

Analysis of the reimbursement process at TU/e using Process Mining BPI Challenge 2020

Trifonov Dmitry, Kalinin Yuri, Khabarova Ekaterina, Evtushenko Aleksandr,
Zhemchuzhnikov Evgeny
Sberbank of Russia

{Trifonov.Dm.Vl, Lupin.Yu.A, Khabarova.Ek.Ma,
AAEvtushenko, EVZhemchuzhnikov}@sberbank.ru

Abstract: Process mining is a modern technique for visualizing and analyzing the entire range of processes in an organization, thanks to which we can get detailed information about bottlenecks, possible risks for the business, and the performance of individual systems and employees. Understanding the ongoing processes allows us to reduce costs, improve the quality of services provided, and speed up production.

As part of the 2020 Business Process Intelligence Challenge, using a similar method, we reviewed the experience of TU/e of organizing business trips for its employees. To answer the process owner's questions, we used a combination of manual data analysis (SQL) and own developed mining process software "myPM" (section 5). This report contains directly the answers to the questions posed, recommendations for improvement, as well as an overview of some deviations in the process that were identified during the analysis

Keywords: BPI Challenge, Process Mining, Process Analysis

1 Introduction

In the 21st century, a digital boom has occurred: the count of units for automated data processing and logging tools is growing every day, and at the same time, the amount of information they produce is increasing. It is already difficult to imagine any process in an organization that would not leave a digital footprint. All this has led to the fact that such a phenomenon as the process mining is receiving more and more interest, and a clear example of this interest is the annual Business Process Intelligence Challenge.

In BPIC2020, from the point of view of the process mining, it is proposed to consider the experience of Eindhoven University of Technology (TU / e) in organizing domestic and international travel of its employees, who travels a lot to conferences or to other universities for project meetings and / or to meet up with colleagues in the field.

The main data layer is represented by a slice of several logs for 2017-2018 with a partial capture of 2019. It should be noted, however, that for 2017, there are some differences as this was a pilot year and the process changed slightly on several occasions. Thanks to the available logs, the process owner expects to receive answers to several basic and additional questions from the participants in the BPI challenge, as well as to hear recommendations on possible ways to improve the current travel model.

2 Data and Processes Overview

For analysis, 5 logs are presented for three processes:

- International travel;
- Domestic travel;
- Requests for payment of non-travel expenses.

Logs Permit, international declaration, PrepaidTravelCost to the international travel process.

For domestic trips - log Internal declarations, for requests for payment - Request for payment

International travel process in 5 stages:



Fig. 1. Main stages for international travels

For internal business trips, only the “Declaration” stage works, and for requests for payment only “Request For Payment”, followed by Payment.

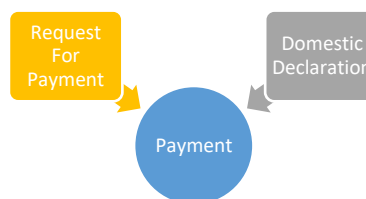


Fig. 2. Two-stage processes Request For Payment and DomesticDeclarations

All stages (except Trip and Payment) follow the following route:

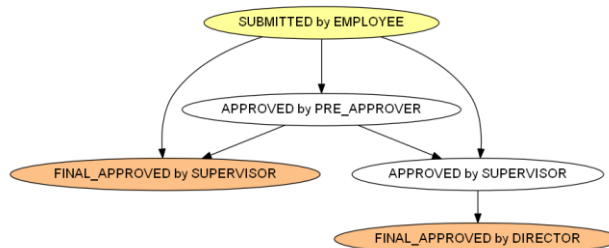


Fig. 3. Main trace for 2017

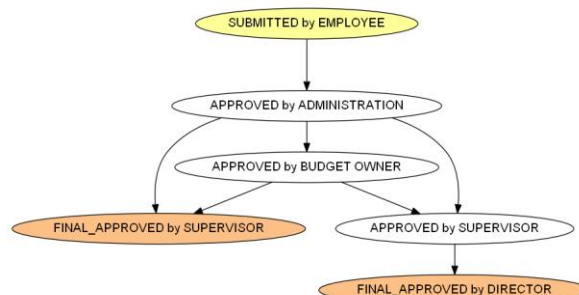


Fig. 4. Main trace for 2018

Payment stage contains 2 steps:



Fig. 5. Payment stage

After declaration approval and request for payment approval, the “Request Payment” stage begins - payment approval. This stage is executed by the system.

During the first 3 days, 64.56% (15,839) of approved requests and declarations are approved. 22.94% (5627) during the week. Up to 30 days - 11.68% (2865) and more than a month - 0.81% (198), the wait is up to 284 days. There are also three declarations approved prior to final approval. this indicates that this stage is being performed by employees, as well as the presence of errors in documents, absence of documents, etc.

Payment Handled stage is held twice a week on Tuesdays and Thursdays, once a day, usually at 16-30 or 15-30. On 31.07.2018 and 03.01.2019, there were unscheduled mass payments. All declarations, requests for payment, and prepayment are paid in a lump sum.

For 2017-2018, 8 days of payments were missed, which led to an increase in the days of waiting for payments after approval (Request Payment status) to 8 days.

Dates on which planned payments were not made:

2017:	2018:
17.04.2017 Monday	02.04.2018 Monday
27.04.2017 Thursday	10.05.2018 Thursday
25.05.2017 Thursday	21.05.2018 Monday
05.06.2017 Monday	30.07.2018 Monday

Waiting for payment after approval (Request Payment status) for more than 8 days was revealed in 490 (2%) applications. The wait was up to 63 days. Suggestion: it is necessary to fully automate this process - collecting the details of employees, it is possible to create a database with the details of recipients and the formation of lists of enrollment of approved applications for payment.

Process change:

In 2017, some requests were preliminary approved (role Pre_approver). In 2018, there was mandatory coordination of all requests with the administrator (the Administration role).

In 2018, an approving with the Budget owner appears in the process, 36% of requests have been approved by him.

The Missing role does not occur after 03/29/2018, it can be assumed that this event was excluded from the target route in 2018.

When analyzing the PrepaidTravelCost log, an interesting fact was revealed. Before moving on to it, let's reveal the essence of this process a little, illustrating the TOP5 popular routes.

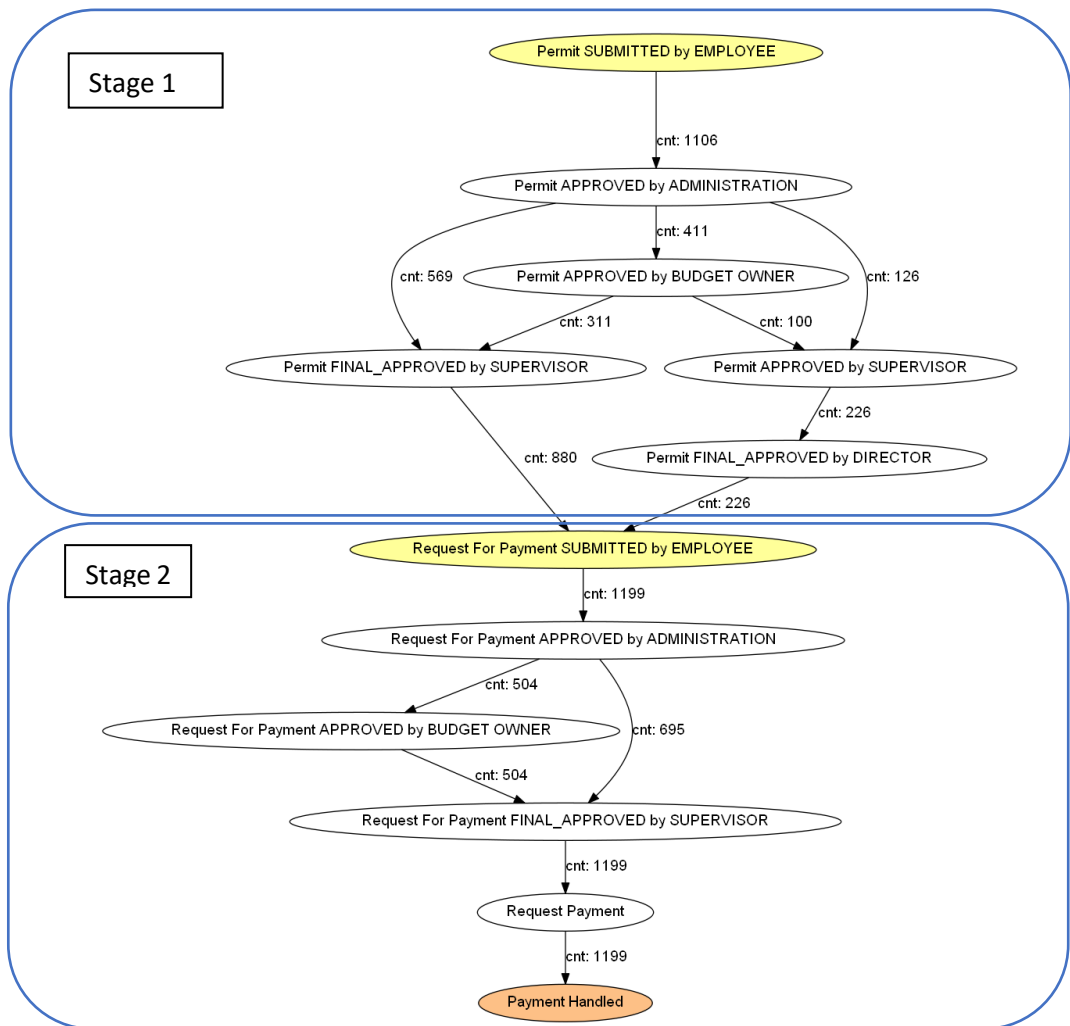


Fig. 6. Top 5 traces PrepaidTravelCost.xes

As you can see, this process is divided into 2 parts. The trip must be permitted by FINAL_APPROVED (stage 1) prior to claiming prepaid travel expenses. After receiving permission, the employee registers a request for payment and, if approved, receives a payment - Payment Handled (stage 2).

It is noteworthy that all "Request For Payment" pass a similar path as the parent Permit, i.e. if the Permit was agreed through BUDGET OWNER, then the Request For Payment also goes through BUDGET OWNER; if Permit has passed the stage of approval with BUDGET OWNER, then Request For Payment does not need approval from BUDGET OWNER (all this can be seen in the graph above).

Several prepaid travel reimbursements may be made within the same trip. This can also be seen in the process graph above (1106 cases Permit SUBMITTED => 1199 cases Request For Payment) and this is also mentioned in the brief description of the process, but this is where the fun begins.

The fact is that a certain amount is allocated for each trip (the Permit RequestedBudget attribute) and it is logical that prepaid expenses should not exceed this amount (Permit RequestedBudget > sum (RequestedAmount)). However, for 71 business trips, there is a fact of going over budget. Here are some interesting cases:

Table 1. Examples of cases with payments over the limit

Permit Id	Requested Budget	Handled Sum	Pay count	Total Declared	Overspent Amount
travel permit 14775	0,000	924,422	1	1040,873	1040,873
travel permit 24054	7613,328	8793,394	11	9246,303	3720,616
travel permit 39341	12,759	15,311	1	15,311	-5,793
travel permit 423	923,081	57749,436	83	763,518	-164,583
UNKNOWN	NULL	117714,948	124	NULL	NULL

Especially interesting is the situation with caseId "travel permit 423", where the amount paid is ten times higher than the declared one with 83 payments. 449 declarations have been assigned to this permit (1-2 permit declarations are the norm). Perhaps this permission is technical and on its basis requests for payments in excess of the limit determined by the plan are created.

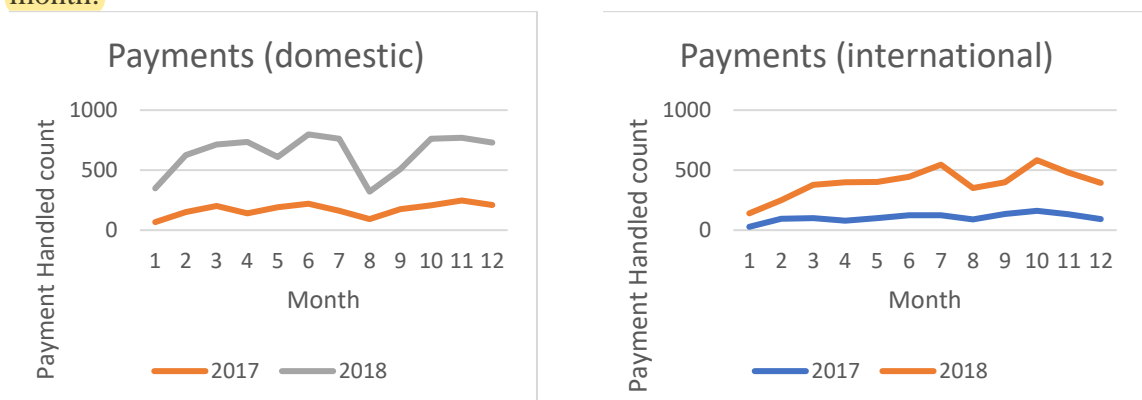
The analysis suggests that the system does not control the issue of overspending the limit of prepaid travel expenses during the trip. Such an assumption can entail at least difficulties in accounting for the organization's budget, and as a maximum ensure the possibility of irrational spending on the part of the employee (for example, when an employee instead of an economy class takes a business or stays in a hotel of increased comfort)

Recommendation: The process needs to be revisited so that the organization takes care of the transportation and reservations itself, or provides a list of travel options based on the travel limit. Such activity can be undertaken by external providers who monitor prices and availability and can always find the most comfortable and inexpensive travel conditions. This will reduce costs, as well as relieve employees of some of the difficulties in preparing a business trip.

3 Answers on questions

1) What is the throughput of a travel declaration from submission (or closing) to paying?

To analyze the issue, we examined the dependence of the number of payments per month:

**Fig. 7.** Payments count per month

The graphs show that by the end of the year and in August-September there is a noticeable decline in the number of payments. 2017 was a pilot year, the number of applications is significantly less. In 2018, an average of 640 payments was made per month for internal applications and 400 for external ones.

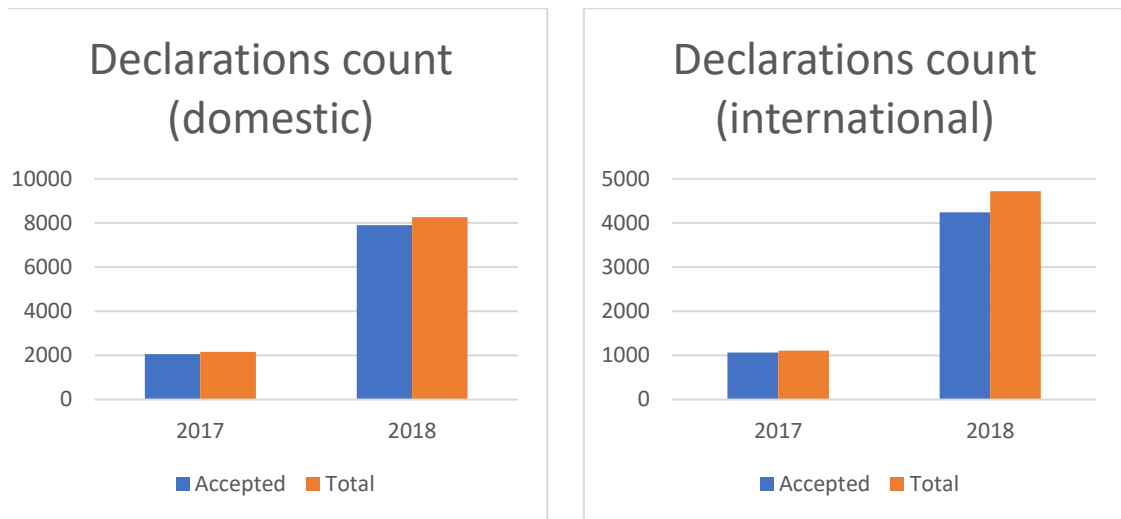


Fig. 8. Number of payment requests and the number of handled payments

2017: 2061 of 2155 (96.5%) requests reached payment. Median 7days 40min
 2018: 7903 of 8260 (96.5%) requests reached payment. Median time 8 days 1 hour

2017: 1065 of 1108 (96.1%) requests reached payment. median time 58d 7h
 2018: 4241 of 4720 (89.9%) requests made it to payment. Median 63d

The processing time of the request in 2018 increased by almost a day, this is due to an increase in the number of requests by 4 times and changes in the approval process.

Request processing time increased by 5 days in 2018. The number of rejected requests has more than doubled (3.9% -> 10.1%).

2) Is there are difference in throughput between national and international trips?

International declarations take much longer to process domestics. Median 64 days. The process includes the trip itself, the median time of which is 7 days, so we can say that the difference in the duration of the processes excluding the trip itself = $(64-7) - 8 = 49$ days.

There are 4 main stages in International Declarations: Obtaining a permit, prepayment, departure, filing a declaration, and its approval. Domestic only filing a declaration.

In international cases, the process of obtaining permission and approving the declaration in some cases takes less than 10 seconds. Most likely this is the submission of documents on behalf of the administrator (approver), or the process is partially automated.

3) Are there differences between clusters of declarations, for example between cost centers/departments/projects etc.?

By department:

Total 27 organizational units

Table 2. Statistics by org_unit(heatmap)

orgunit	Adjusted Amount summ	Declarations count	Mean Adjusted Amount	Mean trip duration	Mean Adjusted Amount per one day
Organizational unit with maximal Amount					
organizational unit 65458	1 158 829,04	1397	829,51	17,87	138,69
organizational unit 65454	781 934,75	911	858,33	5,20	193,15
organizational unit 65456	704 157,50	840	838,28	8,49	157,78
organizational unit 65455	630 676,92	1042	605,26	16,98	110,71
Organizational unit with minimal Amount					
organizational unit 65467	9 071,32	17	533,61	2,53	167,85
organizational unit 65468	8 631,25	10	863,13	4,70	164,85
organizational unit 65465	6 874,65	8	859,33	5,38	194,08
organizational unit 65471	5 716,26	7	816,61	4,14	215,86
organizational unit 65482	5 164,19	7	737,74	5,14	136,15
organizational unit 65473	3 452,61	12	287,72	3,67	86,08
organizational unit 65484	2 041,09	3	680,36	3,67	243,92
organizational unit 65477	1 976,35	4	494,09	4,00	110,29
organizational unit 65488	1 529,71	1	1 529,71	3,00	509,90
organizational unit 65486	641,24	1	641,24	8,00	80,16
organizational unit 65478	492,74	1	492,74	6,00	82,12
organizational unit 65462	178,83	1	178,83	3,00	59,61
Total	4 934 066,49	6 186	797,62	11,56	148,42

4 organizational units 65458, 65454, 65456, 65455 account for more than 66% of the international travel budget (3.276 million). And the number of trips in these departments is the maximum - 4190 declarations, 68%. But the indicators of expenses for 1 business trip day of these departments are not the maximum, which is associated with the duration of the trip - the more the number of business trip days, the lower the cost of 1 business trip day.

Example: Organizational unit 65458 has the highest number of trips and the maximum total amount. This department accounts for 24.5% of the total budget for 2017-2018. The indicator cost of 1 business trip day is 138.69, the average indicator for the system is 148.42.

At the same time, out of 10 departments with a minimum number of trips (from 1 to 9), 5 have an indicator of expenses per 1 business trip day in the TOP-10 and exceed this indicator in the system.

Example: Organizational unit 65488 - 1 international trip for 3 days. On average, there are 509.9 per day, which is more than 3 times higher than the average for the system - 148.42.

Proposal: Consider the possibility of replacing expensive short-term trips to participation in a remote format.

By cost center:

The Domestic Travel Cost Center keeps track of costs per BudgetNumber - budget 86566.

For international travel, there are BudgetNumber and Permit_BudgetNumber. Permit_BudgetNumber is the cost centers, BudgetNumber is the travel number for the departure group.

Permit_BudgetNumber is a reference, there are 207 values in it, we can assume that Permit_BudgetNumber corresponds to the purpose of the trip.

The most popular is budget 425, 455 trips (declarations with the Final Approved status) in the amount of 406 314.36 took place.

For each Permit_BudgetNumber, the cost planning accuracy was calculated - the difference between the amount for declarations and the Permit_RequestedBudget for Permit, per trip. For 123 Permit_BudgetNumber, this indicator ranges from -500 to 500 per trip. For 26, this figure exceeds 1000 and reaches 6562.38 for budget 11232, which speaks of planning excess amounts for travel.

Suggestion: update the calculation of the planned amount Permit_RequestedBudget. Use data from previous years for updating. Introduce monitoring of planning indicators and actual costs.

4) What is the throughput in each of the process steps, i.e. the submission, judgment by various responsible roles and payment?

To understand the main process, a graph of the top 4 popular routes was built:

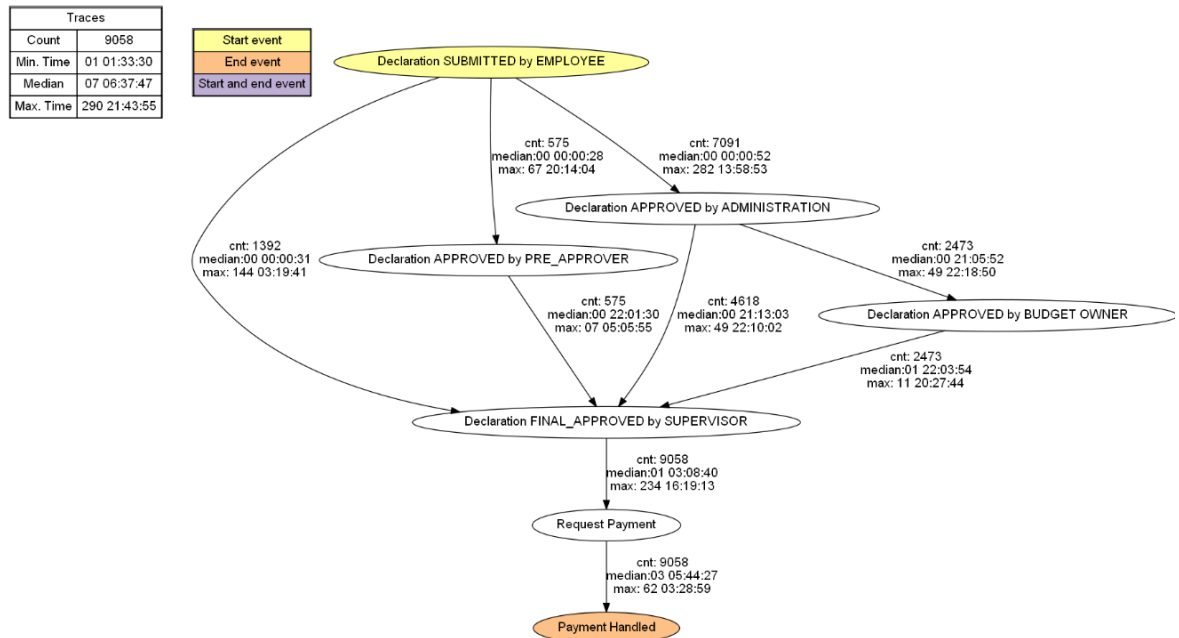
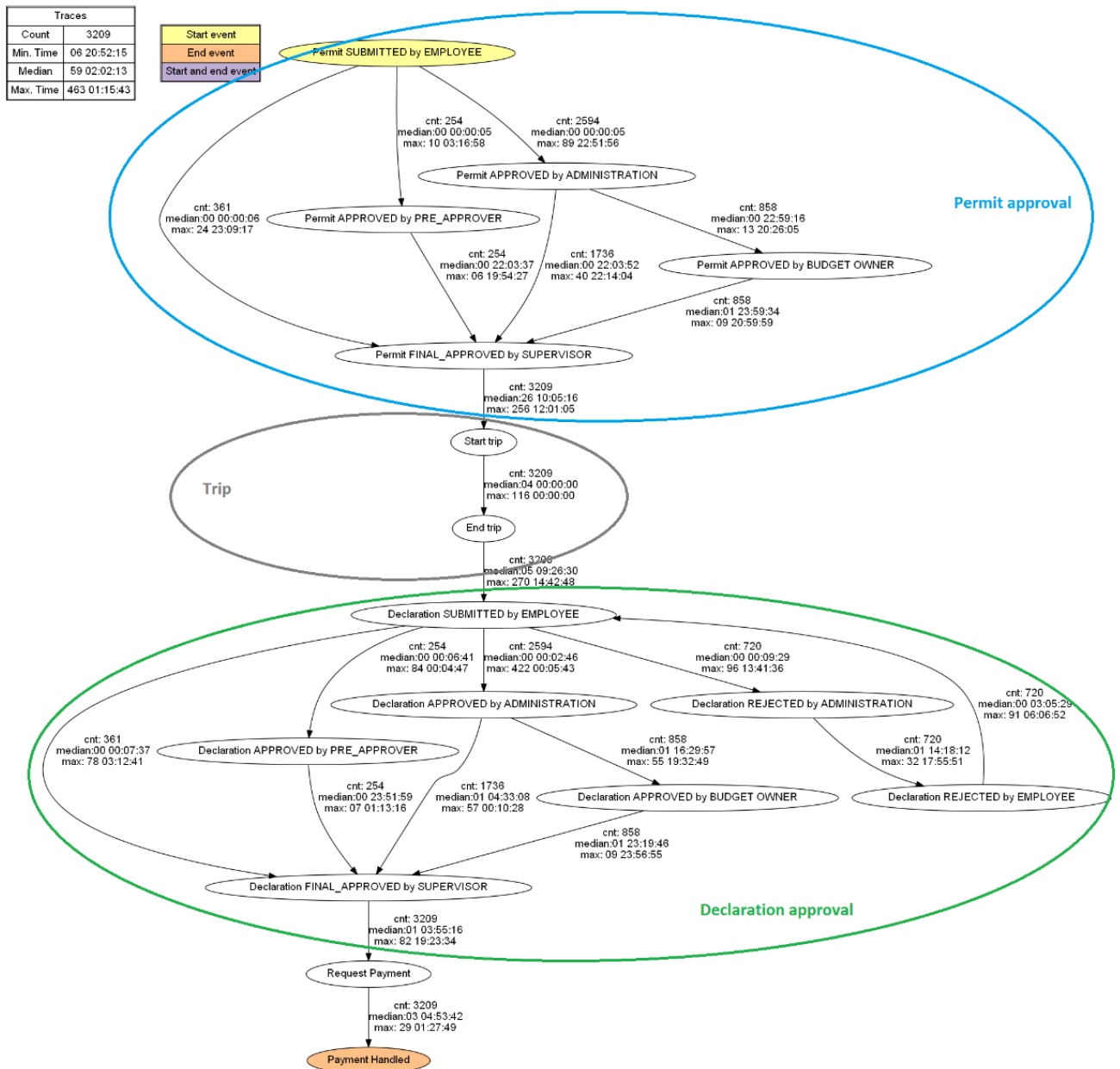


Fig. 9. Domestic declarations top 4 popular trace

Table 3. Domestic declarations event descriptions

Stage	Events	Desc	Duration
Submission of the Declaration	Declaration SUBMITTED by EMPLOYEE	The initial submission of the declaration, repeated submission of the declaration after refusal	-
Level 1 approval	Declaration APPROVED by ADMINISTRATION Declaration APPROVED by PRE_APPROVER	Approval by the ADMINISTRATION / PRE_APPROVER	med:41sec max: 282d
Level 2 approval (budget)	Declaration APPROVED by BUDGET OWNER	The activity appeared in 2018. Not necessarily if the first approver is the owner of the budget	med:20h max:49
Level 3 approval (director)		Не встречается	-
Level 4 approval	Declaration FINAL_APPROVED by SUPERVISOR	Process owner approval	med:9h max:144d
Not approved	Declaration REJECTED by ADMINISTRATION	Reject is possible at any of the approval stages. After reject, a second request is possible	med:0d 0:1:35
	Declaration REJECTED by BUDGET OWNER		med: 0d 9:56:23
	Declaration REJECTED by MISSING		med:3d 1:26:40
	Declaration REJECTED by PRE_APPROVER		med: 0d 0:1:05
	Declaration REJECTED by SUPERVISOR		med:0d 22:02:51
Cancellation by initiator	Declaration REJECTED by EMPLOYEE	The employee removes the request from approval	med:0d 22:07:56
Payment request, payment	Request Payment Payment Handled		med:4d 9:11:08

International declarations:



InternationalDeclarations.xes

Fig. 10. International declarations top 6 popular trace

Table 4. International declarations event descriptions

Stage	Events	Desc	Duration
Request for travel permit	Permit SUBMITTED by EMPLOYEE	Initial filing, resubmission of the request after refusal	-
Level 1 approval	Permit APPROVED by ADMINISTRATION Permit APPROVED by PRE_APPROVER	Approval by the administrations / pre_approver	med: 4sec
Level 2 approval (budget)	Permit APPROVED by BUDGET OWNER	Not necessary if the first approver is the budget owner	med: 0d 23:29:42
Level 3 approval	Permit FINAL_APPROVED by SUPERVISOR	Process owner approval	med:0d 21:38:29
Level 4 approval (director)	Permit APPROVED by DIRECTOR		med:0d 22:00:36
Not approved	Permit REJECTED by ADMINISTRATION	Reject is possible at any of the approval stages. After reject, a second request is possible	med:0d 04:04:04
	Permit REJECTED by DIRECTOR		med:20d 01:48:02
	Permit REJECTED by MISSING		med:0d 20:47:56
	Permit REJECTED by PRE_APPROVER		med:6sec
	Permit REJECTED by SUPERVISOR		med:0d 22:13:27
Cancellation by initiator	Permit REJECTED by EMPLOYEE	The employee removes the request from approval	med:1d 6:01:40
Trip	Start trip	Dates are approximate. There are no actual check-out dates	min:6d 17:31:11
	End trip		med:66d 01:00:00 max:742d
Submission of the Declaration	Declaration SUBMITTED by EMPLOYEE		-
Approved	Declaration APPROVED by ADMINISTRATION	often negotiation takes place in a couple of seconds. It looks like combining the roles of an employee and an Administrator.	min: 1sec med:0d 0:03:14
	Declaration APPROVED by BUDGET OWNER		min: 17 sec med:1d 3:06:27
	Declaration APPROVED by PRE_APPROVER		min: 2sec med:0d 00:04:14
	Declaration APPROVED by SUPERVISOR		min: 3sec med:0d 0:18:49
	Declaration FINAL_APPROVED by DIRECTOR		min: 0sec med:1d 4:00:37
	Declaration FINAL_APPROVED by SUPERVISOR		med:1d 3:35:17
Not approved	Declaration REJECTED by ADMINISTRATION		min: 1sec med:0d 00:09:27
	Declaration REJECTED by BUDGET OWNER		med: 1d 6:52:19
	Declaration REJECTED by DIRECTOR		med: 4d 21:25:50
	Declaration REJECTED by EMPLOYEE		med: 0d 3:03:27
	Declaration REJECTED by MISSING		med: 2d 19:01:29
	Declaration REJECTED by PRE_APPROVER		min: 4sec med: 0d 0:10:36
	Declaration REJECTED by SUPERVISOR		med: 2d 18:25:41
Secondary Submission of the Declaration	Declaration SUBMITTED by EMPLOYEE		med: 0d 3:14:23
Request payment, payment	Request Payment Payment Handled		med:1d 4:32:54

A short time for passing the approval (up to 10 seconds) says the approval of their requests on their own (the request is submitted by an employee with the approval role). In some cases, the approval takes much longer. It is also worth noting that APPROVAL by BUDGET OWNER always occurs in manual mode.

Recommendation: Employees should not reconcile their applications. This can lead to fraud. The requests must go to a superior manager for approval.

It is not uncommon for applications to freeze at various stages of approval. It is recommended that you define a maximum time for each step and send reminders to participants in the process to take action. In the absence of action, send repeated reminders and duplicate them to the manager.

5) Where are the bottlenecks in the process of a travel declaration?

Domestic:

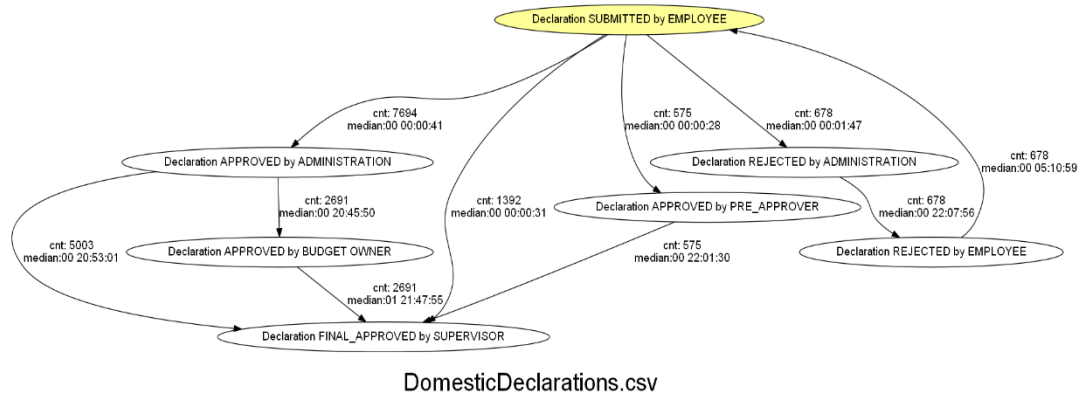


Fig. 11. Domestic top 6 popular traces (more 100 cases in every trace)

The longest steps are Budget Owner Declaration approval and Supervisor final approval. If the route includes a Budget Owner, the route increases in duration by an average of 1.5-2 days. Also, the process is idle if the declaration is rejected.

Recommendation: add a notification system for those responsible for the next step, in case of rejection - for the Employee role.

International:

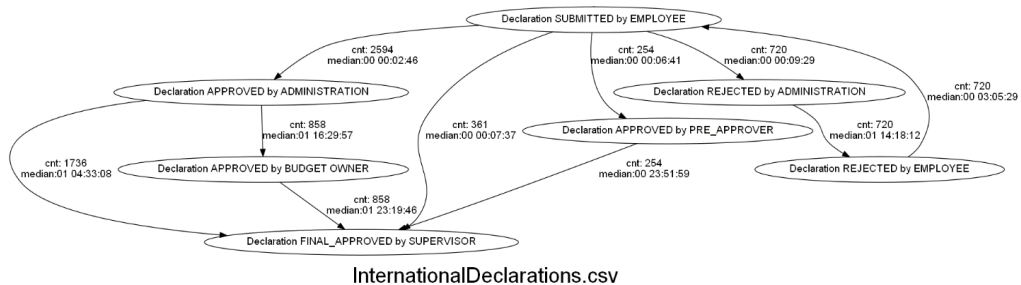


Fig. 12. International top 5 popular traces (main route - from permission to payment)

The routes are similar. Bottlenecks and recommendations are the same.

6) Where are the bottlenecks in the process of a travel permit (note that there can be multiple requests for payment and declarations per permit)?

Check-out start and end dates are a formality. Some employees, when preparing documents, indicate dates with a "large margin", for example, there are business trips lasting more than 1 year. This complicates the real picture and confuses the graph. these events can appear at any of the stages.

Recommendation: upon completion of the departure, oblige the employee to record the actual dates of the trip.

Within the framework of one obtained travel permit, several declarations for reimbursement of expenses can be submitted (there were cases in which 14 declarations were submitted within the framework of one permit). If the amount of compensation (RequestedAmount) is less than the allowed budget (Permit RequestedBudget), the approving employee may not analyze the previously submitted declarations and agree on the compensation even if the limit is exceeded.

Recommendation: At the stage of filing the declaration, automatically compare the amount of all submitted declarations within the travel authorization with the agreed budget.

Analysis of the duration of the main routes:

Below are the most popular routes in ascending order of duration. The Start trip and End trip activities have been removed to remove noise.

Route 1 (fastest, does not include third party approval)

Median duration	6 sec
Count traces	556
Graph	<pre> graph TD A([Permit SUBMITTED by EMPLOYEE]) -- "cnt: 556 median: 00 00:00:06 max: 24 23:09:17" --> B([Permit FINAL_APPROVED by SUPERVISOR]) </pre>
Date interval	09.01.2017 - 20.12.2017
Comment	since 2018 the scheme has been changed.

Route 2

Median duration	23 Hours
Count traces	2633
Graph	<pre> graph TD A([Permit SUBMITTED by EMPLOYEE]) -- "cnt: 2663 median: 00 00:00:04 max: 193 23:38:52" --> B([Permit APPROVED by ADMINISTRATION]) A -- "cnt: 470 median: 00 00:00:05 max: 262 23:02:12" --> C([Permit APPROVED by PRE_APPROVER]) B -- "cnt: 2663 median: 00 22:16:46 max: 40 22:14:04" --> D([Permit FINAL_APPROVED by SUPERVISOR]) C -- "cnt: 470 median: 01 00:02:09 max: 06 20:32:45" --> D </pre>
Date interval	10.01.2017 - 17.12.2018
Comment	Adding a verification step by the Administrator \ Pre approver increased the duration by 23 hours.

Route 3

Median duration	5 days 10 hours
Count traces	1462
Graph	<pre> graph TD A([Permit SUBMITTED by EMPLOYEE]) -- "cnt: 1462 median:00 00:00:05 max: 76 00:55:21" --> B([Permit APPROVED by ADMINISTRATION]) B -- "cnt: 1462 median:00 22:18:19 max: 20 20:13:54" --> C([Permit APPROVED by BUDGET OWNER]) C -- "cnt: 1462 median:02 07:07:44 max: 44 20:18:24" --> D([Permit FINAL_APPROVED by SUPERVISOR]) </pre>
Date interval	06.01.2018 - 18.12.2018
Comment	This scheme has been introduced since the beginning of 2018. Adding an approval stage by the Budget owner increased the time of processing by more than two days.

It can be concluded that the more approvers, the longer the supervisor's approval is the bottleneck (2 times longer than the administrator checks). A situation is possible when, with a uniform flow of requests, the supervisor will not have time to issue permissions.

7) How many travel declarations get rejected in the various processing steps and how many are never approved?

Domestics declarations
Total cases: 10500

Table 5. Domestics declarations reject statistics

Reject stage	REJECTED count, %	Next "Declaration REJECTED by EMPLOYEE"
Declaration REJECTED by ADMINISTRATION	952 / 9%	204
Declaration REJECTED by MISSING	91 / 0.9%	0
Declaration REJECTED by PRE_APPROVER	86 < 0.8%	8
Declaration REJECTED by EMPLOYEE (without prior rejection)	7 / 0%	1
Declaration REJECTED by SUPERVISOR	293 / 2.7%	73
Declaration REJECTED by BUDGET OWNER	59 / 0.5%	14
Total:	1488 / 14%	300 / 2.9%

134 requests have not moved from the Declaration SAVED by EMPLOYEE stage

10043 (95.6%) requests reached the Payment Handled stage, some remained incomplete (freeze in the rejection stage), so we can conclude that $10500 - 10043 = 457$ requests (4.4%) remained unfinished

International declarations:

Total cases: 6 449

8 (<1%) did not reach the stage of submitting the declaration

Table 6. International declarations reject statistics

Reject stage	Reject count, %	Next "Declaration REJECTED by EMPLOYEE"
Declaration REJECTED by ADMINISTRATION	1287 / 20%	140
Declaration REJECTED by MISSING	98 / 1.5%	0
Declaration REJECTED by PRE_APPROVER	82 / 1.3%	8
Declaration REJECTED by EMPLOYEE (without prior rejection)	0	0
Declaration REJECTED by SUPERVISOR	122 / 1.9%	15
Declaration REJECTED by BUDGET OWNER	40 / 0.6%	8
Declaration REJECTED by DIRECTOR	4 / 0%	1
Total:	1633 / 25%	2.667080167

6187 (96%) requests reached a payment

We also noticed that there is an extra step, in our opinion, "Declaration REJECTED by EMPLOYEE" when editing an application. The general scheme for editing a declaration is as follows:

Table 6. Domestic declarations correction scheme

way	submit1	reject1	reject2	submit2
1	SUBMITTED by EMPLOYEE	REJECTED by ADMINISTRATION	REJECTED by EMPLOYEE	SUBMITTSED by EMPLOYEE
		REJECTED by BUDGET OWNER		
		REJECTED by DIRECTOR		
		REJECTED by SUPERVISOR		
2	SUBMITTED by EMPLOYEE	REJECTED by EMPLOYEE		SUBMITTED by EMPLOYEE

Here's an example for the 2018 Domestic declarations edit process:

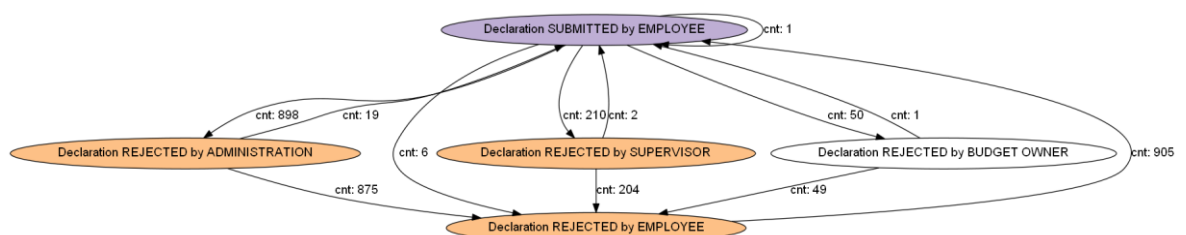


Fig. 13. Domestic declaration correction (2018)

If any of the participants in the process rejects the declaration which needs to be corrected, then first the employee must also reject it and only after that he can re-submit the declaration.

Recommendation: it is necessary to consider the possibility of excluding the "Declaration REJECTED by EMPLOYEE" step when editing declarations so that after the "Declaration REJECTED by SOMEBODY" status, "Declaration SUBMITTED by EMPLOYEE" immediately follows. "Declaration REJECTED by EMPLOYEE" status can be left only for the case of a complete cancellation of the declaration by the employee. The elimination of unnecessary steps from the process makes the process more interpretable, faster and more fault-tolerant.

More detailed questions:

1) How many travel declarations are booked on projects?

The following checks were used to determine whether the international declaration is booked for the project:

- According to the declaration, the project number must be determined;
- Advance payment has been made on the declaration - the presence of data in PrepaidTravelCost.xes;
- The permit and declaration must be finally approved.

As a result, it was possible to determine - a total of 1254 international declarations were submitted for prepaid (362 projects). Of them:

- 63 declarations did not receive final approval for prepaid (37 projects);
- 10 declarations were canceled by the events "Declaration REJECTED by MISSING" or "Permit REJECTED by MISSING" in 2017 (8 projects);
- prepaid was received for 427 declarations, but no approvals were received (1 project).

As a result, 754 declarations (357 projects) received correct approvals and prepaid, i.e. are considered booked for projects.

2) How many corrections have been made for declarations?

For making changes to the declaration, we will make the transition from "refusal of declaration" to "submit a declaration by an employee", i.e. next transition:

Declaration REJECTED BY SOMEONE => Declaration SUBMITTED BY EMPLOYEE

Table 8. International declarations reject counts

reject_init	international_cnt	domestic_cnt	total
by ADMINISTRATION	24	19	43
by BUDGET OWNER	0	1	1
by DIRECTOR	0	not_included	0
by EMPLOYEE	1610	1081	2691
by MISSING	87	61	148
by PRE_APPROVER	0	1	1
by SUPERVISOR	3	2	5
	1724 (by 1405 cases)	1165 (by 1019 cases)	2889 (by 2423 cases)

Thus, answering the question, we can say that for the entire period of available logs, there were 2889 edits of declarations (of which 1724 edits were made for 1405 international business trips and 1165 edits were made for 1019 domestic business trips). Some declarations have been edited more than once.

As you can see, most edits are initiated by the employees themselves, but this was not always the case. Since we know that 2017 was a pilot year, let's take a look at what changed in 2018:

Table 9. Declarations change initiators by year

change initiator	International trips			Domestic trips		
	2017	2018	2016-2020	2017	2018	2016-2020
ADMINISTRATION	not_included	23	24	not_included	19	19
BUDGET OWNER	not_included	0	0	not_included	1	1
DIRECTOR	0	0	0	not_included	not_included	not_included
EMPLOYEE	84	1406	1610	113	920	1081
MISSING	0	not_included	87	56	not_included	61
PRE_APPROVER	0	not_included	0	1	not_included	1
SUPERVISOR	1	1	3	0	2	2

Looking at the data, the following conclusions can be drawn:

- 1) The main part of the log is data for 2017-2018. Data for 2019-2020 is available, but relatively few.
- 2) Domestic trips do not include the stage of coordination with DIRECTOR
- 3) In 2018, along with the increase in the total number of business trips, there is an increase in the number of cases of editing declarations by EMPLOYEE
- 4) In 2018, such statuses as «REJECTED by MISSING» and «REJECTED by PRE_APPROVER» left, but new statuses appeared – «REJECTED by ADMINISTRATION» and «REJECTED by BUDGET OWNER» (valid for all types of business trips)

You can roughly calculate the date of the process change by finding out when the «REJECTED by PRE_APPROVER» and «REJECTED by MISSING» statuses ceased to appear in the logs and the «REJECTED by ADMINISTRATION» and «REJECTED by BUDGET OWNER» statuses began to appear:

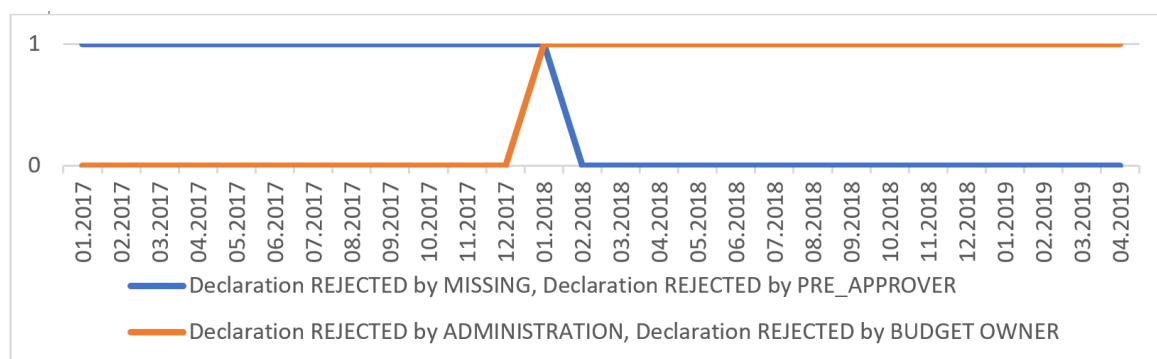


Fig. 14. Change process period

January 2018 is a transitional year. In particular, the last transition "REJECTED by MISSING" => "SUBMITTED by EMPLOYEE" is dated 01/08/2018. This date can be considered the end of the pilot process.

3) Are there any double payments?

Since it is possible to link the entities of prepaid travel costs and international declarations through travel permits, an attempt has been made to identify duplicate payments. There were none.

Domestic declarations, international declarations and requests for payment included payments with the same amount, in some cases with the same payment time. But such payments did not cause suspicion.

There was not a single case that two payments were made on one declaration or on one request for payment.

Among the prepaid travel costs, payments with the same requested amount were identified, which may meet the definition of double payments:

Table 10. Double payments

Log	Case id	Requested amount	Distinctive feature
Prepaid Travel Cost	request for payment 186361 request for payment 186366 request for payment 186371	210.21	Permit unknown
	request for payment 30540 request for payment 24355	16.13	Same time of payment handled, different permit
	request for payment 185607 request for payment 185746	45.43	Same time payment handled, permit unknown
	request for payment 185938 request for payment 186058	19.96	Permit unknown

4) Are there declarations that were not preceded properly by an approved travel permit? Or are there even declarations for which no permit exists?

The declaration submission event is the "Declaration SUBMITTED by EMPLOYEE" activity. In the DomesticDeclarations.xes and InternationalDeclarations.xes logs, in all cases for this activity, org_resource is specified as STAFF MEMBER, and org_role is equal to EMPLOYEE. This means that when this activity is recorded, the real role of the employee is not saved. Therefore, it is necessary to identify other signs of a substitute reaction for an employee by a mandated person.

After analyzing all cases where there is at least one declaration submission event (Declaration SUBMITTED by EMPLOYEE), we get subsets - for DomesticDeclarations.xes it is 10365 cases out of 10500, and for InternationalDeclarations.xes it is 6375 cases out of 6449. Next, we calculate the time intervals between the last event of submission of the declaration (Declaration SUBMITTED by EMPLOYEE) and the next event after it with org_resource = STAFF MEMBER and org_role != EMPLOYEE, i.e. the next event is set by a mandated person, not an employee (EMPLOYEE) and not a system (SYSTEM).

Quite short periods of time may mean that the declaration was filed by a mandated person (not an EMPLOYEE or SYSTEM) for the employee.

We divided the data into clusters according to the duration of the transition (0 - 5 seconds, 5 - 30 seconds, 30 - 90 seconds, more than 90 seconds) and presented in the form of graphs.

Table 11. Time intervals between the last submit and following event declaration from a mandated person in the DomesticDeclarations.xes log

Second Activity	0-5 sec	5-30 sec	30-90 sec	>90 sec	all
Declaration APPROVED by ADMINISTRATION	1963	1889	939	3185	7976
Declaration FINAL_APPROVED by SUPERVISOR	364	392	205	526	1487
Declaration APPROVED by PRE_APPROVER	142	211	69	243	665
Declaration REJECTED by ADMINISTRATION	35	49	28	89	201
Declaration REJECTED by SUPERVISOR	5	6	5	9	25
Declaration REJECTED by PRE_APPROVER	2	0	2	4	8
Declaration FOR_APPROVAL by PRE_APPROVER	0	0	0	1	1
Declaration FOR_APPROVAL by SUPERVISOR	1	0	0	0	1
Declaration REJECTED by EMPLOYEE	0	1	0	0	1

In DomesticDeclarations.xes, there is an exclusive case "declaration 119402" with a declaration submission event (Declaration SUBMITTED by EMPLOYEE), according to which the next event after the extreme event of submission of the declaration is an event committed by the employee (Declaration REJECTED by EMPLOYEE). We attribute this case to the actions of the employee and it will not be displayed on the diagram.

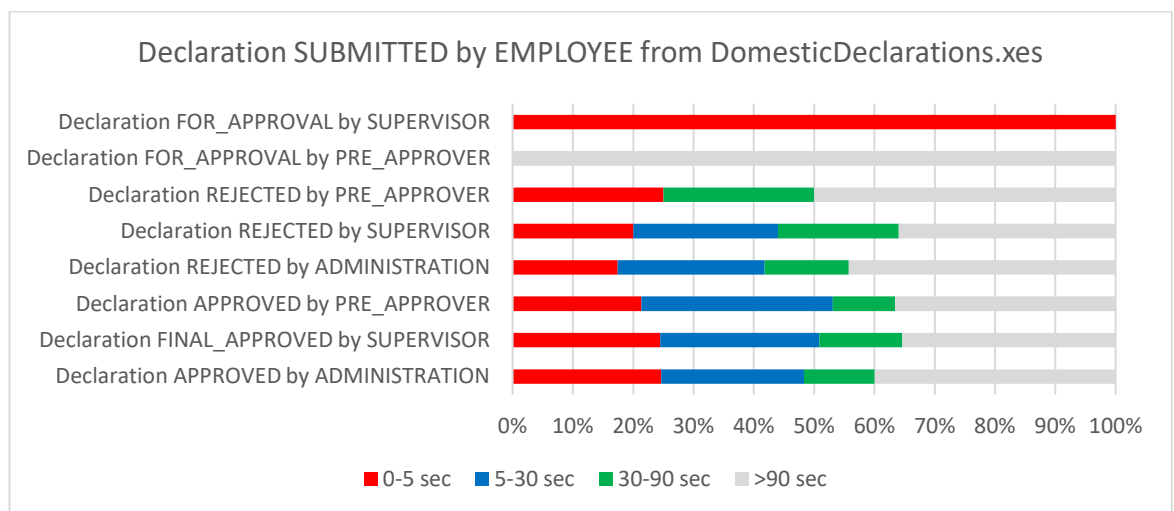


Fig. 15. Distribution of time intervals between the events by domestic declaration

Table 12. Time intervals between the last submit and following event declaration from a mandated person in the InternationalDeclarations.xes log

Second Activity	0-5 sec	5-30 sec	30-90 sec	>90 sec	all
Declaration APPROVED by ADMINISTRATION	685	1069	471	2693	4918
Declaration FINAL_APPROVED by SUPERVISOR	66	148	41	416	671
Declaration APPROVED by PRE_APPROVER	56	147	59	291	553
Declaration REJECTED by ADMINISTRATION	19	34	9	77	139
Declaration APPROVED by SUPERVISOR	6	13	0	61	80
Declaration REJECTED by PRE_APPROVER	0	2	2	6	10
Declaration REJECTED by SUPERVISOR	2	0	0	0	2
Declaration FINAL_APPROVED by DIRECTOR	0	0	0	1	1
Permit REJECTED by MISSING	0	0	0	1	1

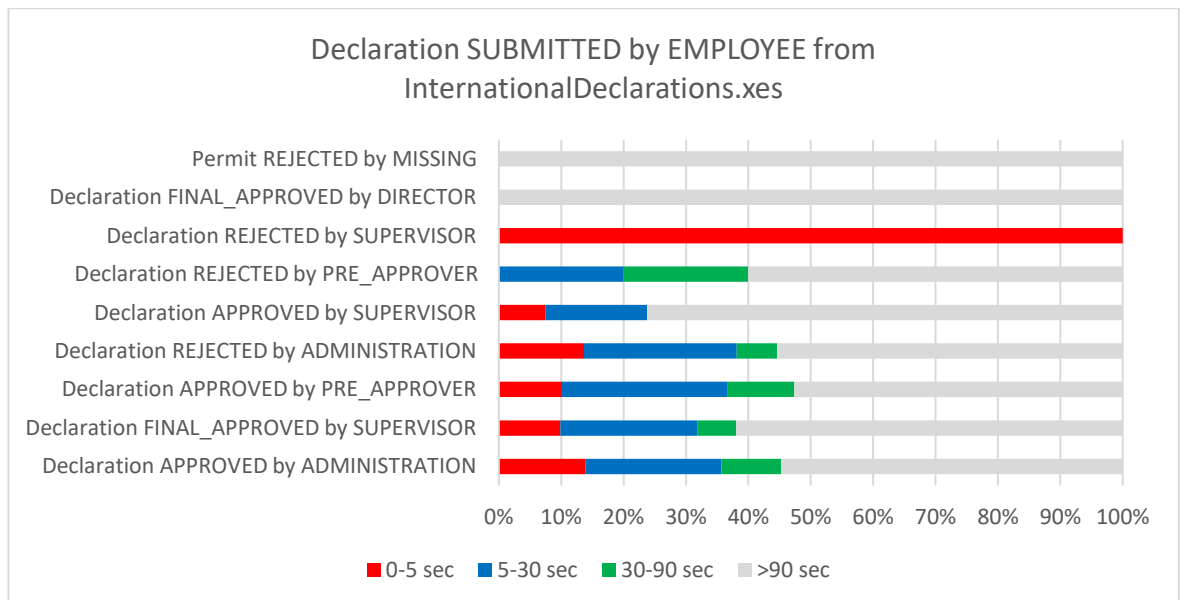


Fig. 16. Distribution of time intervals between the events by international declaration

As a result, we will count the number of all cases of submitting declarations, classified by clusters.

Table 13. The number of all cases of submitting declarations

Second Activity	0-5 sec	5-30 sec	30-90 sec	>90 sec	all
Declaration SUBMITTED (Domestic)	2512	2547	1248	4057	10364+1 (exception)
Declaration SUBMITTED (International)	834	1413	582	3546	6375

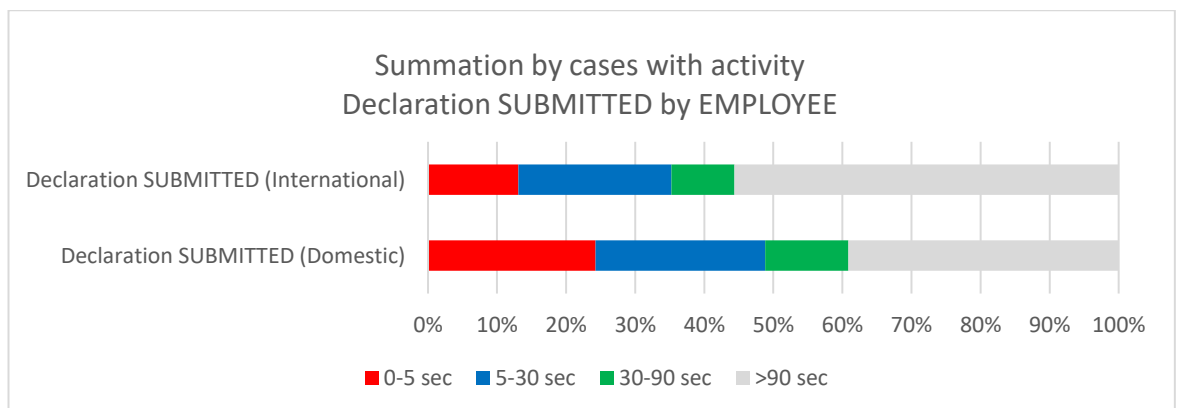


Fig. 17. Final distribution of time intervals by submitting declarations

Assuming that all time intervals of less than 90 seconds between two events (the last Declaration SUBMITTED by EMPLOYEE and any subsequent event from a mandated person (not EMPLOYEE or SYSTEM)) are the submitting of a declaration by a mandated person instead of an employee, we get the final results:

- Submission of declarations for internal travel - 4058 from the employee (taking into account one exclusive case "declaration 119402"), 6307 from a mandated person;
- Submission of declarations for international travel - 3546 from the employee, 2829 from a mandated person.

5) How many travel declarations are submitted by the traveler and how many by a mandated person?

Because no travel permit is required for domestic declarations, only international declarations have been analyzed.

Key features for analysis:

- The submission of the declaration is considered correct if by this time the travel approval by the head or director has been submitted and finally approved;
- The final approval of the international declaration is considered correct if by that time the travel permit has also been finally approved.

Log analysis results:

- From the list of all cases InternationalDeclarations.xes received 448 cases that lack the process of submitting travel permit:
 - a) For 427 cases, the final approval of the international declaration was received;
 - b) For 11 cases, international declarations were rejected;
 - c) Drafts of international declarations were created for 10 cases without sending them for approval;
- Also, one case was identified in which the final approval of the international declaration was received after the rejection of the approval (Permit REJECTED by MISSING) for the trip.

Conclusions:

- One international declaration (caseId = declaration 6199) was issued with a rejected travel permit;
- 448 international declarations were filed without an approved travel permit, of which 427 were finally approved, 11 were rejected and 10 are drafts.

6) How many travel declarations are first rejected because they are submitted more than 2 months after the end of a trip and are then re-submitted?

During the reviewed period, 276 declarations were sent after more than 2 months of the planned departure date (from 63 to 398 days of delay). Of them:

- 69 were rejected (24%), were re-submitted and as a result passed for final agreement and payment of expenses;
- 207 were agreed without deviations (76%);

Since March 2018, reminders have been sent to employees about the need to submit a declaration for international travel. A reminder is sent monthly from the 1st day, for departures completed more than 1 month ago. After that, the share of declarations sent more than 2 months later decreased by almost 2 times.

The figure shows diagrams of the distribution of the period of detention for the submission of declarations.

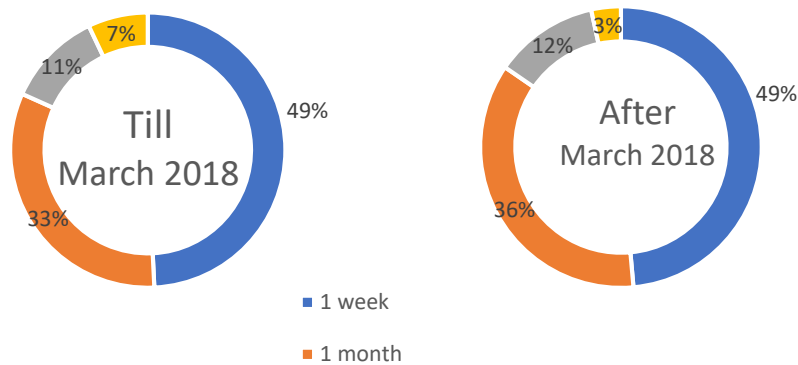


Fig. 18. Distribution of the period of detention

Recommendation: To improve the effectiveness of reminders (reduce the number of declarations of subjects after 2 months or more), we recommend changing the frequency of reminders.

For example, weekly after the end of the check-out, with an indication in the copy of the administrator who approved the trip.

7) Is this different between departments?

The issue was considered in 3rd question of answers block

8) How many travel declarations are not approved by budget holders in time (7 days) and are then automatically rerouted to supervisors?

Domestic declarations:

125 (1%) requests were processed by the Budget owner for more than 7 days

672 (6%) applications were sent to a supervisor bypassing the Budget owner's approval (after 7days +)

International declarations:

207 (3%) requests were processed by the Budget owner for more than 7 days

958 (15%) requests were sent to a supervisor bypassing the Budget owner's approval (after 7 days +)

Automatic redirection to a supervisor does not work in all cases. These cases are quite rare (less than 3% of cases), but they can lead to significant delays in coordination.

9) Next to travel declarations, there are also requests for payments. These are specific for non-TU/e employees. Are there any TU/e employees that submitted a request for payment instead of a travel declaration?

There are cases in the Permit Log where Request for Payment ends in reject by ADMINISTRATION and then reject by EMPLOYEE follows. Next, the declaration is drawn up in accordance with the process. Each case is different, but the main feature is the following part of the process:

Request For Payment SUBMITTED by EMPLOYEE >> Request For Payment REJECTED by ADMINISTRATION >> Request For Payment REJECTED by EMPLOYEE >> Declaration SUBMITTED by EMPLOYEE

There were about 38 such examples. They differ greatly in steps. The process usually ends with Payment Handled.

4 myPM – software for ProcessMining

Within the framework of the competition, software for the analysis of magazines was developed. We named it myPM.

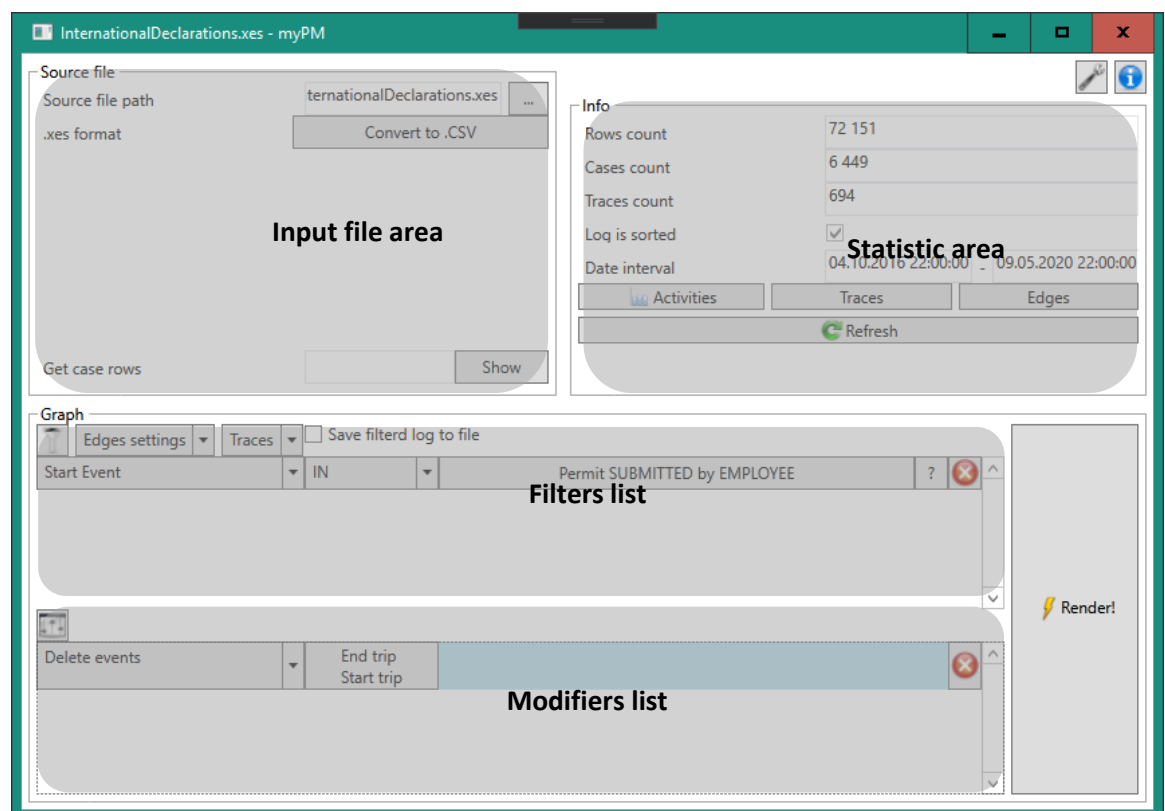
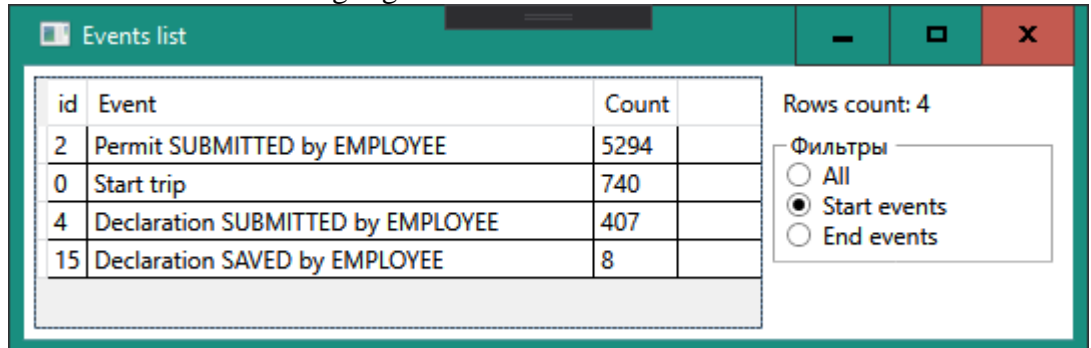


Fig. 19. myPM main window

Using myPM we filtered/modified logs, built graphs and made conclusions. myPM provides a minimalistic user interface, is well optimized, allows you to apply filters/modify the log, etc. We also added a function to export .xes to .csv format for loading into a database and exploring data using SQL.

myPM Key Features List:

- Handling .xes\,csv files
- GUI: building a graph in a few clicks
- The entire log is not loaded into RAM at once, which allows you to process logs larger than the amount of RAM
- Calculating statistics
 - Events counts
 - Cases count
 - Min\max event date
 - List of activities with highlighted start and end



id	Event	Count
2	Permit SUBMITTED by EMPLOYEE	5294
0	Start trip	740
4	Declaration SUBMITTED by EMPLOYEE	407
15	Declaration SAVED by EMPLOYEE	8

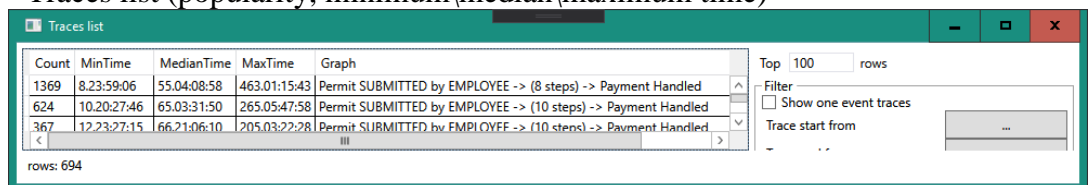
Rows count: 4

Фильтры

- ☐ All
- ☒ Start events
- ☐ End events

Fig. 20. List of activities window

- Traces list (popularity, minimum\median\maximum time)



Count	MinTime	MedianTime	MaxTime	Graph
1369	8.23:59:06	55.04:08:58	463.01:15:43	Permit SUBMITTED by EMPLOYEE -> (8 steps) -> Payment Handled
624	10.20:27:46	65.03:31:50	265.05:47:58	Permit SUBMITTED by EMPLOYEE -> (10 steps) -> Payment Handled
367	12.23:27:15	66.21:06:10	205.03:22:28	Permit SUBMITTED by EMPLOYEE -> (10 steps) -> Payment Handled

rows: 694

Top 100 rows

Filter

☐ Show one event traces

Trace start from

Fig. 21. List of traces window

- Edges list (popularity, minimum\median\maximum time, visualization in the form of graphs)

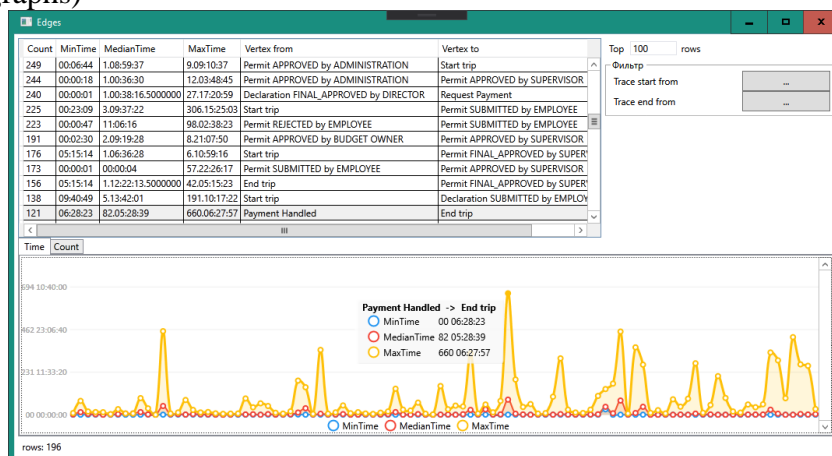


Fig. 22. List of edges window

- Log filtering. List of available filters in the picture

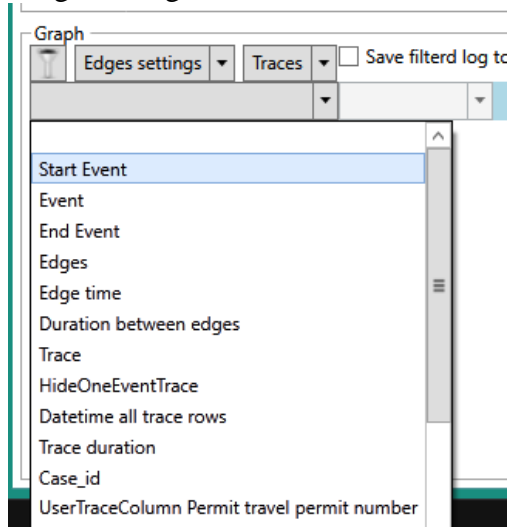


Fig. 23. Log filters list

- Log modifiers
 - Trim log
 - Remove events
- Rendering the graph
 - Display quantity, minimum\median\maximum edge time
 - Highlighting start and end events
 - Display of the number of routes, minimum\median\maximum duration of the route

The filtered/modified log can be saved to a .csv file and processed separately by myPM. This allows you to decompose the task.

myPM is written in c # (.net Framework 4.5.2), uses the GraphViz library to draw graphs. Performance tested on Windows 10 x64

You can try our development. Project link: <https://github.com/Freem27/myPM>

5 Conclusion

As a result of the analysis of five files with information about events, a description was compiled, and then an analysis of the processes of approval and registration of cost recovery. In addition to this, the software was created, which was actively used and will be further developed as an independent project. In addition to the created software, Process Mining methodologies and basic Data Analysis methods were widely used. This includes the Python programming language and its libraries - Pandas and pm4py, the SQL language.

The data provided contained limitations that did not allow for more complete coverage and analysis of all events as part of a single system. The lack of connections between some entities devalued their contribution to the overall picture.