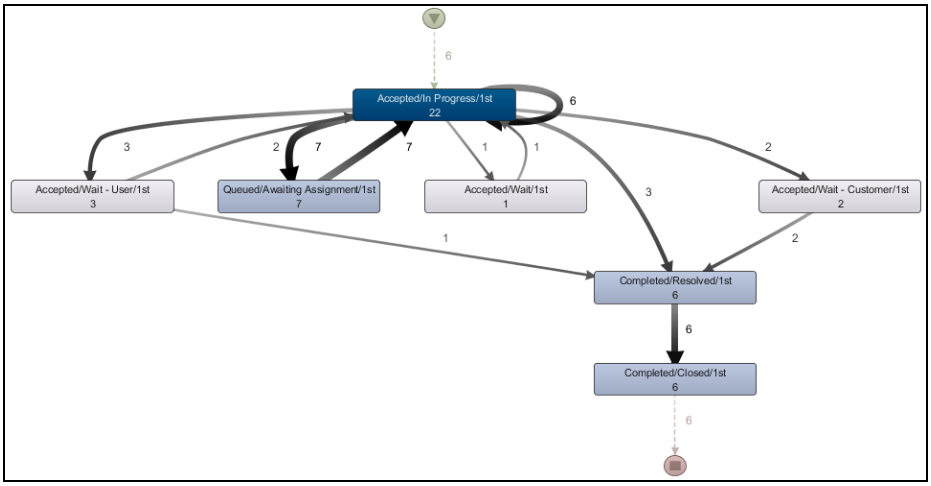

Figure 32. Process model obtained from DISCO, showing the absolute frequency of the nodes.

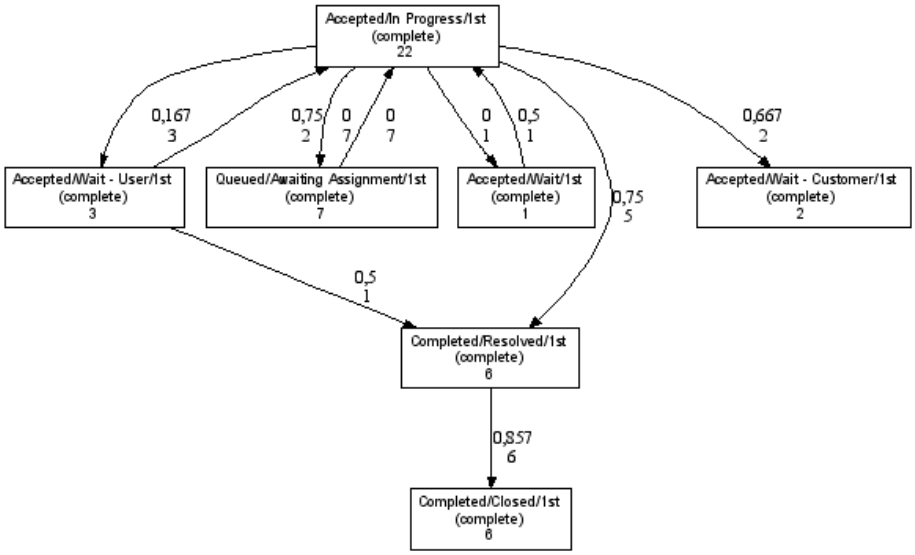

Figure 33. Process model obtained using Heuristic Miner Algorithm in ProM 5.2.

4) *'country' = 'jp'* (Japan)

Japan present 13 cases with 12 variants and 150 events.

As we expected, the most of the cases are solved in the first level and there are not cases in the third level. Just one case is pushed to front from second to first level.

Figure 34 shows the complete process model obtained for Japan using Heuristic Miner Algorithm.

Figure 35 shows a zoom of the substatus *wait-user* in the first level, where we can see the substatus was used in five cases, just once it was used previous to complete a case, but 4 times it were used as a pause between two occurrence of *accepted/in_progress* status, being this a misusage of the substatus.

Table 10. Users with misusage of the substatus *wait-user*

Value	*Frequency*	*Relative frequency*
Lingaraj	2	50%
Uguisu	1	25%
Yukie	1	25%

In the second level, the case is used just once, in one case. The use of the substatus is corrected. Figure 36 shows the zoom of the *wait-user* substatus in the second line.

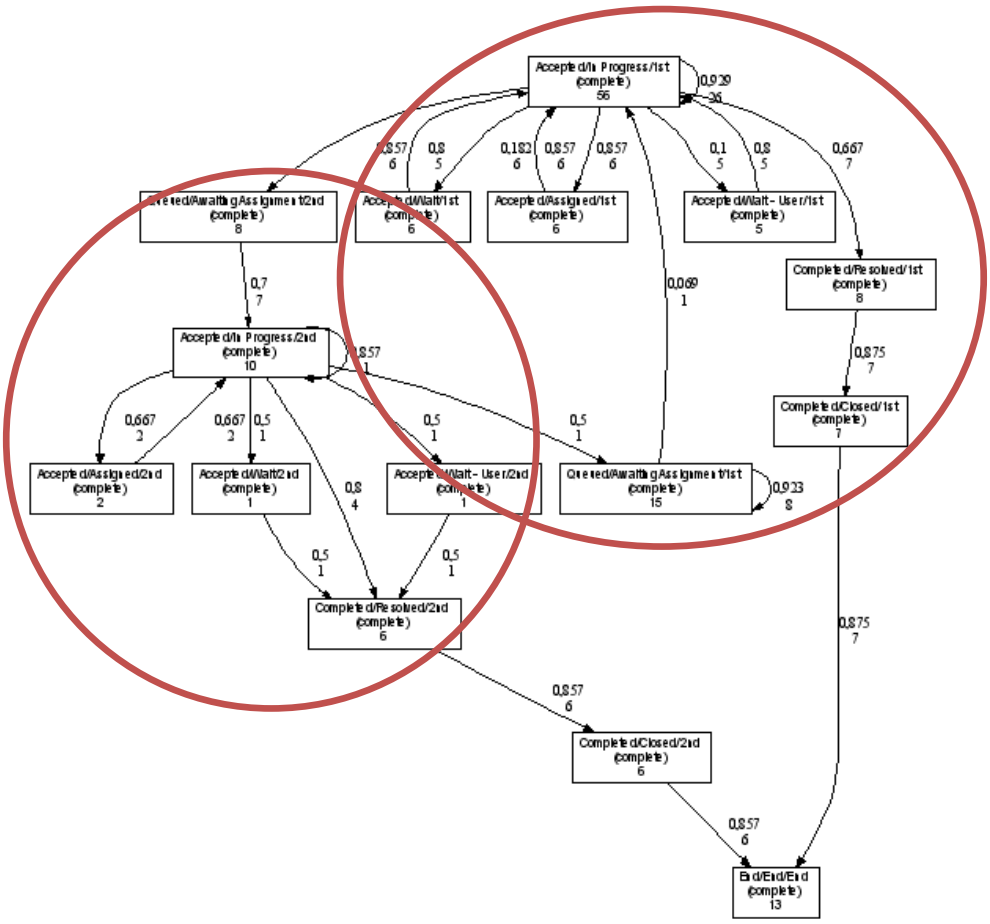

Figure 34. Japan process model obtained using Heuristic Miner in ProM 5.2.

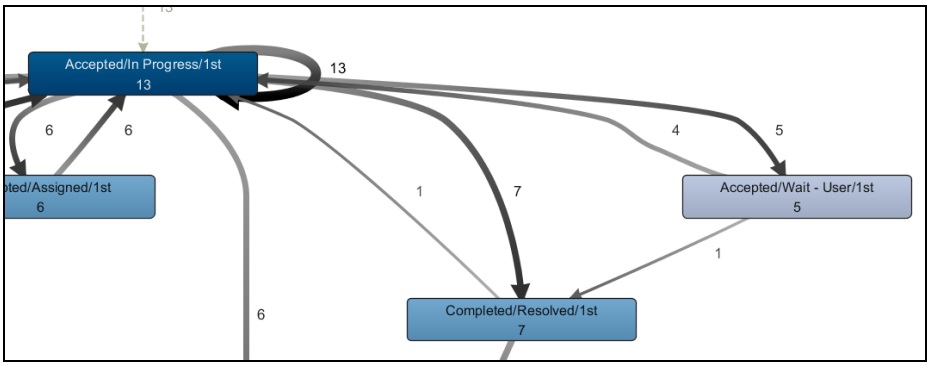

Figure 35. Zoom of the *wait-user* status in the first level of Japan. Numbers indicate the case frequency of the nodes.

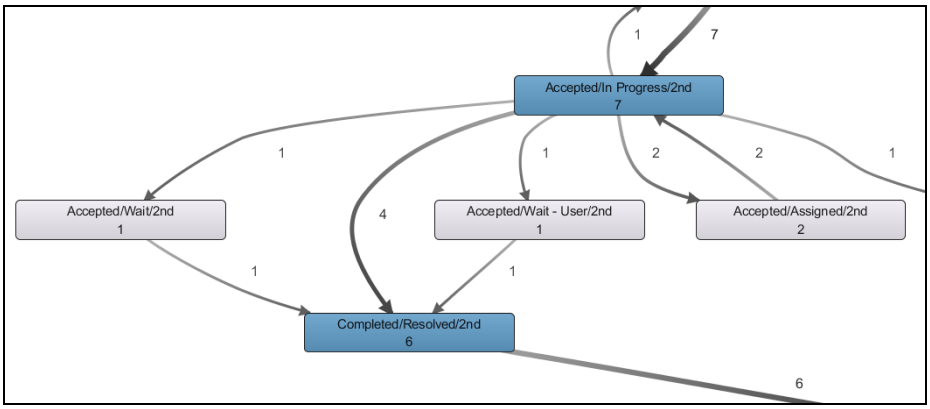

Figure 36. Zoom of the *wait-user* status in the second level of Japan. Numbers indicate the case frequency of the nodes.

5) *'country'* = *'au'* (Australia)

Australia present 24 cases with 16 variants and 232 events. Every case is solved in the first line.

The wait-user substatus is used in 6 cases, five of them is used between the status *Accepted/In_Progress* and *Completed/Resolved* and once is used between the status *Accepted/In_Progress* and *Queued/Awaiting_Assignment*. We do not detect sequences that revels misusage of the substatus as "pause" between two occurrences of the same status.

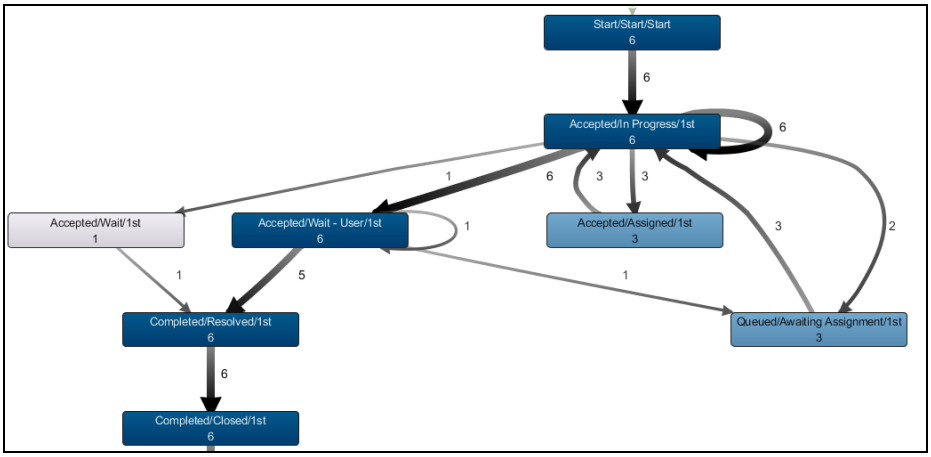

Figure 37. Zoom of the *wait-user* status in the first level of Australia, obtained from DISCO. Numbers indicate the case frequency of the nodes.

6) *'country' = 'de'* (Alemania)

Germany as 2 cases with 55 events. In this country, before the analysis of the substatus wait-user, we noticed the cases overtime was too long, 70 and 160 days in each case. The walkthrough over the diagram showed as both of the cases are derivated from the first to the second line and pushed to front several times (3 and 5 times each case). Finally the cases are solved by the first line.

In the life time of the cases, every status are used several times, in particular the substatus wait-user is used 8 times, 7 in the first level and 1 in the second level. The substatus has different uses gave for the users. We detect the following situations:

- Used as a pause between two occurrences of the *Accepted/In Progress* status.
- Used between *Accepted/In Progress* and Accepted/Assigned (3 times in 1 case). This situation is clearly a misusage of the status.
- Used before Queued/Awaiting_asignment/2nd
- Used before Queued/Awaiting_asignment/1st

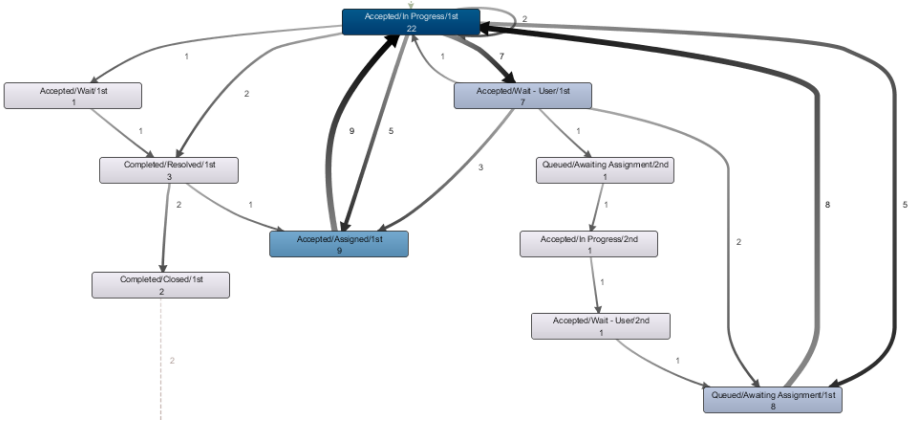

Figure 38. Process model of Germany obtained from DISCO.

7) *'country' = 'gb'* (Great Britain)

Great Britain has 21 cases with 18 variants and 284 events. The substatus is used 8 times (7 in the first level and 1 in the second one), in 4 cases.

In the first level, three times was used as a pause between two occurrence of the status *Accepted/In_progress/1st*. The other 5 times, the status was executed before:

- *Accepted/wait_customer/1st*,
- *Queued/Awaiting_assignment/2nd*,
- *Queued / Awaiting_assignment / 1st,*
- *Accepted/Assigned/1st,*
- *Completed/resolved/1st*

In the second level, the substatus is used just once, between the status *Accepted/In_progress/2nd* and *Queued/Awaiting_assignment/1st*, i.e. is a pause before push to front the case back to the first level.

In any case, we consider there is a misusage of the substatus in this 4 cases. The action_owner involved in this cases.

Table 11. Frequency of the action_owners using wait-user substatus.

Value	*Frequency*	*Relative frequency*
Sunny	4	50%
Krzysztof	2	25%
Marcin	1	12,50%
Steven	1	12,50%

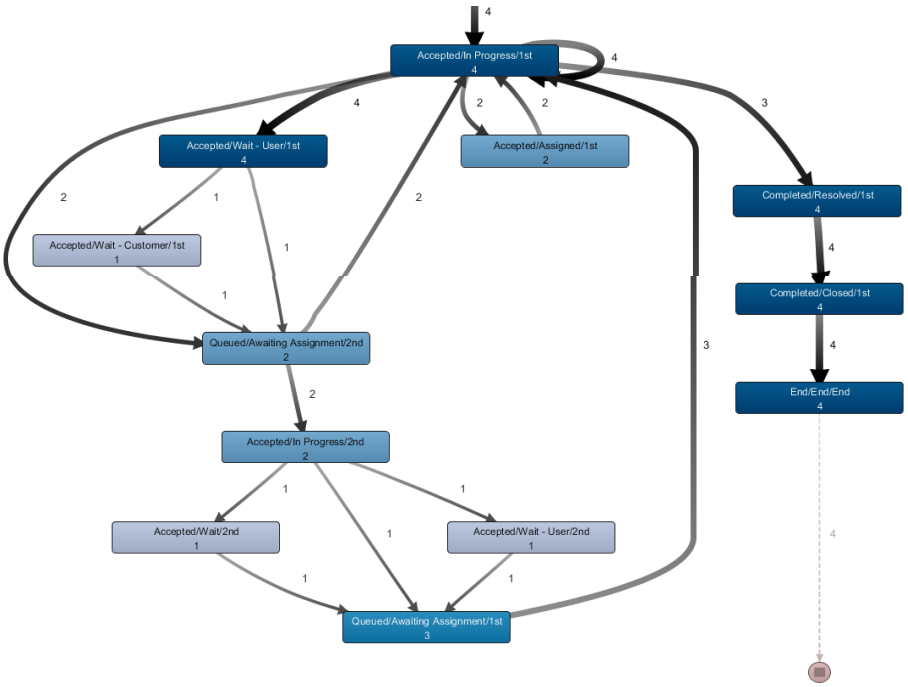

Figure 39. Process model of the cases where the substatus wait-user is used.

8) *'country' = 'ru'* (Russia)

Russia has 45 cases in 28 variants and 391 events. There are 8 cases where the substatus wait-user is used, 4 in the first level, 3 in the second and 1 in both levels.

The cases are solved in the first two levels without intervention of the third one. In the 45 total cases, 23 were solved in the 1st level and 22 cases were scaled to the second level of attention. We expected that a bigger part of the cases were solved in the first line.

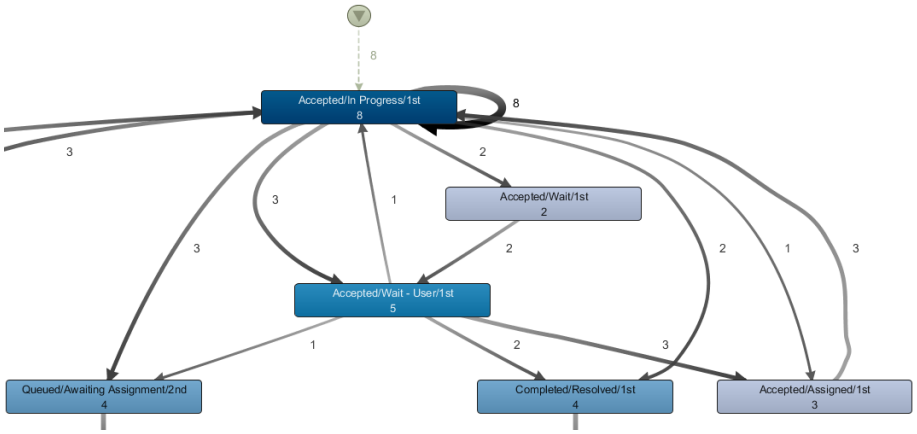

Figure 40. Detail of the wait-user node in the first level of attention.

We can see in the first level, that in 1 case there is evidence of the use the substatus as a pause, but this activity was performed in less than 5 minutes, that is why we do not considered it as misusage.

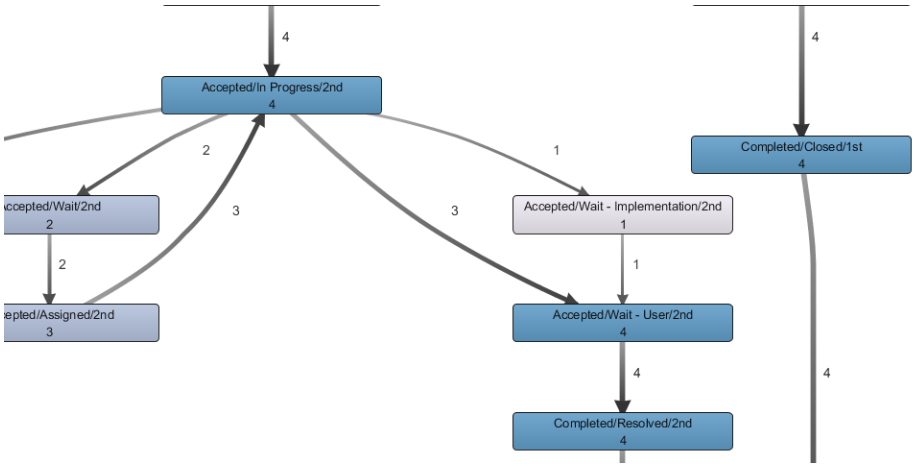

Figure 41. Detail of the wait-user node in the second level of attention.

In the second level, the substatus is used by 4 cases, and in every case it was before the closure of the case.

We can conclude that in Russia there are not misusage of the substatus.

9) *'country'* = *'nl'* (Netherlands)

In Netherlands we found 57 cases, 54 variants and 1.277 events. The cases are solved by first and second level without intervention of the third one.

In the first level, there are 57 occurrence in 25 cases where the substatus *wait-user* is used and in 14 of those (29 times) the use was as a pause between two occurrence of *Accepted/In_progress* status. We interpret this behavior as a misusage.

The *action_owner* involve in this behavior are shown in Table 12. It is remarkable as Olga has half of the occurrence of the substatus. We understand that Siebel is an automatic change of status performed by the system.

Table 12. Frequency in the use of the substatus wait-user

Value	*Frequency*	*Relative frequency*
Olga	26	45,61%
Siebel	6	10,53%
Annick	3	5,26%
Jan	3	5,26%
Abby	3	5,26%
Renaat	3	5,26%
Marco	2	3,51%
Marcin	2	3,51%
Katia	2	3,51%
Ilona	1	1,75%
Shuwen	1	1,75%
Zoi	1	1,75%
Brecht	1	1,75%
Meishan	1	1,75%
Alexandre	1	1,75%
Evy	1	1,75%

We found misusage in Russia, mainly performed by Olga.

10) *'country' = 'cn'* (China)

Available data of China has 96 cases, with 64 variants and 1.348 events. The relation between different lines is the following: There are 96 cases in the available data entering in the first level, from these, 52 are referred to the second level, that means a 46% of resoluteness of the first level, lower than we expected. One case is referred directly to the third level, skipping the second one.

Wait-user status is used 78 times (44 cases) in the first level. In 37 of these times (34 cases), it was followed by *Complete/Resolve* in the first level, showing a correct use of the substatus. In 10 times (8 cases) it was used as pause of the process between two *Accepted/In_progress* occurrences by the *action_owner* **Santosh**, **Kelly**, **May**, **Meishan**, **Peng** and **Max**.

In the second level, the substatus was used 4 times (3 cases where also were use in the first level), The substatus was used before a queued/awaiting_assignment/1st status, i.e. before a push to front action.

In the third level was used just once before complete the case.

We conclude there was misuse of the substatus in 8 cases from the 96 total cases. It represent an 8% of the cases.

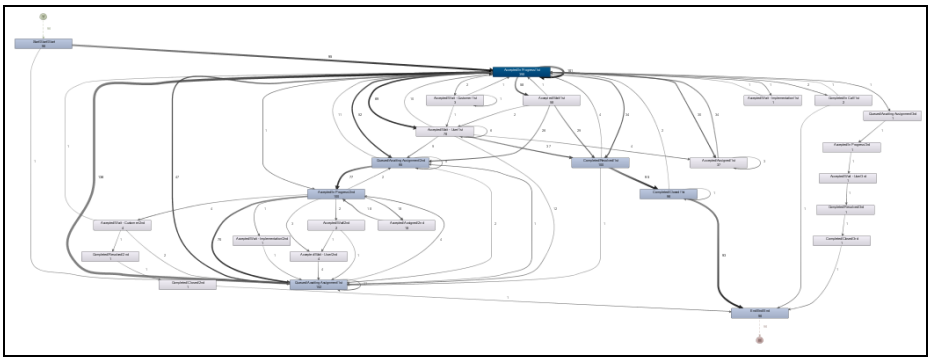

Figure 42. The process diagram shows the complexity of the flows. We understand this figure is unreadable, but we want is to show the complexity of the obtained model.

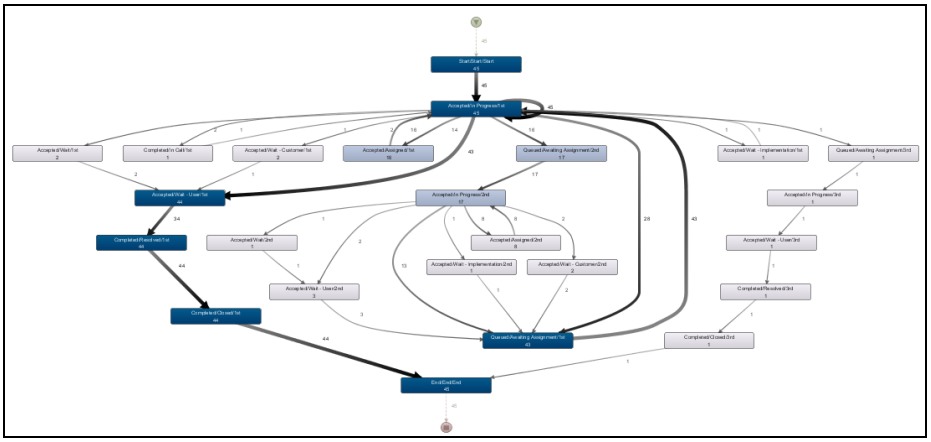

Figure 43. Process model, filtering the log by cases with use of the substatus wait-user.

11) *'country' = 'br'* (Brazil)

Available data of Brazil has 298 cases, with 143 variants and 3.136 events. 240 cases are completed in the first level, 58 in the second level. Third level receive just one case, which is pushed to the second level.

The substatus *wait-user* is used in 61 cases in the first level. In 21 of the cases, the flow goes back to the first activity, indicating a misusage of the substatus. In 33 occurrences the case is solved after the substatus, and in the other 7 cases, the flow goes to other status.

In the second level the substatus is used in 25 cases, 4 of these were consistent with the previous 33 (21 different cases). Half of the cases were solved after the substatus and just 4 (of the 21) we detected misuse of the status.

We conclude there are misuse in 25 cases, it represent an 8% of the cases.

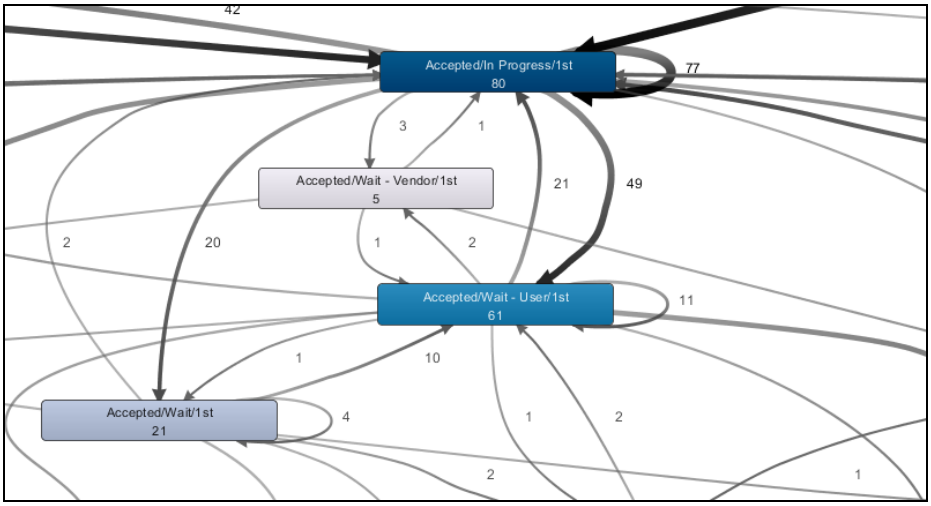

Figure 44. Detail of the process model centered in the status *Accepted/wait-user*.

12) *'country'* = *'fr'* (France)

France has 253 cases in the available data, with 155 variants and 3.126 events. The substatus wait-user is used in 98 cases, in the 3 levels of attention.

In the first level there are 33 cases using the substatus, and in 12 cases there are misusage of it.

In the second level is used in 47 cases, 38 previous to complete the case and 9 to others nodes. In 4 cases we detect misuse of the substatus.

In the third level is used in 29 cases and 16 of those are misusage of the substatus.

We conclude the substatus is misused in 37 cases that represent a 15% of the total cases and 38% of the total cases where the substatus where used.

13) *'country'* = *'us'* (United States)

US has 737 cases, with 211 variants and 6.247 events. The substatus was used in the 3 levels in 201 cases.

In the first level the substatus is used in 166 cases, with 22 evidence of misuse. In the second level there are 29 cases (excluding the cases with use in the first level) that use the substatus, and 7 of its was identified as misusage. In the third level, there are 6 cases (excluding cases where de substatus was used in first and second levels) and just one of its present misusage.

We conclude that there are 30 cases of misusage of the substatus, that represent a 4% of the total cases and an 18% of the cases where the substatus was used.

14) *'country'* = *'be'* (Belgium)

Belgium has 452 cases, with 258 variants and 6.600 events. The substatus was used in first and second level, in a total of 161 cases.

In the first level was used in 151 cases, 349 times. In 99 cases the following status where *complete/resolved*. In 32 cases, the following activity was

queued/awaiting_assignment, to go back to *accepted/in_progress* (previous activity). In 29 cases, the substatus was used as pause in the process also going back to *accepted/in_progress* status. We interpret both former behavior as misusage (61 cases).

In the second level, the substatus was used in 16 cases. In 5 cases, the flow goes directly back to the previous activity, making a pause in the process. In 9 instance, the case goes back to a *queued* status. We interpret that there were 14 cases of misusage of the substatus. 10 cases of the 16 detected were different to the cases detected for the first level, and 8 were misusage.

We conclude that there were 69 cases of misusage. It represent a 15% of the total cases, and 43% of the total cases where the substatus wait-user was used.

15) *'country' = 'in'* (India)

India has 402 cases, with 334 variants and 6.743 events. The substatus was used in every level, in a total of 273 cases.

In the first level was used in 149 cases, 286 times. In 42 cases the following status where *complete/resolved*. In 41 cases, the substatus was used as pause in the process also going back to *accepted/in_progress* status. In 16 cases, the following activity was *accepted/assigned/1st*, to go back to *accepted/in_progress* (previous activity) and finally back to *wait-user*. We interpret both former behavior as misusage (57 cases).

From the 273 cases, 207 are refer to the second level, where 175 use the substatus *wait-user*. From these, 98 go back directly to *accepted/in_progress* status. We interpret that as misusage of the substatus. 108 of the 207 cases where different to the first 149 detected in the first level, and 63 were misusage.

In the third level, the substatus was used in 16 cases and we detected misusage in 3 of those. These case were different to the previously detected.

We conclude that there were 123 cases of misusage. It represent a 31% of the total cases, and 45% of the total cases where the substatus wait-user was used.

16) *'country' = 'pl'* (Poland)

Poland has 1.725 cases, with 658 variants and 18.410 events. The substatus is used in every level, in a total of 618 cases.

In the first level, from the 618 total cases, 587 use the substatus wait-user, and 425 of its are followed directly by *accepted/in_progress* status, being a misusage of the substatus.

In the second level, the substatus is used in 48 cases, 22 are pushed back to first level, 15 go to *complete/resolved* status and 7 cases go back to *accepted/in_progress*. We interpret that 7 cases as a misusage. Only 3 of these 7 cases are different to the previously detected.

In the third level, there are 8 uses of the substatus, and no misusage.

We conclude that there were 428 cases of misusage. Clearly there is a misuse problem in the first line of attention, because it is used in a 25% of the total cases.

The total misusage represent a 25% of the total cases, and 69% of the total cases where the substatus wait-user was used.

17) *'country' = 'SE' and 'country' = 'se'* (Sweden)

Sweden has the most significant participation in the event log. In the available data, there are 2.483 cases, with 756 variants and 23.541.

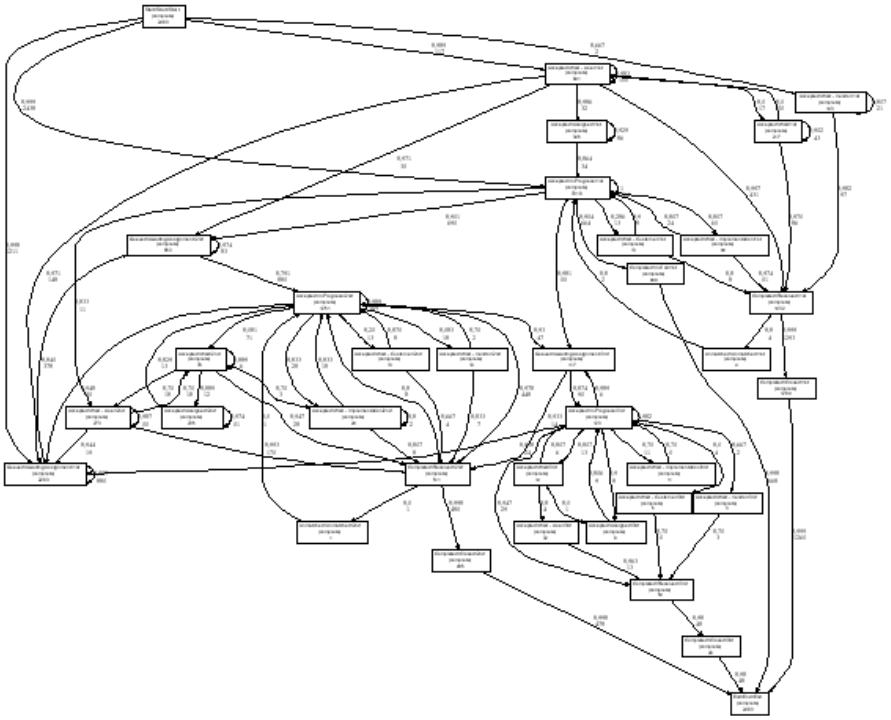

Figure 45. Process model of Sweden incidents, obtained with Heuristic Miner algorithm. The diagram shows the complexity of the model, is not our intention to show the activities.

The wait-user substatus is used in every level of attention, in a total of 722 cases.

In the first level, the substatus is used in 523 cases, 841 times. 329 times is followed by *completed/resolved* status and 197 times the flow go back to *accepted/in_progress* status, being this last behavior a misusage of the substatus.

In the second level, the substatus is used in 206 cases, and 55 of these were misusages of it. Of those 206, only 172 were different from the identified in the first level and 45 were misusage.

In the third level, the substatus was used in 32 cases (27 different to the previous identified), and 3 of these were misusages of it.

We conclude that there were 227 cases of misusage, it represent a 9% of the total cases, and 31% of the total cases where the substatus wait-user was used.