Homework 2

Miriam Wagner
373045

July 6, 2018

Contents

1 Introduction

In this assignment I have to analyze for a fictional financial institute data and make suggestions

2 General insights about the data set

The given data set is generated between 1st Oct 2011 0:38:44 (Saturday) and 14th Mar 2012 16:04:54 (Wednesday) and contains 13087 cases with 262200 events executed. One case contains minimal 3 events and maximal 175. In average 20.035 events.
There are 4366 different variants of cases. There are 36 different events possible "W_Complete_Application + COMPLETE" (9.141%), "W_Complete_Application + START" (8.967%), "W_Quotations + COMPLETE" (8.763%), "W_Quotations + START" (8.545%), "App_Fully_Submission + COMPLETE" (4.991%), "App_Incomplete_Submission + COMPLETE" (4.991%), "W_Handeling_Incomplete_Dossiers + COMPLETE" (4.35%), "W_Handeling_Incomplete_Dossiers + START" (4.348%), "W_Validation + COMPLETE" (3.011%), "W_Validation + START" (3.01%), "App_Rejection + COMPLETE" (2.912%), "W_Complete_Application + SCHEDULE" (2.811%), "App_Pre_Acceptation + COMPLETE" (2.81%), "P_Initiation + COMPLETE" (2.681%), "P_Selection + COMPLETE" (2.681%), "P_Sending + COMPLETE" (2.681%), "W_Quotations + SCHEDULE" (2.53%), "W_Handling_Leads

+ COMPLETE" (2.249%), "W_Handling_Leads + START" (2.249%), "App_Acceptation + COMPLETE" (1.95%), "W_Validation + SCHEDULE" (1.916%), "App_Finalization + COMPLETE" (1.913%), "W_Handling_Leads + SCHEDULE" (1.82%), "P_Cancellation + COMPLETE" (1.394%), "P_Returning + COMPLETE" (1.317%), "App_Cancellation + COMPLETE" (1.071%), "W_Handeling_Incomplete_Doss + SCHEDULE" (0.909%), "App_Initiation + COMPLETE" (0.857%), "App_Approving + COMPLETE" (0.857%), "App_Registration + COMPLETE" (0.857%), "P_Acceptation + COMPLETE" (0.855%), "P_Rejection + COMPLETE" (0.306%), "W_Fraud_Detection + COMPLETE" (0.103%), "W_Fraud_Detection + START" (0.103%), "W_Fraud_Detection + SCHEDULE" (0.047%) and "W_Changing_Contact_Details + SCHEDULE" (0.005%).
The complete, start and shedule are more details about the state of the event. Just about the Workflow events we have more information about the state.
All processes start in "App_Fully_Submission + COMPLETE" (100.0%).
The processes can end in 13 different events:
"App_Rejection + COMPLETE" (26.202%), "W_Validation + COMPLETE" (20.975%), "W_Handling_Leads + COMPLETE" (17.07%), "W_Complete_Application + COMPLETE" (14.816%), "W_Quotations + COMPLETE" (9.849%), "App_Cancellation + COMPLETE" (5.005%), "W_Handeling_Incomplete_Dossiers + COMPLETE" (3.454%), "P_Cancellation + COMPLETE" (2.132%), "W_Fraud_Detection + COMPLETE" (0.436%), "W_Changing_Contact_Details + SCHEDULE" (0.031%), "W_Validation + START" (0.015%), "App_Registration + COMPLETE" (0.008%) and "W_Quotations + START" (0.008%).

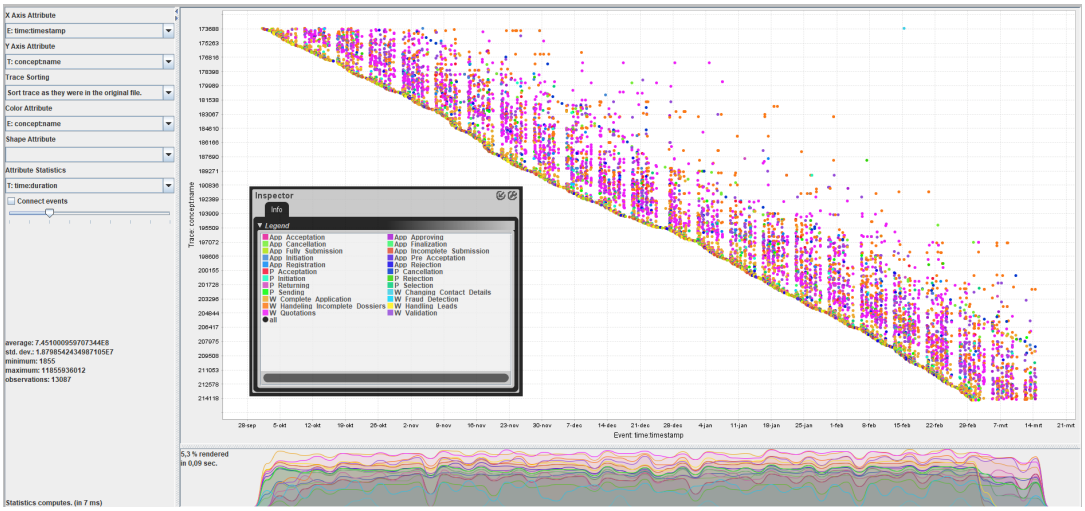


Figure 1: Dotted chart of the data set

Having a look at the whole data it can be seen, that there are just athe following events executes: "W_Complete_Application", "W_Handling_Leads", "App_Incomplete_Submission", "P_Cancellation", "App_Fully_Submission" and "App_Rejection". Furthermore is there a constant input of new cases. The events are mostly during the week executed same amount of times just "W_Fraud_Detection" and "W_Changing_Contact_Details" have really deeps and ups. What is not surprising, because both are outlier behavior actions.

3 Process Models

3.1 Exploring process

For exploring a good model of a data set I first filtered the data with the (**Filter log on event attribute names**) tool for extracting the required. Then I used **Interactive data heuristic miner** and **Inductive Miner** on the filtered data set to discover different models. After comparing the outcomes I decided to concentrate on **Interactive data heuristic miner** and their directly followed graphs and petri nets for understanding the lifecycle. I think the petri net is the best with basic configuration and just different frequency filters. The conformance checking where done with **Replay a log on Petri Net for conformance analysis** tool on the petri net and the filtered data. For the

precision check I applied the **Multi-perspective Process Explorer** tool on the petri net and the filtered data and chose "show precision mode" in the tool with basic configuration.

3.2 Application data set

Just the events beginning with "App_..." are required. The resulting data set is saved as "Filtered App".

3.2.1 General details of the data set

The data set is collected between 1st of 0ct 2011 (saturday), 00:38:44 and 14th of Mar 2012 (Wednesday), 15:33:57. It contains 13087 cases with 60849 executed events.
10 different events appear (occurences relative): "App_Fully_Submission" (21.507%), "App_Incomplete_Submission" (21.507%), "App_Rejection" (12.547%), "App_Pre_Acceptation" (12.107%), "App_Acceptation" (8.403%), ("App_Finalization" (8.242%), "App_Cancellation" (4.613%), "App_Initiation" (3.691%), "App_Approving" (3.691%) and "App_Registration" (3.691%). Just "APP_Fully_Submission" appears to be a start event. However there are 8 end events possible: "App_Rejection" (58.34%), "App_Cancellation" (21.449%), "App_Initiation" (8.573%), "App_Registration" (6.014%), "App_Approving" (2.575%), "App_Finalization" (2.499%), "App_Pre_Acceptation" (0.527%) and "App_Acceptation" (0.023%).
A trace contains maximal 8 different events and minimal 3. The mean is 4.65. In total there are 17 different variants of traces.

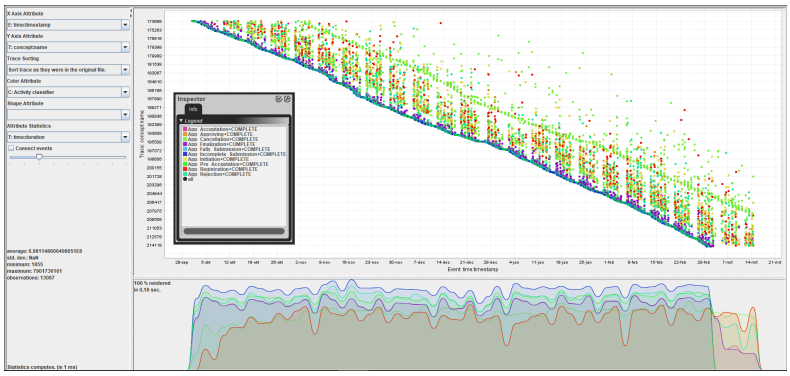
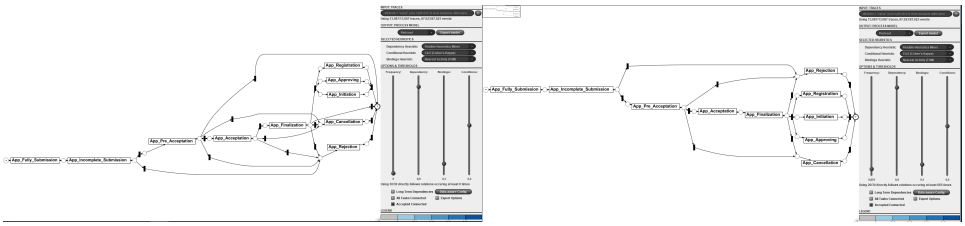


Figure 2: Dotted chart showing the time of events

In figure 2 the dotted chart can be seen. Having a closer look at this chart you see gaps, which are always on a sunday. Those gaps do not appear for "APP_Pre_Acceptation" and "APP_Incomplete_Submission". Furthermore is in the left below corner to see the average duration of a case, 8 days 1 hours 55 minutes and 14.86 seconds, and the maximum duration, 91 days 10 hours 55 minutes and 36.16 seconds. Both are given in milliseconds.

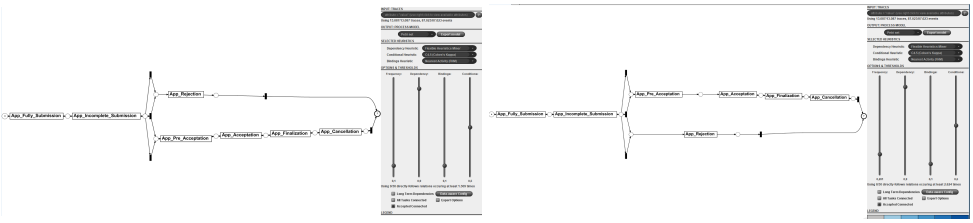
3.2.2 Discover and evaluate a model of the application lifecycle

To find an accurate model I tried different frequency filters. First I chose the frequency 0.1 and had a look at the resulting petri net. I also checked 0.2 and 0.51 for having a better understanding (not to see here is the petri net with 0.6 filtering, which was the next filtering after 0.1, that changed the net, but I decided, that it does not contain enough cases). For comparsion in the end I had also a look at the original petri net found by the **Interactive data heuristic miner**. In figure 3 the 4 considered petri nets can be seen. My first choice was the petri net with 0.1 as threhold for frequency, because it was a simple model that still tells us a lot about the main process (90% of the cases) and is not to specific (still has an acceptable generalization). Obviously the original graph does not fullfill the criterium of simplicity and also the graph with frequency 0.051 still looks not as simple. I do not consider the 0.2 filtering in the next steps, because it is the same petri net as the 0.1 filtered one.
The results of the conformance checking are shown in figure 4.
Having a look at the precision, figure 5, I could see, that the precision of 0.51 filtering is 100%, but of the 0.1 filtering just 88.1%. In combination with the results before, I came to the conclusion, that 0.1 filtering is not good enough as model and 0.051 would be good enough, but did not fullfill my simplicity



(a) Petri net without frequency filtering

(b) Petri net with 0.051 frequency filter



(c) Petri net with 0.1 frequency filter

(d) Petri net with 0.2 frequency filter

Figure 3: Considered petri nets

Property	Value
Calculation Time (ms)	1.5730877970505082
Raw Fitness Cost	0.038053029724153 73
Max Move-Log Cost	4.649575915030195
Num. States	8.634293573775542
Trace Fitness	0.9950123168393202
Move-Model Fitness	0.9944410483686092
Move-Log Fitness	0.9984908688011009
Max Fitness Cost	7.649575915030197
Trace Length	4.649575915030195

(a) Conformance for 0.051 frequency filter

Property	Value
Calculation Time (ms)	2.3281882784442574
Raw Fitness Cost	1.105677389776112
Max Move-Log Cost	4.649575915030182
Num. States	7.797508978375503
Trace Fitness	0.8819591437743015
Move-Model Fitness	0.9263009092993 05
Move-Log Fitness	0.9820325513868706
Max Fitness Cost	7.649575915030184 5
Trace Length	4.649575915030182

(b) Conformance for 0.1 frequency filter

Figure 4: Conformance checking

Avg activity precision	100%
# Moves Observed	144.789
# Moves Possible	144.789
Avg fitness	99,6%
% Violations	0,8%
% Event Violations	0,8%
% Data Violations	0%
# Correct Events	60.849
# Wrong Events	0
# Missing Events	498

(a) Precision for 0.051 as frequency filter

Avg activity precision	88,1%
# Moves Observed	97.074
# Moves Possible	110.161
Avg fitness	88,6%
% Violations	21,7%
% Event Violations	21,7%
% Data Violations	0%
# Correct Events	52.096
# Wrong Events	8.753
# Missing Events	5.717

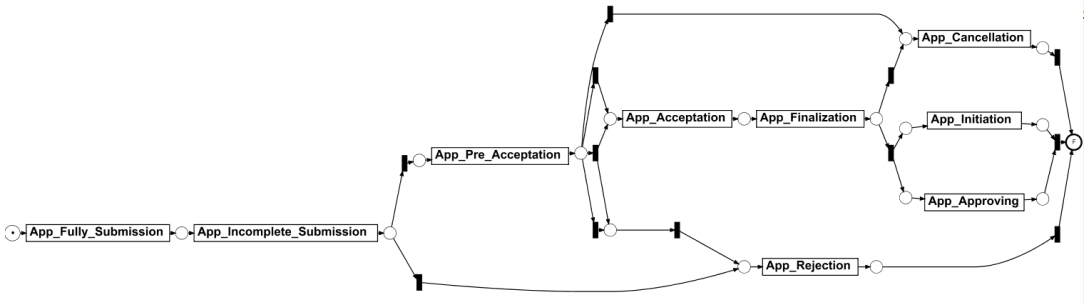
(b) Precision for 0.1 as frequency filter

Figure 5: Precision checking

Property	Value
Raw Fitness Cost	0.33223809887674755
Max Move-Log Cost	4.64957591503021
Num. States	8.79514021548101
Trace Fitness	0.9657921716262099
Move-Model Fitness	0.9944410483686092
Move-Log Fitness	0.9566108861210915
Max Fitness Cost	7.649575915030202
Trace Length	4.64957591503021
Queued States	22.11140826774671

Avg activity precision	94,8%
# Moves Observed	133.936
# Moves Possible	141.304
Avg fitness	96,6%
% Violations	7%
% Event Violations	7%
% Data Violations	0%
# Correct Events	57.772
# Wrong Events	3.077
# Missing Events	1.271

(a) Conformance and Precision



(b) Petri net

Figure 6: Frequency 0.069

criterium completly. Outgoing by 0.051 filtering I again tried different frequency filters starting by 0.075 to find a model fullfilling both, a similar simplicity as the 0.1 frequency model, but a better conformance and precision than this simpel model. The best result I found was with 0.069 filtering. This model is simpel enough to understand the main lifecycle, but still has good result for conformance (fitness 96.58%) and precision(94.8%),6. Based on this I decided to choose this petri net for the lifecycle of the application data.

3.2.3 Discussion of the lifecycle

Simplifying the discussion I will not write "App_.." in begin of all events, but every event mentioned in this section is from the application data set. To also have details about the occurency of the steps in

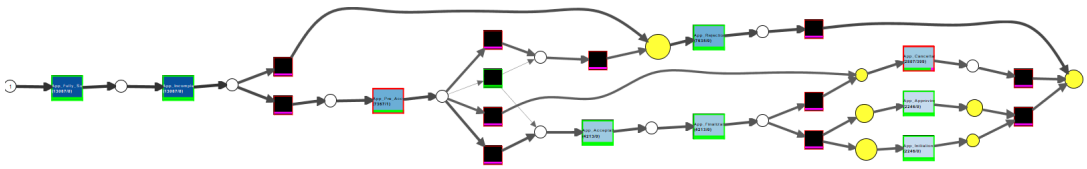


Figure 7: Replay on petri net with 0.069 frequency filter

the model I used the replay result, 7.
The process always starts with "Fully_Submission" and "Incomplete_Submission" (13087 cases). After this there are two different outcomes:

1. "Rejection" (5719 cases, executed in total in 7365 cases)
2. "Pre_Acceptation" (7367 cases)

"Pre_Acceptation" has 4 different successor events

1. "Rejection" (1916 cases)
2. "Rejection" and "Acceptation" (0 cases)
3. "Acceptation" (4213 cases)

4. "Cancellation" (1239 cases)

"Acceptation" is followed by "Finalization" (4213). "Finalization" is followed in (1967 cases) by "Cancellation", which is in total executed 3206 times, and otherwise by "Approving" and Initation (2246 cases). All the differences between incoming and outcoming cases can be explained through the filtering by frequency for building the model.
Looking at the **Time between Transition matrix** of the replay result it can be seen, that the transition to "Approving" and "Initiation" need in average ~16 days, what is the second longest transition time. "Cancellation" has the worst transition time with in average ~20 days.

3.3 Proposal data

Just the events beginning with "P_..." are required. The resulting data set is saved as "Filtered P".

3.3.1 General Details of the data set

The data set is collected between 1st of 0ct 2011 (saturday), 10:44:40 and 14th of Mar 2012 (Wednesday), 15:50:59 and contains 13087 cases with 31244 events. Having a look at the visualization you can see that there a gaps in the workflow.

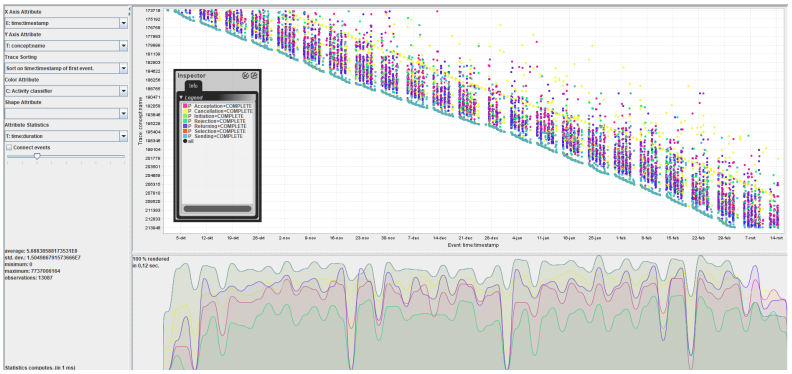
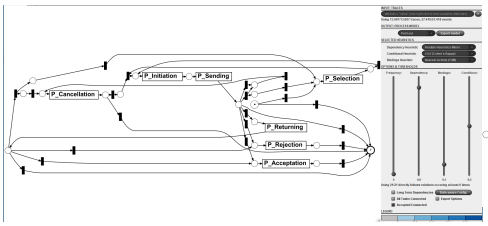


Figure 8: Dotted chart showing the time of events

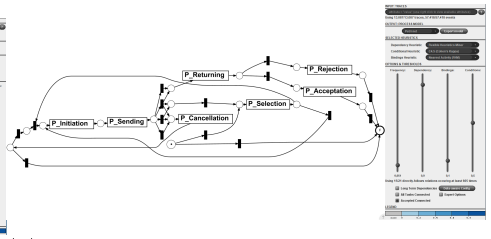
In figure 8 the dotted chart can be seen. Having a closer look at this gaps makes clear, that it is always a sunday. What does not show this behavior so clear is "P_Cancellation" and "P_Initiation". Furthermore is in the left below corner to see what is the average duration of a case, 6 days 14 hours 0 minutes and 38.59 seconds, and the maximum duration, 89 days 13 hours 10 minutes and 6.16 seconds. Both is given in milliseconds.
The data set has 7 events: "P_Initiation" (22.5%), "P_Sending" (22.5%), "P_Selection" (22.5%), "P_Cancellation" (11.698%), "P_Returning" (11.055%), "P_Acceptation"(7.179%) and "P_Rejection" (2.567%), where the percentage is the relative occurence. In total this data set contains 13087 cases with in total 31244 events. It always stars with "P_Selection" and has 5 different end events: "P_Acceptation" (44.726%), "P_Cancellation"(32,702%), "P_Rejection" (15.992%), "P_Sending" (4.806%) and "P_Returning" (1.775%).
There are 169 different variants of traces.
Maximal 30 events are executed in a case and minimal 0, what is a hint, that there are cases ending directly. The mean of events per class is 2.387.

3.3.2 Discover and evaluate models of the proposal lifecycle

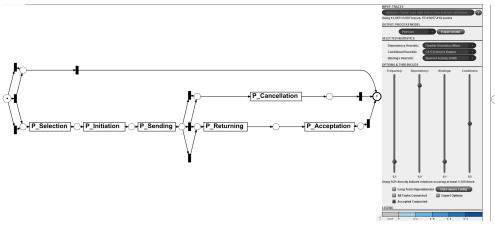
Applying the steps on the proposal data set it gave me 4 models I wanted to have a closer look at.
Because of simplicity reasons I picked out of those, figure 9, just the 0.051 and 0.1 filtered petri nets and because of genralization not 0.126 or 0.2.
The conformance and precision outcomes, to see in figure 10, told me, that the 0.1 frequency model (fitness 91.55%, precision 80%) has not a good enough precision. And also the precision of the 0.051 filtered model is not really high. So I searched for an other model, which has a similar simplicity as the



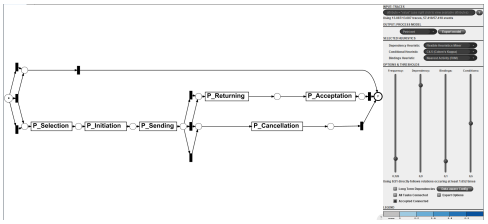
(a) Petri net without frequency filtering



(b) Directly followed graph 0.051 frequency filter



(c) Petri net for 0.1 frequency filter



(d) Petri net for 0.126 frequency filter

Figure 9: Considered petri nets

Property	Value
Raw Fitness Cost	0.07778711698632224
Max Move-Log Cost	2.3874073508061433
Num. States	6.670130664017713
Trace Fitness	0.9882694969197188
Move-Model Fitness	0.9937748638349697
Move-Log Fitness	0.9959403350088936
Max Fitness Cost	2.3874073508061433
Trace Length	2.3874073508061433
Queued States	16.723542446702886

(a) Conformance for 0.051 as frequency filter

Property	Value
Raw Fitness Cost	0.7597615954764291
Max Move-Log Cost	2.387407350806137
Num. States	6.684114006265772
Trace Fitness	0.9154572453183918
Move-Model Fitness	0.9867960571559562
Move-Log Fitness	0.9298519372518611
Max Fitness Cost	2.387407350806137
Trace Length	2.387407350806137
Queued States	12.33177962863913

(b) Conformance for 0.1 as frequency filter

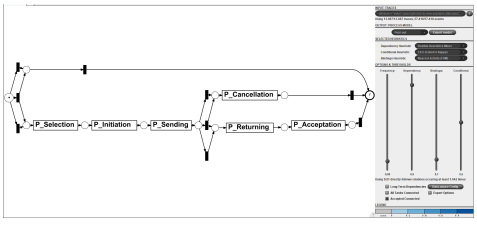
Avg activity precision	83,6%
# Moves Observed	103.538
# Moves Possible	123.899
Avg fitness	37,4%
% Violations	3,2%
% Event Violations	3,2%
% Data Violations	0%
# Correct Events	30.740
# Wrong Events	504

(c) Precision for 0.051 as frequency filter

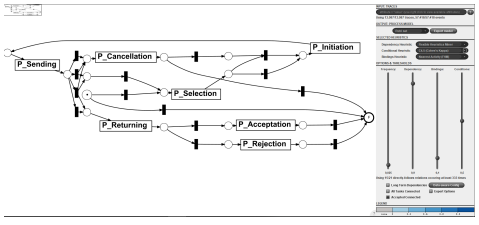
Avg activity precision	80%
# Moves Observed	72.228
# Moves Possible	90.330
Avg fitness	30,3%
% Violations	31%
% Event Violations	31%
% Data Violations	0%
# Correct Events	22.119

(d) Precision for 0.1 as frequency filter

Figure 10: Conformance and precision checking



(a) 0.08 frequency filtering



(b) 0.025 frquency filtering

Figure 11: Petri net

Property	Value
Raw Fitness Cost	0.7597615954764291
Max Move-Log Cost	2.387407350806137
Num. States	6.684114006265772
Trace Fitness	0.9154572453183918
Move-Model Fitness	0.9867960571559562
Move-Log Fitness	0.9298519372518611
Max Fitness Cost	2.387407350806137
Trace Length	2.387407350806137
Queued States	12.33177962863913

Avg activity precision	80%
# Moves Observed	72.228
# Moves Possible	90.330
Avg fitness	30,3%
% Violations	31%
% Event Violations	31%
% Data Violations	0%
# Correct Events	22.119
# Wrong Events	9.125
# Missing Events	818

(a) 0.08 frequency

Property	Value
Raw Fitness Cost	0.07778711698632224
Max Move-Log Cost	2.3874073508061433
Num. States	6.670130664017713
Trace Fitness	0.9882694969197188
Move-Model Fitness	0.9937748638349697
Move-Log Fitness	0.9959403350088936
Max Fitness Cost	2.3874073508061433
Trace Length	2.3874073508061433
Queued States	16.723542446702886

Avg activity precision	83,6%
# Moves Observed	103.538
# Moves Possible	123.899
Avg fitness	37,4%
% Violations	3,2%
% Event Violations	3,2%
% Data Violations	0%
# Correct Events	30.740
# Wrong Events	504
# Missing Events	514

(b) 0.025 frequency

Figure 12: Cnformance and precision

0.1 and 0.051 filtered models and has maybe a better performance. The models fullfilling the simplicity criterium had 0.08 or 0.025 frequency filtered.
After having a look at all results in 12 and comparing it with earlier results I chose the model with 0.051 frequency filter. The precision of 0.025 is with 83.6% the same and the fitness is also the same, 98.82%. With 0.08 filtering it had a worse fitness (91.54%) and precision (80%). All filtering inbetween were the same.

3.3.3 Discussion of the lifecycle

Simplifying the discussion I will not write "P_.." in begin of all events, but every event mentioned in this section is from the proposal data set.

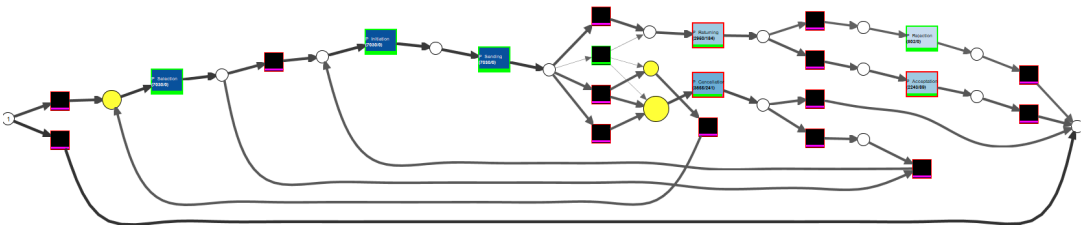


Figure 13: Replay on petri net with 0.051 frequency filter

This model has the conspicuousness, that 8072 cases directly end, but recheck it with the original graph this can be found, too. The other 5015 cases start with "Selection", which is executed in total 7030 times, because of a loop. Then "Initiation" event is performed (7030 cases), while in 2015 cases it also has to be executed "Cancellation" before "Initiation" can be executed. "Initiation" is followed by "Sending" (7030 cases), which has 4 different succesor actions.

1. "Returning" (3143 cases, while in 184 cases a part of the trace is not to see in the model)
2. "Returning", "Selection" and "Cancellation" (0 times)
3. "Selection" and "Cancellation" (2015 times)
4. "Cancellation" (1881 times)

"Returning" is in 802 cases followed by "Rejection" and in 2243 cases by "Acceptation". "Cancellation" is in 1881 cases the last action or it leads to "Initiation" (2015 times).
So there are two backloops to see in this model, which are executed in at most 25% of the traces.
Furthermore are there

3.4 Combinded Model

For combinded Models I first filtered the data to have a dataset with all proposal data combined with the application data. This data I filtered with Heuristic filter (all configurations to 100% and just deciding what the endstate is) with the outcomes/endstates "APP_rejected" or "APP_cancelled" or "APP_approved". I saved them under the names "Filtered P App with approv", "Filtered P App with canc" and "Filtered P App with rej".

3.4.1 Endstate APP_Approved

General Details of the data set
The data set is collected between 1st of 0ct 2011 (saturday), 12:36:08 and 14th of Mar 2012 (Wednesday), 15:23:32. The set contains 301 and 4550 events executed.

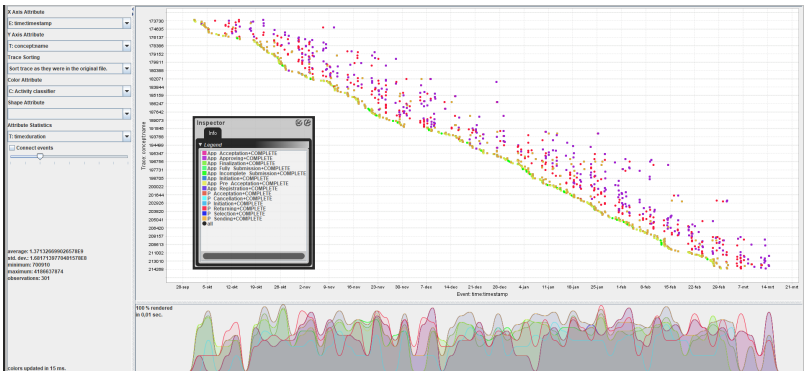


Figure 14: Dotted chart showing the time of events

In figure 14 the dotted chart can be seen. There is a gap to see on every sunday. This is more clearly to see, when changing the color to days of the week. There you can see, that it is always a sunday. Furthermore is in the left below corner to see what the average duration of a case is, 15 days 20 hours 55 minutes and 26.70 seconds, and the maximum duration, 48 days 27 hours 23 minutes and 14.69 seconds. Both is given in milliseconds in the figure.
There are 14 different events: "P_Initiation" (10.0%), "P_Sending" (10.0%), "P_Selection" (10.0%), "P_Returning" (7.099%), "App_Pre_Acceptation" (6.615%), "App_Initiation" (6.615%), "App_Approving" (6.615%), "App_Registration" (6.615%), "App_Finalization" (6.615%), "App_Fully_Submission" (6.615%), "App_Incomplete_Submission" (6.615%), "App_Acceptation" (6.615%), "P_Acceptation" (6.593%), "P_Cancellation" (3.385%). The percentages are the relative occurences. There are 73 different variants of traces.
Maximal 34 events are executed in a case and minimal 12. The mean of events per class is 15.116.
Discover and evaluate models of lifecycle with Endstate APP_Approved
Like for approved and proposal I first checked different frequency filter setting to have a first idea, which models fullfill simplicity. Then for every chosen frequency in begin I checked conformance and precision.

	Frequency			
	0	0.1	0.2	0.3
Simplicity	-	-	+	++
Fitness	99.84	99.47	93.93	93.93
Precision	93.3	94.5	91.7	94.8

Figure 15: Results for approved as endstate

Based on the results, 15, I chose the model with 0.3 filtering. This one has a high simplicity, but still has surprisingly good results.
Discussion of the model
In figure 16 the petri net and the result of the replat can be seen. It is hard to clearly read the details, but I describe it in detail here.

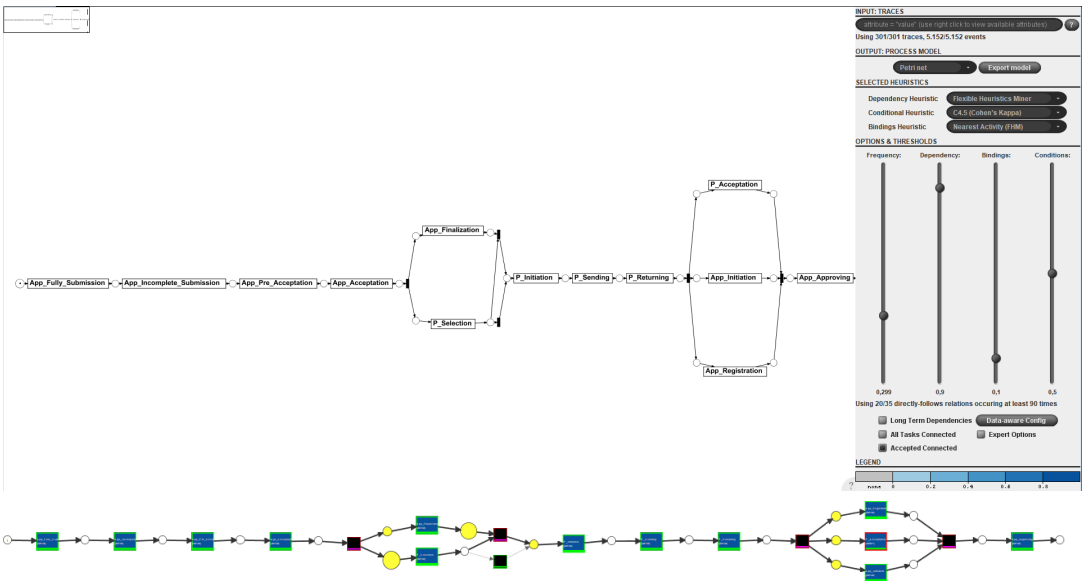


Figure 16: Endstate APP_Approved

All 301 cases start with the same 4 events: "App_Fully_Submission" →"App_Incomplete_Submission" →"App_Pre_Acceptation" →"App_Acceptation". Then there are two events parallel: "App_Finalization" and "P_Selection". Before and after those actions are events skipped in the model, that occur in the original log. Afterwards "P_Initiation" →"P_Sending" →"P_Returning" are executed in all cases. Again 3 events are parallel "App_Registration", "App_Registration" (just in 300 cases and "App_Initiation". The last event is "App_Approving".
When I controlled the time matrix for the transitions I saw, that there are a lot of transitions taking between 12 and 16 days.

3.4.2 Endstate APP_Cancelled

General Details of the data set
The data set is collected between 1st of 0ct 2011 (saturday), 09:45:25 and 14th of Mar 2012 (Wednesday), 15:30:47. The set contains 1937 and 13524 events executed.

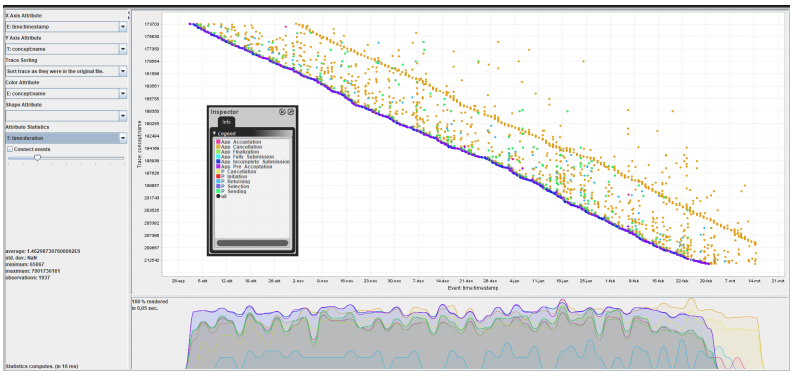


Figure 17: Dotted chart showing the time of events

In figure 17 the dotted chart can be seen. On every sunday there are just two events executed "App_Pre_Acceptation" and "App_Incomplete_Submission". Furthermore is in the left below corner to see what the average duration of a case is, 16 days 22 hours 23 minutes and 7.31 seconds, and the maximum duration, 90 days 3 hours 48 minutes and 38.29 seconds. Both is given in milliseconds in the figure.
There are 11 different events (relative occurences): "App_Cancellation" (14.323%), "App_Fully_Submission" (14.323%), "App_Incomplete_Submission" (14.323%), "App_Pre_Acceptation" (14.315%), "P_Cancellation"

(7.572%), "P_Initiation" (7.572%), "P_Sending" (7.572%), "P_Selection" (7.572%), "App_Acceptation" (6.182%), "App_Finalization" (5.694%) and "P_Returning" (0.555%). There are 46 different variants of traces.
Maximal 32 events are executed in a case and minimal 3. The mean of events per class is 6.982.

Discover and evaluate models of lifecycle with Endstate APP_Cancelled

	Frequency		
	0	0.049	0.1
Simplicity	-	+	++
Fitness	99.96	99.39	96.84
Precision	92.5	91.8	94.6

Figure 18: Results for cancelled as endstate

Like for APP_Approved I checked different configurations and based on the results, 18, I decided to pick 0.1 filtered frequency model. The fitness and precision is still higher than 90%, but it is also the most simpel model.

Discussion of the model

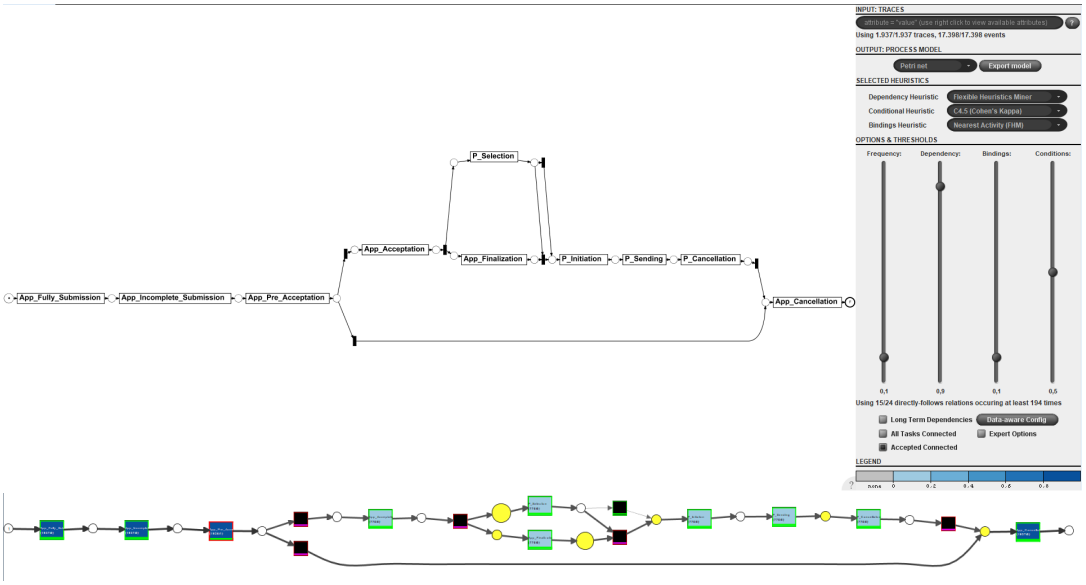


Figure 19: Endstate APP_Cancelled

In figure 19 the petri net and the result of the replat can be seen.
All cases start with "App_Fully_Submission" →"App_Incomplete_Submission" →"App_Pre_Acceptation" (this is just executed in 1936 cases). Then in 1167 cases "App_Cancellation" follows directly. Otherwise "App_Acceptation" (770 cases), followed by "P_Selection" and "App_Finalization", but those both can appear in both orders. After those are executed, "P_Initiation" →"P_Sending" →"P_Cancellation" →"App_Cancellation" is executed.
The time transition matrix showed me, that the transition to "App_Finalization" and "P_Cancellation" took the most time (between 17 and 22.5 days).

3.4.3 Endstate APP_Rejected

General Details of the data set

The data set is collected between 1st of 0ct 2011 (saturday), 08:11:08 and 14th of Mar 2012 (Wednesday), 15:20:23. The set contains 7252 and 26691 events executed.
In figure 20 the dotted chart can be seen. On every sunday there are just two events executed "App_Rejection" and "App_Incomplete_Submission". Furthermore is in the left below corner to see what the average duration of a case is, 1 days 6 hours 50 minutes and 35.01 seconds, and the maximum duration, 76 days 22 hours 42 minutes and 38.05 seconds. Both is given in milliseconds in the figure.

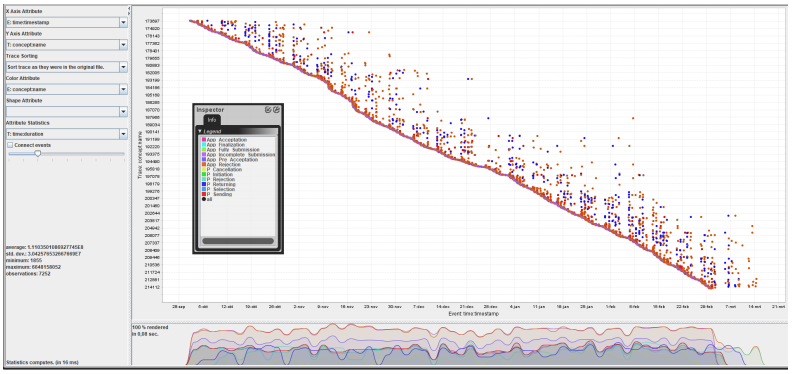


Figure 20: Dotted chart showing the time of events

There are 12 different events (relative occurences): "App_Rejection" (27.17%), "App_Fully_Submission" (27.17%), "App_Incomplete_Submission" (27.17%), "App_Pre_Acceptation" (5.744%), "P_Initiation" (1.997%), "P_Sending" (1.997%), "P_Selection" (1.997%), "App_Acceptation" (1.678%), "App_Finalization" (1.57%), "P_Rejection" (1.57%), "P_Returning" (1.51%), "P_Cancellation" (0.427%). There are 30 different variants of traces.
Maximal 12 events are executed in a case and minimal 3. The mean of events per class is 3.681.
Discover and evaluate models of lifecycle with Endstate APP_Rejected
The last analysis is of the models ending in APP_Rejected. In the first step I checked different frequency filters and decided based on simplicity and traceability I had a closer look at 0, 0.025 and 0.1.

	Frequency		
	0	0.025	0.1
Simplicity	-	+	+++
Fitness	1.00	99.60	96.91
Precision	99.3	98.7	100

Figure 21: Results for cancelled as endstate

Based on 21 I chose 0.1 filtered frequency model as the best. It is really easy to follow and has a good fitness.
Discussion of the model
In figure 22 the model and the replay can be seen.
The process starts with "App_Fully_Submission" →"App_Incomplete_Submission" (7252 cases). This is or followed directly by "App_Rejection" (5719 cases) or first by "App_Pre_Acceptation" (1533 cases) and then "App_Rejection". There are for "App_Rejection" details not presented in the model because of the filtering.
The worst transition is with 5.51 days from "App_Pre_Acceptation" to "App_Rejection".

3.5 C-net of the proposal process

Based on the results of my analysis of the model before I chose the same frequency filtering for the C-net of the proposal process.
In 23 I show two C-nets of the proposal process for comparsion. 23a is the one I picked and I will discuss in detail. In 23b the original C-net of the process is to see and obviously it is much more complicated and not so intuitive than the filtered one.

3.5.1 Analysis of the C-net

For simplicity reasons I will not write "P_" as prefix of every activity.
The first thing I did is having a look at the maximal number of bindings. This is for sending with 4 possible bindings. The input is always from initiation, but there are 4 different outputs possible. Just checking the possible traces shows you, that there are infinite many possible traces, because of a loop between sending and initiation. All trace possibilities:

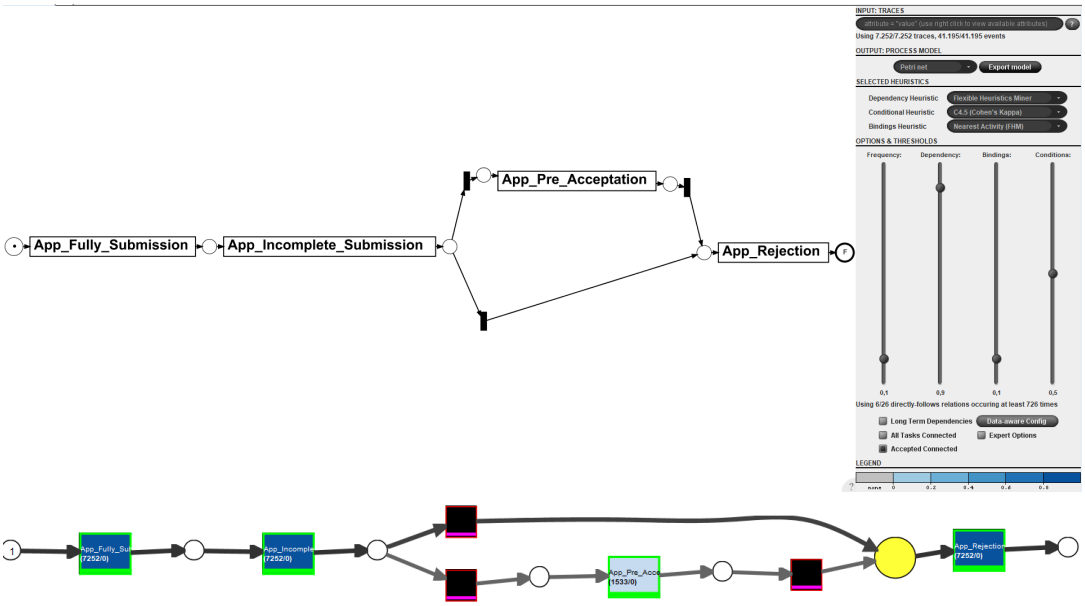
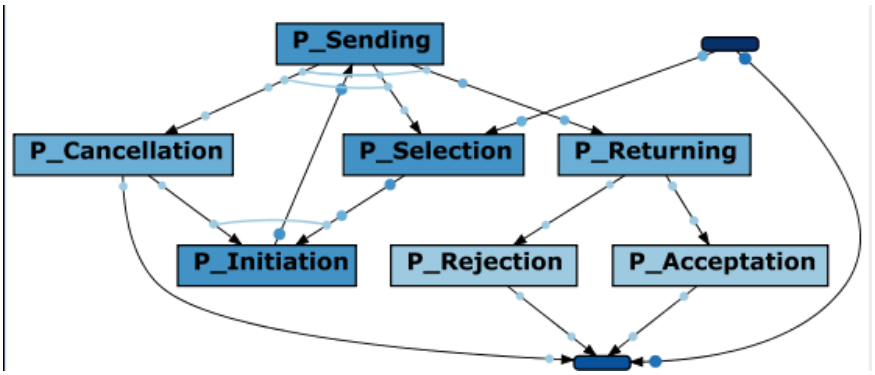
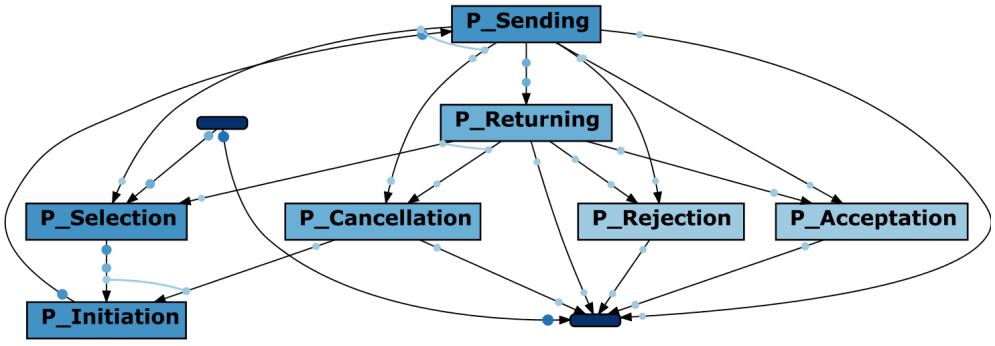


Figure 22: Endstate APP_Rejected



(a) C-net with 0.051 filtered frequency



(b) C-net without filtered frequency

Figure 23: C-Nets of the proposal process

- Done
- Selection →Initiation →Sending →
  - Returning →
    - Acceptation →Done
    - Rejection →Done
  - ((Cancellation →Selection) or (Selection →Cancellation)) →Initiation →Sending ...
  - Cancellation →Done

This pathes are also what you would expect. The sending step has the biggest variance when looking at the succesor events and we have a fixed prefix for every case, that not directly ends.

3.6 Own Petri net of the proposal process

Based on my analysis of the C-net and also the petri net.The first problem I saw, was that the start event can occur in between again. Because of that you have a deadlock if you would build it without transitions, that can occur more than once. Also the cancellation and selection or just cancellation or split makes problems.
Then just checking the different variants and the occurences I found out, that already just the traces occur at least 1% of the time have "P_Selection" as start state and as in between state. So I decided just to emphasize the main three main variants. I think this is the behavior as it should be probably.
The traces that I considered were:
"P_Selection"→"P_Intiation" →"P_Sending"→

- "P_Returning" →"P_Acceptation" (1.537 cases, 11.74%)
- "P_Returning" →"P_Rejection" (1.132 cases, 8.65%)
- "P_Cancellation" (574 cases, 4.39%)

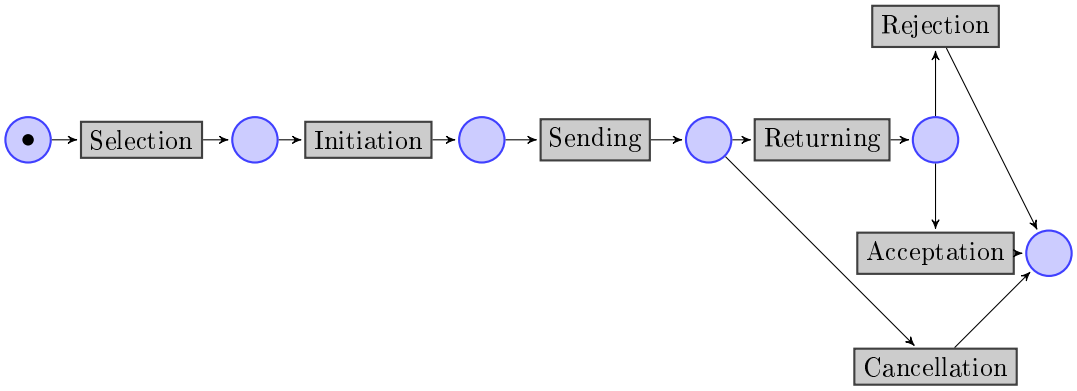


Figure 24: Own petri-net

3.7 Analysis of the performance of Application and work process

Just the events beginning with "P_..." or "W_..." are required. The resulting data set is saved as "Filtered Work and App".

3.7.1 General Details of the data set

The data set is collected between 1st of 0ct 2011 (saturday), 00:38:44 and 14th of Mar 2012 (Wednesday), 16:04:54 and contains 13087 cases with 230956 events.

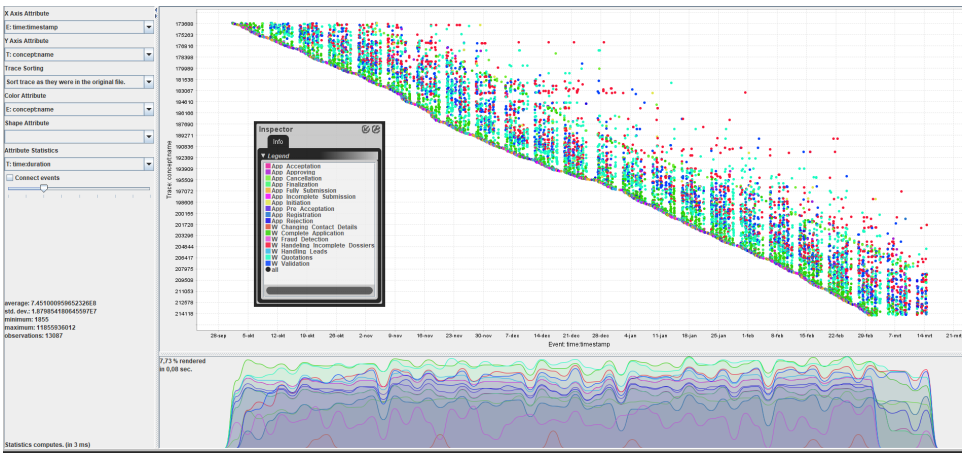


Figure 25: Dotted chart showing the time of events

In figure 25 the dotted chart can be seen. Having a closer look at this I saw again gaps on sunday. On sundays just "App_Incomplete_Submission" and "App_Rejection" is been executed. Furthermore is in the left below corner to see what is the average duration of a case, 8 days 14 hours 58 minutes and 20.10 seconds, and the maximum duration, 137 days 5 hours 18 minutes and 56.01 seconds. Both is given in milliseconds.
The data set has 29 events (occurence relative): "W_Complete_Application" (10.377%), "W_Complete_Application" (10.18%), "W_Quotations" (9.948%), "W_Quotations" (9.701%), "App_Fully_Submission" (5.666%), "App_Incomplete_Submission" (5.666%), "W_Handeling_Incomplete_Dossiers" (4.939%), "W_Handeling_Incomplete_D (4.936%), "W_Validation" (3.418%), "W_Validation" (3.417%), "App_Rejection" (3.306%), "W_Complete_Application" (3.192%), "App_Pre_Acceptation" (3.19%), "W_Quotations" (2.872%), "W_Handling_Leads" (2.554%), "W_Handling_Leads" (2.553%), "App_Acceptation" (2.214%), "W_Validation" (2.175%), "App_Finalization" (2.171%), "W_Handling_Leads" (2.066%), "App_Cancellation" (1.215%), "W_Handeling_Incomplete_Dossiers" (1.032%), "App_Initiation" (0.972%), "App_Approving" (0.972%), "App_Registration" (0.972%), "W_Fraud_Detection" (0.117%), "W_Fraud_Detection" (0.117%), "W_Fraud_Detection" (0.054%) and "W_Changing_Contact_Details" (0.005%).
There are 3668 different variants of traces.
Maximal 162 events are executed in a case and minimal 3. The mean of events per class is 17.648.
All cases start with "App_Fully_Submission", but there are 12 different outcomes: "App_Rejection" (26.202%), "W_Validation" (20.975%), "W_Handling_Leads" (17.07%), "W_Complete_Application" (14.816%), "W_Quotations" (9.895%), "App_Cancellation" (7.091%), "W_Handeling_Incomplete_Dossiers" (3.454%), "W_Fraud_Detection" (0.436%), "W_Changing_Contact_Details" (0.031%), "W_Validation" (0.015%), "App_Registration" (0.008%) and "W_Quotations" (0.008%).

Questions

Conclusion