Process Mining techniques in complex Administrative Processes

Jan Suchy, Milan Suchy

GRADIENT ECM, Kosicka 56,
82108 Bratislava, Slovakia
{Jan.Suchy, Milan.Suchy}@gradientecm.com

Abstract. This research demonstrates the process mining techniques used for analyzed data in the five Dutch municipalities. Thirteen roles were identified between employees involved in various stages of the processes. Organizational structures were identified and an algorithm was developed for the purpose of creating an organizational chart for each municipality. Possible points of improvement were made evident for an organizational structure, their organizational charts, as well as their allocated responsibilities for their tasks. With the help of four custom metrics, we were able to identify the differences in throughput times between all municipalities. A control flow was presented for each municipality and similar as well as different behaviors were identified. Our findings indicate what effects outsourcing some of the procedures will have on the organizational structure, as well as the employee roles involved in the process. Finally, we highlighted the changes in the processes that could possibly be due to the relocation of employees.

Keywords: Process Discovery; Administrative processes; Organizational chart; Performance Measurement of Processes; Process Mining

1 Introduction

In today‘s modern society, current business processes tend to become increasingly complex. One of the reasons being, as business processes increase their use for technology also increases making decisions to many multi-level environments. This complexity often grows to the extent that none of the involved decision makers are able to have a total overview of the complete end-to-end processes [1]. We can gain immense insights in complex scenarios/cases with technology where the average human being does not have the means to deduce all factors, predictions, observations, and conclusions by themselves. By using technology that is far more advanced than what our standard means as a human can operate at, we can perform higher than average, even exceed expectations as we can observe more details through research that can be repeated countless times without having any strain on us as people. One way in which we can provide process owners an “as-is” overview/outlook on processes within an organization is to use process mining techniques [2].

With the gained use of process mining techniques and developed algorithms on data [3], [4], [5], [6], [7] provided by five Dutch municipalities as a part of the 2015 Business Process Intelligence Challenge [8], we were able to provide process owners crucial finds and in essence answers to the questions asked. We did not only answer the questions that were asked, but even more so, we did our very best to exceed and convey more in-depth information to each answer for a better understanding of the process.

For municipalities, each of the identified roles of the people involved in various stages of the process were examined. Roles for people were identified not only for stages of the process but also for the contributing attributes of the all analyzed processes. Findings contributing to the possible points of improvement in the organizational structure, besides the structure being shown itself, we developed an algorithm to create an organizational chart that we could reflect the roles of obtained attributes. Analyzed results were also those that pertained to a social network based on independent municipalities and their identified overloaded resources.

Differences in throughput times between the municipalities were identified through four various procedures. For each procedure, specific areas were identified by which independent municipalities were recorded with highest throughput times.

Process activities were presented for each municipality, as well as together and also showing their differentiations among each other. Research was able to indicate municipalities that were similar yet also different in their modes of processing building permits through modified variants. Control flow was presented by showcasing composite events. For each municipality, a process map was created to visualize the findings that could compare and contrast similar indicators in each control flow between municipalities. The effects on organizational structures by outsourcing some procedures were later made evident. Also to be viewed in this paper are identified influences on the employee roles.

Lastly, we have shown the changes in the processes that could possibly influence and conclude to when the employees of two of the municipalities have physically moved into the same location.

2 Tools

In order to achieve the goal of this challenge, we used process mining tool, data mining tool, database management system, and spreadsheet applications for purposes of understanding the data better and be able to answer that process owners questions.

2.1 RapidMiner

We used RapidMiner 6.4 (Starter Edition) as a data mining tool. RapidMiner is a widely used data mining toolset which provides us with rich data visualizations and statistical information about attributes of datasets.

2.2 Minit

Minit is a process mining tool by which we can import datasets for a wide range of possibilities in analyzing processes precisely. With this tool, we were able to analyze all process models in the building permit applications and their statistical characteristics. What was of great importance in our analysis was to analyze the social network by which we could bring out the fine details of each zoned resources in all processes. The greatest advantage was also that Minit has the capability to produce custom charts through which we could analyze data and compare and contrast their characteristics based on any needs.

Minit was used also for data validation of partial results throughout the development of algorithms and methods.

2.3 Graphviz

In order to visualize organizational charts that reflected organizational structures, we used an open source graph visualization software Graphviz.

2.4 Microsoft Excel, .NET Framework and C# language

To demonstrate results which we have achieved during this challenge, we choose Microsoft Excel (Microsoft Office 2013; Microsoft Corporation). This spreadsheet application is useful to draw charts and offers a lot of “Add-Ins” suitable for data handling.

Purpose of this challenge was to answer to all process owners questions and provide a new outlook on analyzing complex administrative processes. In order to meet the target, we needed to apply different data preprocessing, conversions, and mathematical functions. This is why we decided to use .NET 4.5 programing framework and C# language.

3 Data preparation

For this year’s challenge the logs from five Dutch municipalities were prepared. The logs contain information of building permit applications. Each of this five logs indicate different steps which are to be taken in order to process the buildings permits together with other information logged in 12 case attributes and 11 event attributes. The time period captured by datasets is approximately four years. Each step which is to be taken for the issuing of building permit is represented by codes and specifically labelled in English and Dutch. Process instances registered in logs describe the processing of different building permits. Such a process is represented as a flow of events throughout the application, which consists of a main process and subprocesses.

The original files were imported to the Minit software, where all case and event level attributes were defined as well as loaded into an SQL database in order to perform further analysis. Visualizations and summary calculations were processed by Minit. Table 1 contains summary of insights on analyzed municipalities.

Table 1. Shortened datasets information.

	M1	M2	M3	M4	M5
Log	BPIC15_1	BPIC15_2	BPIC15_3	BPIC15_4	BPIC15_5
Cases	1 199	832	1 409	1 053	1 156
Events	52 217	44 354	59 681	47 293	59 083
Start	5.2.2010	26.9.2010	1.1.2010	18.11.2009	23.11.2009
End	1.8.2015	4.3.2015	5.3.2015	5.3.2015	3.3.2015
taskNameEN	289	304	277	272	285
action_code	398	409	383	355	386

After the initial processing of data we were able to identify the main process flow and its aim in comparison to subprocesses. After detailed recognition of processes performed in different municipalities, we decided to apply the attribute “action_code” as an identifier of activities. We chose this approach because it is obvious already from the basic characteristics of logs that the number of unique values of attribute “action_code” differs from the number of unique values of labels (attribute taskNameEN, taskNameNL). This results in the following situation, where one label represented more than one “action_code”. An example given, in every municipality the label “procedure change” represented codes 01_HOOFD_180 and 01_HOOFD_330.

4 Roles in municipalities

4.1 Searching roles in attributes

To answer the question on roles in different municipalities we primarily searched for the common characteristics of resources according to various attributes. At first, we started searching within the case level attributes, which have the same value for every event in a case. As a first attribute we chose “(case) parts”. Due to higher number of unique values which occurred rarely we focused only on the more frequented values of this attribute. The result of this selection was that it is not possible to clearly identify roles by a given attribute, because the majority of resources deals with building permits where the value of attribute “(case) parts” is mostly “Bouw”, “Kap”, “Sloop”. Small differences were noticed in values which contained word “Milieu”, as we found resources which did not deal with such values. Actually, we found that resource “560596” in Municipality-5, has only dealt with cases with the value “Kap”. Therefore, this attribute it is not sufficient to identify the role of the resources.

We obtained the same result after the observation of attributes “(case) caseStatus”, “(case) caseProcedure”,”(case) termName”,”(case) last_phase” and “(case) Includes_subCases”. Nevertheless, we have encountered a few interesting aspects while analyzing the attribute “(case) Responsible_actor” even though we have not used this attribute for characterization of roles. In different municipalities we can find resources, which also bear the role of the “Responsible_actor”. On the other hand, there are also some “Responsible actors”, which have never acted as a resource (e.g. in

Municipality-1 - 4901428). In Municipality-1 for example, we can identify employee 560464, who acted as a resource in 1% of all cases, while the same employee also acted as a “Responsible actor” in 26% of all cases. In Municipalities-2/3/4/5 we can find such employees who acted as a resource and at the same time also as a “Responsible actor” for themselves (Municipality-4 – 560852, the employee was active in 99% of his activities where he were also a resource and “Responsible actor”). However, this may be caused also by smaller number of employees. To shine light on the outcome from the analysis of this attribute, we identified employees who act as the following: only as a resource, only as a “Responsible actor”, as a resource and as a “Responsible actor” at the same time. Subsequently, we have tried to define roles by event level attributes. To begin, we have chosen an attribute “question”. Once again, we have focused only on the most frequent values due to a huge number of the latter. The outcome was identical in the result for case level attributes “(case) parts”. Attribute “monitoringResource” was next to be examined. While analyzing this attribute, we have encountered the same behavior as previously mentioned in “(case) Resonsible_actor”. Also in this attribute, were found employees who acted as a resource and concurrently as a “Monitoring resource” (Municipality-5 – 560600, where in 95% of his activities were acting in the resource role and a “Monitoring resource”).

Since searching for roles according to the above mentioned attributes brought no precise results, we have attempted to identify roles by combination of two different attributes, namely “Responsible actors” and “monitoringResource”. We chose these two attributes because their values contained ID numbers of employees and their title evoked some connection between employees. Hence, we have prepared an algorithm that processed all values in attributes “resource”, “monitoringResource” and “Responsible_actor”. Afterwards this algorithm, we have assigned to each “employee number” the quantity of occurrence within individual attributes. Roles were afterwards assigned to employees, depending on in which attribute or in what combination of attributes the employee has found himself. By taking this step we have identified the following roles: Role 1 (responsible actor, monitoring resource, resource), Role 2 (responsible actor, monitoring resource), Role 3 (monitoring resource), Role 4 (responsible actor), Role 5 (monitoring resource, resource), Role 6 (resource), Role 7 (responsible actor, resource).

Fig.1 indicates number of employees at different roles in all municipalities and a percentile proportion to the total number of employees in all the municipalities.

Interestingly, we have come to the discovery that in Municipality-3, three employees were found whom acted during the process only in the position of Role 4 while in no other municipality was this role identified. The largest number of employees who acted in Role 5 and Role 6 were noticed in Municipality-5, which leads to the conclusion that a group of employees who are responsible for building permits is defined in a more restricted way in this municipality. The exact reversed situation was registered in Municipality-2. Another interesting find was, that Municipality-2 had the highest number of employees who acted in Role 1. This may be caused by the fact that this municipality has the lowest number of employees in comparison to the rest of municipalities. We have observed that in Municipality-2, employees 560532 and 560530 act in Role 1. These employees simultaneously were also in Municipality-5 (employee 560532 since 12.5.2014, employee 560530 since 14.5.2014) and both of them were subsequently assigned to Role 5 in Municipality-5.

	M1 (c)	M1 (%)	M2 (c)	M2 (%)	M3 (c)	M3 (%)	M4 (c)	M4 (%)	M5 (c)	M5 (%)
Role 1	16	**55%**	7	**64%**	10	**40%**	7	**50%**	8	**35%**
Role 2	4	**14%**	0	**0%**	7	**28%**	2	**14%**	0	**0%**
Role 3	2	**7%**	0	**0%**	1	**4%**	2	**14%**	1	**5%**
Role 4	0	**0%**	0	**0%**	3	**12%**	0	**0%**	0	**0%**
Role 5	4	**14%**	2	**18%**	4	**16%**	1	**7%**	7	**30%**
Role 6	3	**10%**	2	**18%**	0	**0%**	2	**14%**	7	**30%**
Role 7	0	**0%**	0	**0%**	0	**0%**	0	**0%**	0	**0%**

Fig. 1. Overview of all employees' roles among all municipalities

By this analysis we were able to identify 7 roles but it didn't provide us with sufficient information to determine which type of activity of resources were performed while processing building permits in the identified roles.

To define roles we have ultimately chose event level attribute "action_code". Primarily we have to define an appropriate method suitable to represent roles, given that this attribute assumes 500 unique values throughout all municipalities (in case of any empty values we have used the following titles of attribute "taskNameEN"). If we take a closer look at resource 560872 in Municipality-1, we can see that this resources has operated with 230 unique values of the above mentioned attribute. To simplify representation of these roles by attribute "action_code", we have created a system which aggregated processes and their phases in different municipalities on the basis of the description of logs. To explain our representation of a process and its phases by values of this attribute as follows: we have taken the first two digits and a variable number of characters, which represented process; and last three digits which indicate an order, where we took first digit of a trinity as a representative of that phase. For examples out of codes "01_BB_540, 01_BB_545, 01_BB_546, 01_BB_550, 01_BB_550_0, 01_BB_550_1, 01_BB_550_2, 01_BB_550_3, 01_BB_550_3a, 01_BB_560, 01_BB_590" which indicate "objections and complaints" subprocess, a new aggregate code "01_BB_5" has arisen. By using this procedure we have created three new aggregated codes "01_BB_5, 01_BB_6, 01_BB_7" out of subprocess "objections and complaints". Analogically, we have processed all codes and thus we have obtained overall 45 new aggregated codes. Table 2 contains a summary overview of aggregated codes throughout all municipalities.

Consequently we have created an algorithm, which has calculated an absolute frequency of performed aggregated codes for each resource in different municipalities. By this step, we have obtained an overview of activities of resources who participate in the building permits process.

According to the above described process we have prepared an analysis of the workload of resources for the aggregated codes for each municipality. For the purposes of analysis we have chosen only those resources, whose relative frequency of performed activities was higher than 1%. Resources with a relative frequency lower than 1% were not included into the present analysis given that some of them have engaged into the process only in the beginning of 2015. An example seen was, (e.g.

resource 12941730 in Municipality-1), where the activity of some of those resources were registered only at the beginning of the observed period (e.g. resource 560528 in Municipality-2; the last activity of this resource was registered on 7.12.2010), or their activity was registered only in a small number of cases (e.g. resource 560713 in Municipality-3; the activity was registered only in one permit).

Table 2. Summary count of aggregated codes throughout municipalities.

	Aggregate codes	"action_code" (count)
Municipality-1	37	398
Municipality-2	43	409
Municipality-3	40	383
Municipality-4	42	355
Municipality-5	40	386

4.2 Roles of resources involved in various stages of the process

We have identified roles by names of aggregated codes, because the titles of codes captured in the attributes (English and Dutch) didn't have sufficient informative capability and we leave their full denomination to the process owner.

Out of all municipalities we have chosen such resources whose relative frequency is higher than 1%. Subsequently we have calculated a relative frequency of all performed aggregated codes for each of the resources. For example, we have calculated relative frequency of aggregated code 01_HOOFD_0 for resource 560872 whose absolute frequency captured in log is 12 117 as follows: we have summed up frequency of all codes which formed aggregated code performed by this resource and divided its' captured frequency[1]. Figure 2 indicates the top 7 aggregated codes which employee 560872's activity were divided in. By this calculation we have obtained an overview of the portions of activities of one resource which is devoted to concrete aggregated codes, therefore seeing what phases of individual processes is his activity divided in.

	01_HOOFD_0	01_HOOFD_1	01_HOOFD_4	01_HOOFD_5	01_HOOFD_8	02_DRZ_0	04_BPT_0
% Activity in Aggregated Codes	45.03 %	11.03%	7.54%	6.02%	3.75%	4.09%	5.83%

Fig. 2. Top 7 aggregates codes for Employee 560872 activity roles in percentages.

By the above described process, we have identified the distribution of activities of the resources in all municipalities. On the basis of the obtained outcomes we were able to define the aggregated codes among which activity of each resource is divided. Subsequently, we have identified different groups of resources according to the portion of activities among aggregated codes for each municipality. Figure 3 indicates the

[1] Calculated relative frequency of aggregated code 01_HOOFD_0 for resource 560872: 01_HOOFD_010 (f=690) + 01_HOOFD_011 (f=476) + … + 01_HOOFD_099 (f=16) = 5 456 / 12 117 = 0.45027 = 45.03 %.

distribution of activities of resources among most numerous aggregated codes and its' color allocation for Municipality-4.

Resource	TotalCount	01_BB_5	01_BB_7	01_HOOFD_0	01_HOOFD_1	01_HOOFD_2	01_HOOFD_3	01_HOOFD_4	01_HOOFD_5	01_HOOFD_8	08_AWB45_0	10_UOV_0
560781	15734	0.01	0.09	29.04	3.67	2.23	2.83	18.71	23.66	0.48	2.75	0.91
560752	11893	3.5	4.1	27.18	5.22	2.22	1.97	13.6	14.02	4.47	2.07	2.82
560821	3340	0.18	0.27	25.47	6.5	3.35	5.47	17.9	13.86	1.36	1.8	4.71
560852	8263	0	0.01	2.29	15.32	14.21	16.97	22.22	0.39	0.16	7.76	0.52
1550894	6397	0.5	0.66	4.72	15.79	12.01	13.3	22.09	0.8	0.75	10.71	1.59
560849	763	0	0.39	3.67	27.39	15.2	10.09	8.39	1.18	1.97	11.01	0
560812	721	2.22	1.66	3.74	15.53	13.31	12.76	20.8	13.59	1.53	0	0

Fig. 3. Color coded employee allocation of their activities within resources.

By comparison of portions in aggregated codes, we were able to identify roles for each municipality separately; and subsequently we have performed another comparison of identified roles between individual municipalities. As a next step, we have formed a list of 13 roles, which were identified by means of this analysis. Figure 4 indicates a list of roles and their assigned number of resources and a list of aggregated codes which define the given role.

The most significant role is Role 0, which was identified in all municipalities. This role represents resources where approximately 70% of their activity is divided among aggregated codes 01_HOOFD_0, 01_HOOFD_4 and 01_HOOFD_5. This portion is higher in municipalities with more resources engaged in the process, for example 82.6 % of activity for resource 560532 from Municipality-5 is divided between these aggregated codes. In municipalities with less resources engaged in processes where portions were usually lower, e.g. only 57.23% of activity of resource 560821 from Municipality-4 is divided among the above mentioned aggregated codes. This difference may be due to the fact that in municipalities with higher number of resources exists a possibility of appointing a more specific scope of activities to one resource, compared to the municipalities with lower number of resources.

	Num. of re	01_HOOFD_0	01_HOOFD_1	01_HOOFD_2	01_HOOFD_3	01_HOOFD_4	01_HOOFD_5	01_HOOFD_7	01_HOOFD_8	02_DRZ_0	04_BPT_0	08_AWB45_0	01_BB_5	01_BB_7	10_UOV_0
Role 0	13	✓				✓	✓								
Role 1	9	✓	✓	✓	✓	✓						✓			
Role 2	8		✓	✓	✓	✓						✓			
Role 3	5	✓													
Role 4	2	✓	✓	✓	✓	✓	✓					✓			
Role 5	2	✓	✓			✓					✓	✓			✓
Role 6	1					✓	✓								
Role 7	1		✓	✓	✓	✓	✓		✓			✓			
Role 8	1					✓	✓					✓			✓
Role 9	1	✓	✓	✓	✓	✓				✓	✓	✓			
Role 10	1	✓	✓	✓	✓	✓	✓		✓						
Role 11	1		✓	✓	✓	✓	✓								
Role 12	1	✓				✓	✓	✓	✓				✓	✓	

Fig. 4. Summary of identified roles with their aggregated codes.

Role 1 and Role 2 were identified in three municipalities. Role 3 and Role 5 were identified only in two municipalities. Other roles are particular for individual municipalities. We can notice divergence in the procedure of processing the building permits in cases of roles, which were identified only in one of the municipalities. For example, for Role 12 in Municipality-5, a typical 30% of activity of resource 560613 is divided between aggregated codes 01_BB_5 (15%) and 01_BB_7 (14%). Another example may be of Role 8 in Municipality-3, where resource 5025869 represents their activity including also the performance of aggregated code 10_UOV_0 (12%). Analogically, as seen cases for Role 8 and also for Role 5, both encountered performing the aggregated activity 10_UOV_0. These two roles are performing the same aggregated activity, which can be seen with a specific example of resource 560462 from Municipality-1. In addition, this resource also performed aggregated codes 0_HOOFD_0 (10%) and 0_HOOFD_1 (16%). Figure 5 indicates list of roles and list of municipalities in which the roles were identified.

	M1	M2	M3	M4	M5
Role 0	✓	✓	✓	✓	✓
Role 1	✓		✓		✓
Role 2	✓	✓		✓	
Role 3	✓				✓
Role 4	✓				
Role 5	✓		✓		
Role 6	✓				
Role 7		✓			
Role 8			✓		
Role 9			✓		
Role 10		✓			
Role 11				✓	
Role 12					✓

Fig. 5. List of roles seen in each municipality.

In Municipality-4 we identified the lowest number of roles from the point of view of aggregated codes. This may indicate that the resources are not closely specified to concrete actions in the course of processing the building permits. This may influence the time of processing the building permit, while a particular objection or complaint may arise and should be resolved by a resource who has experience and is specified in this particular area to do so. In this municipality we can notice that all resources which have been included in the analysis are involved in the processing of codes, that make part of aggregated codes 01_BB_5, 01_BB_6 and 01_BB_7 (these aggregated codes represent the subprocesses “objections and complaints”). As a result, if a specific objection or complaint occurs, the resource is obliged to study the area to which the objection or complaint relates, in order to find a solution and this may take more time. In this case, a narrower specialization of the resources for different subprocesses could bring some improvement.

2

5 Organizational structure

In this section, we looked at ways in how to enhance organizational structure in every municipality. To visualize the organizational structure, we selected an organization chart. This section also includes a short description of the algorithm for creating the organizational chart and the resulting comparison of municipalities’ organizational charts. Improvements proposed in this section are only recommendations, because we can only demonstrate what we observed in the data. However, without knowledge of the methodology of application building permits, we will not be able to identify the behavior clearly and assess how good or bad it is. Overall assessment of facts is the process which should be done by the process owner.

5.1 Creating organizational structure as an organizational chart

Firstly, it was necessary in respect to create an organizational structure and organizational chart to determine hierarchical relation, which was determined by the direction of responsibility flow. In the recorded data we observed sort of hierarchy, where the entire case was responsible for “Responsible_actor”. “Responsible_actor” is also an attribute on the case level, and stays that way for the entire case as a fixed entity. Subsequently, was observed an event level attribute “monitoringResource”, which is responsible for part of a case. The lowest hierarchical unit was designated as a resource responsible for conducting the event. The hierarchy is illustrated in Figure 6. Our main goal was to show an organizational chart, which would describe the hierarchical status of each person involved and also the relationship to other persons in the hierarchy in the process. Relationships within the hierarchy describes who is superior as a Responsible or Monitoring and who is a subordinate as Monitoring or Resource.

In order to create organizational structure and the organizational chart, an algorithm was developed so that inputs were identified in the above paragraph. For the operation of this algorithm, it is essential to have identified a hierarchy of attributes in the log (to have this hierarchy situated in the log).

The number of hierarchical levels, which enters the algorithm is $\boldsymbol{N}$. The number of hierarchical relations in which the algorithm creates, is one less of the number of hierarchical levels. We define the relation between levels always downwards to a lower level in the hierarchy. Relations between level 3 and level 2 are identified as $\rightarrow_{\mathbf{n-1}}$, where $\boldsymbol{n}$ is specified always by the higher level of the hierarchy. Relation can only be created between hierarchical levels $\boldsymbol{i_n}$ and $\boldsymbol{i_{n-1}}$.

As mentioned above, the number of hierarchical levels are three and the number hierarchical relations are two. Designed algorithm consists of four steps, which are described below.

In the first step, an algorithm is passing through all the cases and their respective events and it identifies employees and creates a relations. For employees included in the process are noted a certain hierarchical level and the absolute event frequency of occurrence at that level. The relations in which algorithm creates are “Responsible$\rightarrow_{\mathbf{2}}$Monitoring” and “Monitoring$\rightarrow_{\mathbf{1}}$Resource” and it counts their absolute frequencies on an event level. Created relations $\rightarrow_{\mathbf{2}}$ and $\rightarrow_{\mathbf{1}}$ are completely independent from each other. The result of the first step is the registration of information for every employee, which is the occurrence and frequency of occurrence in one of the hierarchical levels and clearly identified above are explained relations and their quantity.

In the second step, unique hierarchical relations are created for each employee whereby passing through the identified relations and especially for relations “Responsible$\rightarrow_{\mathbf{2}}$Monitoring” and also particularly for the relations “Monitoring$\rightarrow_{\mathbf{1}}$Resource”. Unique relations are created only in cases of unique relationships between employees so that their relation “Responsible$_{\mathbf{A}}\rightarrow_{\mathbf{2}}$Monitoring$_{\mathbf{B}}$”

is between two employees created only in a case if there exist casual dependency $\mathbf{A}\rightarrow_{2}\mathbf{B}$ and at the same time, dependency $\mathbf{B}\rightarrow_{2}\mathbf{A}$ does not exist. The same applies to "$\text{Monitoring}_{\mathbf{B}}\rightarrow_{1}\text{Resource}_{\mathbf{C}}$", where a relation is created only if the existing casual dependency $\mathbf{B}\rightarrow_{1}\mathbf{C}$ exists and at the same time $\mathbf{C}\rightarrow_{1}\mathbf{B}$ dependency does not exist. Discovered in the analyses were facts that in some cases an employee was acting in all hierarchical levels. Because of this found fact, identifying the employees more precisely created relations $\mathbf{A}\rightarrow_{2}\mathbf{A}$ and relations $\mathbf{A}\rightarrow_{1}\mathbf{A}$, whose multiplicity determines the activity of employee in processes within hierarchical levels in more detail.

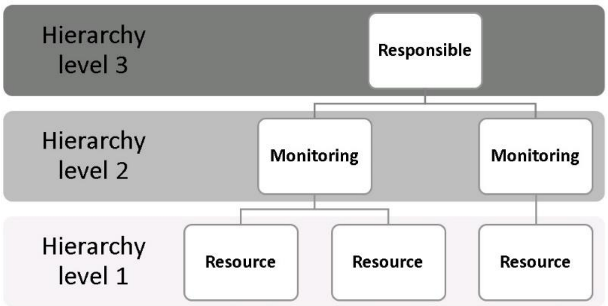

Fig. 6. Illustration of the hierarchy for process instances.

In the third step, for every employee there was a calculated absolute frequency of occurrence in hierarchical levels. For the relations identified, an absolute and average values were calculated. The average frequency have not been calculated nor values of the relations $\mathbf{A}\rightarrow_{2}\mathbf{A}$ and $\mathbf{A}\rightarrow_{1}\mathbf{A}$, because their job is above mentioned in their identification of activities. These calculations work mainly for the visualization properties, which describes an employee and his relations, which are created in a fourth step.

The last fourth step creates the visualization of organizational chart. Employees are represented in a chart as circles and absolute multiplicity of performed events of every employee is visualized by the size of the circle. Visualized for an employee is an organizational chart whereby their specific role is described in section 4.1 of this work and a specific role is represented by coloring. Colors are connected to roles in the following manner:

- Role 1: responsible actor, monitoring resource, resource
  -
- Role 2: responsible actor, monitoring resource
  - Gradient coloring and
- Role 3: monitoring resource
  -
- Role 4: responsible actor

  –
- Role 5: monitoring resource, resource
  - Gradient coloring and
- Role 6: resource
  –
- Role 7: responsible actor, resource
  - This role is not present in any municipality.

Relation is represented by oriented edge and an absolute event count of a particular relation is visualized by the thickness of the edge. Relation $\rightarrow_2$ in an organizational chart that is represented by a solid edge and relation $\rightarrow_1$ is represented by a dotted edge.

The algorithm was implemented in C# and data were read from SQL databases. To visualize the organizational chart, an open source graph visualization software was used, Graphviz. Please note, description of the actual implementation of the algorithm is not the purpose of this part of the work.

The resulting organizational chart of Municipality-1 is shown in Figure 7. Count of employees is 30, from which one employee is identified without a label and who performed in 19 events as Responsible. Among the employees is the count of $\rightarrow_2$ relations 93 and count of $\rightarrow_1$ relations is 88, which the total count of relations is not included in a count of $\mathbf{A}\rightarrow_2\mathbf{A}$ and $\mathbf{A}\rightarrow_1\mathbf{A}$ relations.

The municipality chart is very complex because some employees are featured in a process in 2 or 3 roles, which causes them to have many relations in both levels with low event frequencies. This feature was observed across all organizational charts for every municipality. Based on this property, we decided to reduce the complexity of the organizational chart. To reduce complexity, we used the value calculated in the third step of the algorithm, the average value of the frequency of relations for each level was calculated separately.The result displays only those relations at both levels, which are equal to or higher than the average value. Between the employees is a count of $\rightarrow_2$ relations 42 and a count of $\rightarrow_1$ relations is 32. The total count of relations were not included in a count of $\mathbf{A}\rightarrow_2\mathbf{A}$ and $\mathbf{A}\rightarrow_1\mathbf{A}$ relations. The count of relations between employees with low frequency was still high and for better visualization and legibility of the chart we have filtered the relations with less frequency than a value of 50 at both levels. The value of 50 has been calculated by a median value of other median values from event counts in cases for each municipality. Our intention was to visualize the most frequent relations at both hierarchical levels. The resulting chart is visualized on Figure 8. Chart includes 17 $\rightarrow_2$ relations and 12 $\rightarrow_1$ relations, and a total count of relations were not included count of $\mathbf{A}\rightarrow_2\mathbf{A}$ and $\mathbf{A}\rightarrow_1\mathbf{A}$ relations. Eleven unconnected employees can be observed in the chart. The reason being, is their generally low event frequency between 9 and 215, as well as the total event frequency for both types of relations is less than the 50.

Organizational charts for every municipality were produced thereby showing the resulting data in charts (see Appendix Figures 15-19). A comparison of the characteristics of organizational charts is contained in Table 3, where the total count of relations are calculated without **A→$_2$A** and **A→$_1$A** relations.

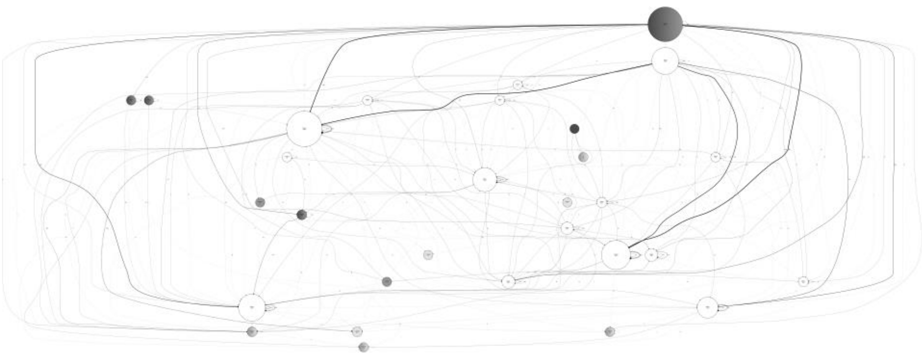

Fig. 7. Municipality-1 organizational chart including all relations.

In the organizational charts of Municipality-2 and Municipality-5 one will see similarities in the roles as described previously section 4.1, where employees fit into roles 1, 5 and 6, with the exception of employee 560583 with 1 event count in role 3. An organizational chart of the two municipalities is also decomposed only to 3 horizontal lines compared with other municipalities.

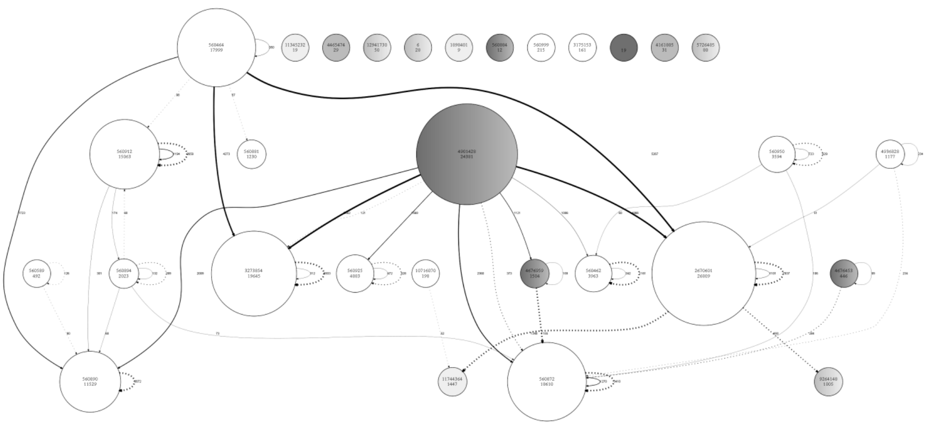

Fig. 8. Organizational chart for Municipaliyt-1 with limitation on average relation event frequency and filtered relations with less frequency than a value of 50 at both levels.

In all reduced organizational charts of municipalities, we have observed at least one employee, who performed with high frequency in relations **A→$_2$A** and **A→$_1$A**. An employee essentially operates within the organizational structure mostly alone. This really impressed us so much that, we have calculated data for each municipality and each employee and what percentage of their work as a resource is done as a monitoring resource alone, or a responsible actor, or a monitoring and responsible actor simultaneously. We observed how many performed events an employee is responsible for as a resource, to which we saw that no one was responsible for his/her work, but themselves. Again, we have focused on employees whose relative frequency in the role of resource is more than 1%. Displaying the top 5 employees according to event frequency is found in Figure 9.

Table 3. Compared characteristics of relations between original and reduced datasets.

	Resources	$\rightarrow_2$ count		$\rightarrow_1$ count	
		Original	**Reduced**	**Original**	**Reduced**
M1	30	93	17	88	12
M2	11	19	6	18	6
M3	25	62	15	49	10
M4	14	26	7	25	8
M5	23	51	4	61	18

Figure 9 enables us to notice differences between Municipality-1 compared to other municipalities. Interestingly enough, notable differences are seen as well in regards to employees in compared municipalities. Employee 560912 in Municipality-1 has responsibilities up to 28% as responsible actor for himself, compared to Employee 560852 in Municipality-4 where they had a 99% responsibility as a responsible actor for himself.

The main difference in such a wide range in percentiles in Municipality-1 compared to other municipalities is the fact that Municipality-1 includes employee 4901428 in their organizational chart, whom is exclusive in a role as a responsible actor and monitoring resource. In this diversity we see possible points for improvement on the organizational structure for municipalities 2, 3, 4, and 5.

For Municipality-5, once can see that there is room for improvement and capability to dedicate people for the last hierarchy level in the organizational chart. The reason being, in their organizational structure there are only 2 hierarchy levels where there are missing employees noted as having only a resource role.

Subsequently, we examined the allocation of performed events between employees working as resource roles. We were interested in the count of resources, who handled over 50% of the total event counts. Table 4 displays a total count of resources, the count of resources which show for the sum of the relative frequency being more than 50% and the relative frequency of their share of the performed events.

Interestingly, in Municipality-4, two resources performed at 58.56% of the total countable events in the process, compared to Municipality-5, where 3 resources performed at 53.72% of the total countable events. Another interesting fact is that in Municipality-2, three resources cover almost 70% of all total events in the process. What grew of interest was the question, “What is the load on the three resources and how did they communicate together in implementing building permits?” Figure 10 shows a social network for Municipality-2 with a displayed metric case count, where resources only with higher relative event frequencies than 1% are shown. In Figure 10, it is possible to quickly identify these 3 resources through the most expressive “halo/echo effects”.

M1

Name	EventCount	MR%	RA%	MR-RA%
560872	12117	28,14%	10,27%	9,66%
2670601	9886	79,27%	3,69%	3,16%
3273854	9075	76,40%	1,73%	1,66%
560890	7399	49,63%	0,00%	0,00%
560912	5346	49,74%	28,23%	14,74%

M2

Name	EventCount	MR%	RA%	MR-RA%
560530	11479	26,74%	26,59%	24,53%
560532	10080	2,68%	0,19%	0,19%
560458	9082	58,67%	56,51%	52,33%
560519	7821	83,52%	79,76%	78,40%
560521	3459	93,18%	90,23%	89,82%

M3

Name	EventCount	MR%	RA%	MR-RA%
560454	14620	78,11%	78,78%	68,65%
560673	10457	18,57%	0,00%	0,00%
2013365	8819	78,56%	85,44%	71,40%
560749	8763	12,90%	0,00%	0,00%
560741	7429	10,76%	0,00%	0,00%

M4

Name	EventCount	MR%	RA%	MR-RA%
560781	15748	2,13%	0,00%	0,00%
560752	11948	3,39%	0,49%	0,45%
560852	8264	98,94%	99,43%	98,94%
1550894	6452	94,34%	95,35%	93,04%
560821	3344	2,90%	0,00%	0,00%

M5

Name	EventCount	MR%	RA%	MR-RA%
560604	12225	99,51%	94,24%	93,91%
560602	10505	2,79%	0,14%	0,00%
560600	9008	94,98%	95,63%	94,13%
560429	7590	65,59%	61,95%	59,75%
1254625	6234	4,57%	0,00%	0,00%

Fig. 9. Responsibility for Top 5 resource for every municipality.

The social network aspect can indeed show that these 3 resources are overloaded. Employee 560532 has worked an entire 64% of all building permits. Another interesting fact is that these 3 resources were first in 91% of all building permits. High values were reached as well as, their case frequencies when they worked together and cooperated as a team on all building permits. Specifically, resource 560532 had dedicated/pushed their work to resource 560458 in 25% of all building permits. Another distinctive communication is where resource 560530 had dedicated their work to resource 560458 in 20% of all building permits.

Table 4. Relative event coverage for resources which sum of the relative frequency being more than 50%.

	Resources count	Count of top resources	Relative event coverage
M1	23	3	59,52 %
M2	11	3	69,09 %
M3	14	3	56,8 %
M4	10	2	58,56 %
M5	22	3	53,72 %

Similar overloaded resources were seen in every municipality. In Appendix Figures 20-24, social networks of all municipalities' shows resources with their relative event frequency higher than 1% with a displayed case count metric. In Municipality-1 resource 560872 worked a total of 63% of all building permits. Resource 560872 dedicated their work to resource 2670601 in 18% of all building permits. In Municipality-3, resource 560673 worked on 56% of all building permits and in 16% process instances dedicated their work to resource 560454, which has on another hand, the highest event frequency. Resource 560752 in Municipality-4 worked a total of 88% from all building permits and resource 560781 worked a total of 77% of all building permits. Resource 560781 dedicated their work to resource 560752 in 50% of all building permits. In Municipality-5, resource 560602 worked on 62% of all building permits.

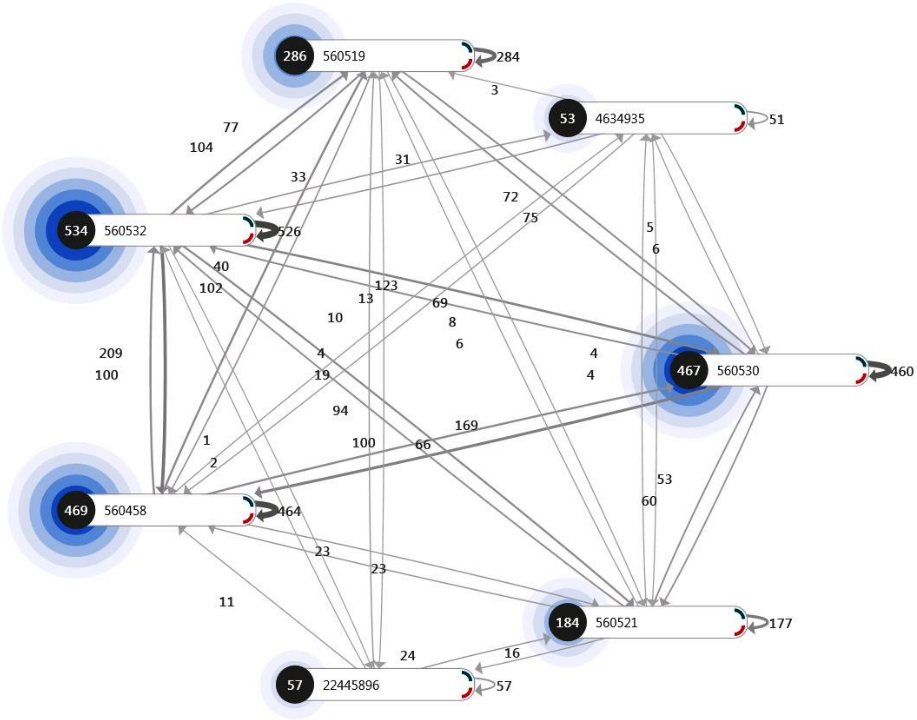

Fig. 10. Social network of Municipality-2, resources with relative event frequency greater than 1 % are displayed.

On the basis of the above described facts, we can consider the next possible point for improvement in the organizational structure in dedicating equally shared work between resources.

6 Differences in throughput times between the municipalities

In this section of the paper we will describe how we have searched for answers to the question, what are the differences in throughput times between the municipalities, what kind of analysis we used and we will describe the observed differences. The biggest obstacle in the initial analysis was that the timestamps captured in the log, were mostly the same for different groups of activities. For example, in case 13274281 we captured in log BPIC15_5 that, for 15 events, available were only three different timestamps. This fact does not interfere when comparing the durations of individual municipalities, but it does not let us describe in detail the differences between the implementation of individual activities. So, we had to create custom methods as procedures for obtaining the differences in throughput times in the logs, which are extracted from a system which lacks detailed logging occurrence of the activity and their duration.

The first step to obtain a basic overview of differences in throughput times was to compare the median and mean time durations between municipalities. Table 5 shows a basic comparison of the throughput times between municipalities.

With this step, we identified that Municipality-3 explains was the fastest in processing the building permits while Municipality-2 is the slowest in processing all building permits. We could already deduce from this perspective, that Municipality-2 and Municipality-4 processes building permits the longest because they have the lowest count of resources. This observation excludes the fact that Municipality-3 is not the one with the most amount of resources (most resources are in Municipality-1 with 23 resources while Municipality-3 has 14 resources). At the same time, it was brought up to attention that the number of resources with time in each municipality changed. To be able to answer the question, "Why there are such visible differences in throughput times between the municipalities", we started searching for answers using attributes.

Table 5. Comparison of Mean and Median Case Durations for each Municipality.

	Median case duration	Mean case duration
M1	61.37 days	95.72 days
M2	108.5 days	160.1 days
M3	38.46 days	62.23 days
M4	96.41 days	116.8 days
M5	77.17 days	98.34 days

6.1 Different findings with the help of attributes

The first attribute that we chose to identify differences was "(case) last_phases". According to this attribute we tried to group cases in the individual municipalities and subsequently compare between them case durations. Followed by detection of this attribute we found out that it is not convenient to make a comparison because most numerous values of this attribute between different municipalities vary. For example, the most frequent value of attributes in Municipality-4 is "Besluit onherroepelijk" which were included in 442 cases compared to Municipality-2, which was only seen in one case. This difference may be explained by the fact that each of the municipalities works independently from other municipal systems, and thus it uses other values of this attribute.

To make possible for us to compare cases that share common characteristics we chose attribute "(case) parts". For a comparison, we have chosen the most frequent values "Bouw", "Kap", "Sloop" and "Milieu (vergunning)". Table 6 shows a comparison of case durations according to the attributes "(case) parts" between municipalities.

Table 6. Case Duration comparisons according to specified attributes.

	M1(med/avg)	M2(med/avg)	M3(med/avg)	M4(med/avg)	M5(med/avg)
Bouw	56.6 d / 75.4 d	107.8 d / 141.4 d	45 d / 63.8d	95.2 d / 117 d	65.5 d / 81.3 d
Kap	82.5 d / 81.5 d	53.5 d / 59.6 d	19.4 d / 25.2 d	88.5 d / 90.5 d	67.4 d / 95.9 d
Sloop	37.0 d / 50.4 d	116.6 d / 128.8 d	19 d / 27.4 d	72.3 d / 82.5 d	59.6 d / 67.4 d
Milieu(v.)	199.5 d / 278.6 d	270.2 d / 295.4 d	195.3 d / 211.4 d	193.2 d / 198.8 d	236.6 d / 265.3 d

We observed that Municipality-2 in these cases has the second fastest duration after Municipality-3 with the value attribute “Kap”. This means that the municipality is not the slowest in general terms, but certain types of cases create dismay for them. This observation was recorded also for other attribute values and so we decided to look at the differences in throughput times between the municipalities from another perspective.

6.2 Collocation of throughput times to time intervals

When comparing we have found out that in some cases there is a higher difference between average case durations and median case durations. Thus, we decided to verify if the individual municipalities noticed extremely long cases (eg. as in Municipality-3, where the duration of 4 years and 53 days in case 3198296), which could distort the average case duration. For the realization of this comparison, we have prepared an algorithm that divides cases in individual municipalities by duration into groups that represent different time intervals with a length of one week. In total, 217 intervals were created for a one week representation in order to cover all different case durations. Then data for individual municipalities were combined into a single report. Each time interval, which can be seen in the aggregate contain the following information: identifier of time interval, the count of building permits for all municipalities, the count of building permits for individual municipalities, the relative share of building permits for a single municipality, winner (municipality, which had highest relative proportion in a time interval), looser (municipality, which had lowest relative share in a time interval), top 5 Last Phases (5 of the most numerous attribute values “(case) last phases)”, top 5 Parts (5 of the most numerous attribute values “(case) parts”), top 5 Other (5 of the most numerous values for attributes “(case) Includes_subCases”, “(case) caseStatus”, and “(case) caseProcedure”). Figure 11 is overview of time intervals in which the count of building permits for any municipality is greater than 100 (for clarity, this does not contain information about the top 5 values of selected attributes).

Time Interval	PermitsCount	M1C	M2C	M3C	M4C	M5C	M1R	M2R	M3R	M4R	M5R	Winer	Loser
7(49) ~ 8(56)	414	163	28	108	27	88	13.6	3.4	7.7	2.6	7.6	1	4
8(56) ~ 9(63)	316	91	20	74	33	98	7.6	2.4	5.3	3.1	8.5	5	2
2(14) ~ 3(21)	307	39	14	165	33	56	3.3	1.7	11.7	3.1	4.8	3	2
0(0) ~ 1(7)	305	59	37	107	65	37	4.9	4.4	7.6	6.2	3.2	3	5
3(21) ~ 4(28)	293	50	17	144	39	43	4.2	2	10.2	3.7	3.7	3	2
6(42) ~ 7(49)	277	78	28	106	21	44	6.5	3.4	7.5	2	3.8	3	4
9(63) ~ 10(70)	274	85	22	63	44	60	7.1	2.6	4.5	4.2	5.2	1	2
13(91) ~ 14(98)	271	58	46	38	58	71	4.8	5.5	2.7	5.5	6.1	5	3
11(77) ~ 12(84)	241	58	36	36	50	61	4.8	4.3	2.6	4.7	5.3	5	3
1(7) ~ 2(14)	238	31	18	105	37	47	2.6	2.2	7.5	3.5	4.1	3	2
4(28) ~ 5(35)	228	48	15	123	13	29	4	1.8	8.7	1.2	2.5	3	4
5(35) ~ 6(42)	223	58	16	100	23	26	4.8	1.9	7.1	2.2	2.2	3	2
14(98) ~ 15(105)	218	41	52	12	36	77	3.4	6.2	0.9	3.4	6.7	5	3
10(70) ~ 11(77)	213	51	22	35	57	48	4.3	2.6	2.5	5.4	4.2	4	3
12(84) ~ 13(91)	195	32	32	33	36	62	2.7	3.8	2.3	3.4	5.4	5	3
15(105) ~ 16(112)	153	28	27	16	39	43	2.3	3.2	1.1	3.7	3.7	4	3
16(112) ~ 17(119)	116	12	21	8	34	41	1	2.5	0.6	3.2	3.5	5	3

Fig. 11. Overview of time intervals for each municipality regarding count of building permits.

The aggregated time intervals we subsequently put together in an overview are a comparison for the durations of building permits for singular municipalities. Therefore, this report was set up in order to identify the time interval between which municipality decomposes the largest share of building permits. A summary comparison of building permits of time intervals is seen Table 7.

Table 7. Comparison summary of each municipality regarding time intervals in building permits.

Time interval (weeks)	M1 (%)	M2 (%)	M3 (%)	M4 (%)	M5 (%)
0-1	4.9	4.4	7.6	6.2	3.2
1-2	2.6	2.2	7.5	3.5	4.1
2-3	3.3	1.7	11.7	3.1	4.8
3-4	4.2	2	10.2	3.7	3.7
4-5	4	1.8	8.7	1.2	2.5
5-6	4.8	1.9	7.1	2.2	2.2
6-7	6.5	3.4	7.5	2	3.8
7-8	13.6	3.5	7.7	2.6	7.6
8-9	7.6	2.5	5.3	3.1	8.5
9-10	7.1	2.6	4.5	4.2	5.2
10-12	9.1	6.9	5.1	10.1	9.5
12-14	7.5	9.4	5	8.9	11.5
14-16	5.7	9.5	2	7.1	10.4
16-20	4.2	9.4	2	12.6	7.7
20-24	2.6	9.3	1.9	10	3.8
24-28	2.4	5	1.2	6.5	2.2
28-32	1.3	4.8	1	4.2	1.9
32-40	2.6	6.5	1.2	4.2	3.1
40-50	2.1	5.7	0.7	1.7	1.6
50-100	3.2	5.9	1.8	2.2	2.2
100-200	0.7	1.6	0.2	0.7	0.5
200-217	0.1	0	0.1	0	0

We observed that Municipality-3 has the largest share of cases, in which 60.3% were closed within 7 weeks. Municipality-1 had 30.3% shares, Municipality-5 had 24.3%, Municipality-4 had 21.9%, and the lowest amount of share cases were in Municipality-2 at 17.4%. Municipality-2 had the largest share of cases between 7 to 30 weeks at 60.5%. To compare, Municipality-3 contains 35.2% cases with a duration of 7-30 weeks. Interestingly enough, Municipality-4 has 67.1% cases during this time frame. From this observation, we can conclude that Municipality-2 has the largest proportion of cases with more than 30 weeks in duration, seeing a result of 21% whereas other municipalities came up with, Municipality-1 at 10.5%, Municipality-3 at 4.7%, Municipality-4 at 11.2%, and Municipality-5 at 8.4%).

6.3 Searching differences according to the number of steps

The best method to compare differences in throughput times between municipalities would be to compare the duration of cases between municipalities, which represent the same behavior. If we would choose from each municipality the cases in which the same sequence of events occurs, and then subsequently compare them with each other, we would find major differences in throughput time between them. However, unfortunately it is not possible to do so in this case, because we found only one common modified variant for all municipalities (described later in section 7.1). Therefore, we had to find another way to compare throughput times and find where the largest differences between municipalities takes place. For further analysis, we therefore chose a comparison of the duration of cases with the same number of events between different municipalities. For this analysis, we have created an algorithm that identifies 117 different numbers of events in cases covering all municipalities. Subsequently, to each value the algorithm computes a total number of cases over all municipalities, as well as a count of cases in individual municipalities, a calculated average duration of all cases with the same value number of events, and lastly, the average length of cases for individual municipalities. By doing so, interesting insight into what differences in throughput times between the municipalities according to the number of steps made in the processing of building permits were discovered. Table 8 shows sample output of the differences in throughput times between the municipalities according to the number of events in cases (for this illustration, selected were only the most numerous cases in municipalities).

Table 8. Differences in throughput times in days among each municipality.

EventCount	TotalPerm.	M1	M2	M3	M4	M5	Avg.Duration	Avg.D.M1	Avg.D.M2	Avg.D.M3	Avg.D.M4	Avg.D.M5	Winer	Loser
44	385	77	50	135	10	113	86.4	60.7	124.3	35.2	127.7	84.0	3	4
42	285	48	7	109	98	23	122.2	54.0	364.1	31.7	54.7	106.4	3	2
46	282	36	21	116	22	87	90.2	70.3	111.9	38.1	152.9	77.8	3	4
40	222	72	15	90	38	7	97.8	55.8	120.4	41.9	95.0	175.8	3	5
45	189	61	27	25	14	62	99.7	64.5	101.7	60.9	171.2	100.1	3	4
41	188	26	17	71	48	26	87.5	70.1	141.1	33.5	95.9	96.7	3	2
39	166	12	6	49	92	7	106.3	41.6	241.7	42.3	117.9	87.9	1	2
54	163	24	15	11	33	80	107.2	108.5	101.9	101.5	122.1	101.9	3	4
43	160	28	30	26	23	53	98.2	73.7	136.4	55.9	130.8	94.4	3	2
38	160	38	6	90	23	3	62.1	45.4	85.9	41.3	60.9	76.8	3	2
53	157	30	27	30	61	9	102.4	107.0	117.2	62.7	127.1	97.8	3	4
52	147	24	20	6	86	11	96.2	88.5	113.4	80.3	118.1	80.7	3	4
51	142	28	16	13	66	19	96.5	95.2	123.6	94.7	114.2	54.9	5	2
50	128	18	9	7	32	62	110.6	81.3	164.2	120.4	136.7	50.5	5	2
61	126	10	54	19	6	37	164.1	188.5	119.7	156.8	246.6	108.8	5	4
55	115	26	34	12	6	37	118.5	127.0	150.1	100.6	112.6	102.0	3	2
47	113	54	13	16	14	16	90.5	65.9	113.0	65.1	134.2	74.2	3	4
48	109	29	20	13	17	30	86.5	69.8	103.5	64.1	154.8	40.3	5	4
59	108	19	45	14	6	24	130.7	92.4	109.5	149.9	178.5	123.2	1	4

For the analysis, we selected only a number of events that we have identified in all five municipalities. We observed that Municipality-3, was the fastest in 38% of all event count cases. The least/slowest were found in Municipality-4 at 28% of event count cases. A close second was Municipality-2, which was 27 %, Municipality-1 at 26%, and Municipality-5 with 8.2%. Municipality-3 was the slowest in seven penetrations compared to other municipalities and in cases with the number of events 75, 64, 62, 31, 16, 13, and 10. We have identified the top 5 counts of events in cases that have been among individual municipalities the longest and over an average among all municipalities. Top 5 event counts are displayed briefly in Table 9.

Table 9. Top 5 Event Counts among the longest duration times, in days.

Events	Avg. Dur.	Avg. M1	Avg. M2	Avg. M3	Avg. M4	Avg. M5
83	432.9	644.2	512.0	365.0	148.4	495.0
78	284.0	472.3	171.4	150.6	271.0	354.4
82	275.6	218.0	196.8	176.9	464.1	322.4
71	272.1	93.4	421.1	279.4	383.4	183.3
77	263.4	458.2	347.5	115.6	249.5	145.9

We subsequently identified a count of events in cases for which individual municipalities have the longest throughput time, but the condition was dependent on having at least 1% of the cases in an event count value. This overview is displayed in Table 10.

Table 10. Event Count Values for each municipality and their longest duration of throughput times in days.

	Event count	Case count	Relative Case count	Avg. Dur.	Avg.M1	Avg.M2	Avg. M3	Avg.M4	Avg.M5
M1	62	17	1.4	164.6	**140.9**	175.3	251.8	204.5	50.3
M2	80	9	1.1	256.5	375.8	**371.5**	88.0	277.7	169.4
M3	61	19	1.3	164.1	188.5	119.7	**156.8**	246.6	108.8
M4	62	15	1.4	164.6	140.9	175.3	251.8	**204.5**	50.3
M5	35	22	1.3	98.3	53.8	129.6	32.3	95.7	**180.3**

With this observation, we noticed that there was at least one municipality in which they reached the same number of events in shorter throughput times. We would recommend that based on observation, the municipality with the shortest throughput times doing the same number of events could explain to other municipalities how they can enhance their average throughput times to be faster.

Based on the above analysis, we have already gained insight into which of the municipalities is the slowest, the intervals at which is a large proportion of cases for individual municipalities, as well as insight into which cases according to the number of steps have the longest throughputs times in individual municipalities. This created a question of, "What steps in processing building permits contribute to increasing the throughput times the most?"

6.4 Differences in time to complete

We noticed that in most cases, events in which we noticed a time jump were made visible in the behavior of a process. For example, in case 21989885 (BPIC15_2) after the first assets "register request submission date" (01_HOOFD_010) follow-up activities "OLO messaging active" (01_HOOFD_011) with a time shift one day, followed by a series of activities with the same timestamps, follow the activity with time shifts. These time jumps between different series of activities, can be monitored in the logs. In the following analysis, we focused on time jumps so that can identify the individual municipalities, by which their activities cause an increase in throughputs times. To answer this question, we have prepared a metric time to complete. This metric describes the time that is needed to complete two successive activities, the assumption is that an activity starts if the preceding activity ends. The logs captured the time offsets of various lengths, and often even after just one second. However, we were focused on activities that the time to complete is longer than one day in order to detect those which caused the largest increase throughputs times in individual municipalities. For the disclosure of these activities we had to prepare an algorithm that goes through cases of individual municipalities and records these time shifts together with the name of these activities, the length of time to complete, and names of resources who carried out these activities. Again, we have chosen as a representative activity attribute "action_code".

In this analysis we have identified a total of 307 codes at which "time to complete" was noted longer than one day. By analysis, we have included only those codes that have been identified in all five municipalities which were 92. Of the 92 codes, the most numerous code was 01_HOOFD_010. For each code we prepared an output that contains the following information: the total number of occurrences between municipalities, the number of occurrences in individual municipalities, the relative shares for each municipality (this figure reflects what percentage of the overall incidence was code recorded with time to compete longer than 1 day), average time to complete for any municipality, average time to complete for the individual municipality, winner (municipality, where the average time to complete was the shortest), looser (municipality, where the average time to complete was the longest), slowest resource (resource at which the longest time to complete was measured), slowest resource in average (resource at which the longest average "time to complete" was measured). To illustrate the output please see Table 11 overview of top 10 codes, which had the highest overall incidence among municipalities (without information on the resource and the slowest resource in average resource for clarity).

We observed that Municipality-3 had 34% of all analyzed codes with the shortest time to complete. This was followed by Municipality-4, which had the lowest time to complete in 27% of all analyzed codes. The shortest completion time was achieved by Municipality-2 and it was 8% of the analyzed codes. Again, the worst position in this observation remains to Municipality-2.

Table 11. Top 10 codes with highest occurrences recorded between each municipality (average duration is in days).

ActionCode	TotalOcc	M1	M2	M3	M4	M5	M1R	M2R	M3R	M4R	M5R	Avg. Dur.	Avg. M1	Avg. M2	Avg. M3	Avg. M4	Avg. M5	winer	loser
01_HOOFD_010	4328	1034	535	1163	624	972	86.2	64.5	82.5	59.3	84.2	6.0	6.8	8.6	4.6	4.0	6.1	4	2
01_HOOFD_030_2	1966	328	306	907	11	414	31.1	38.9	68.9	8.9	38.4	26.3	33.3	33.9	17.1	26.3	21.0	3	2
01_HOOFD_520	1065	133	167	59	285	421	32.3	52.8	18.1	70.2	91.1	9.7	3.3	15.6	4.8	11.7	13.3	1	2
01_HOOFD_530	927	66	228	19	190	424	16.5	74.5	6.4	48.2	92	56.4	45.7	85.3	48.6	65.6	37.0	5	2
01_HOOFD_065_2	900	172	135	73	474	46	26.3	28.9	12.2	50.2	7.3	31.8	26.8	46.3	43.1	18.5	24.4	4	2
08_AWB45_020_2	812	187	112	181	117	215	64.9	62.9	62.4	61.6	59.6	35.6	43.9	37.8	31.4	31.7	33.1	3	1
01_HOOFD_100	717	134	143	162	246	32	34.5	49.1	33.4	69.5	6.8	23.4	19.0	22.0	18.1	30.9	26.8	3	4
13_CRD_010	705	172	86	143	159	145	45.9	29.5	28.5	36.6	34.6	8.9	9.0	10.7	6.3	11.0	7.4	3	4
01_HOOFD_101	670	325	64	97	182	2	38.6	36.8	37.9	58.3	0.6	18.7	32.5	20.1	14.4	21.8	4.7	5	1
01_HOOFD_510_2	635	126	135	37	185	152	14	19.1	3.5	21.6	15.8	26.8	17.8	35.9	21.5	35.7	23.4	1	2

We carry on with putting together compendium codes that have the highest average time to complete for all municipalities. In the report we selected only those codes which the relative share for each municipality is greater than 5%. Therefore, we have chosen this adjustment since in some codes such as, 01_HOOFD_510_4, the average time to complete for all municipality is skewed, especially seen in Municipality-5, where this code occurred longer than 1 day only in one instance. The duration of this code in Municipality-5 was 412 days and the average of all the other municipalities was 48 days. Code 01_HOOFD_510_4 was recorded in Municipality-5 at 494 times and only once had a time to complete longer than 1 day. Therefore, we have to set a limit order to discover codes that have the highest time to complete for all municipalities. Table 12 contains the top 10 codes for all municipalities with highest average time to complete.

In the next step, we have identified which codes have the longest time to complete for individual municipalities. We identified from Municipality-1 that there were 26 codes in which an average time to complete longer than 50 days was seen. The longest time recorded in code 01_HOOFD_140 was 602 days. In Municipality-2, we identified 41 codes in which an average time to complete was longer than 50 days, and the longest recorded in code 11_AH_II_060c was 493 days. For Municipality-3, 18 codes and the longest time to complete was noticed with code 08_AWB45_051_2 at 373 days. In Municipality-4, we have identified 30 codes with the longest time, which was 523.5 days in code 01_BB_630. In Municipality-5, 30 codes were seen and the longest duration was 538.8 days with code 01_BB_640.

Table 12. Top 10 codes for municipalities with the highest avg. time to complete (in days).

ActionCode	M1	M2	M3	M4	M5	M1R	M2R	M3R	M4R	M5R	Avg. D.	Avg. M1	Avg. M2	Avg. M3	Avg. M4	Avg. M5
10_UOV_030	44	44	30	20	23	57.9	34.6	38	23.8	29.5	**58.7**	69.9	47.9	69.9	48.6	57.4
10_UOV_030a	25	32	14	29	14	92.6	88.9	45.2	82.9	48.3	**57.6**	61.8	63.1	51.6	51.0	60.4
01_HOOFD_530	66	228	19	190	424	16.5	74.5	6.4	48.2	92	**56.4**	45.7	85.3	48.6	65.6	37.0
01_HOOFD_516a	10	17	2	23	3	38.5	47.2	6.1	76.7	10.7	**55.4**	43.9	67.1	56.0	83.4	26.7
10_UOV_035	6	44	21	27	12	17.6	73.3	67.7	90	36.4	**43.8**	51.7	49.3	9.5	56.8	52.0
04_BPT_030	9	24	33	9	10	9.6	12.3	23.2	7.8	11.8	**42.8**	38.0	76.4	39.7	25.6	34.3
08_AWB45_025	51	29	46	30	62	28.2	28.4	21.9	27.5	29.2	**41.2**	41.7	53.7	23.1	36.0	51.6
08_AWB45_010	14	9	34	11	24	5	5.1	11	5.6	6.6	**39.1**	74.8	71.1	26.8	4.1	18.8
08_AWB45_020_2	187	112	181	117	215	64.9	62.9	62.4	61.6	59.6	**35.6**	43.9	37.8	31.4	31.7	33.1
01_HOOFD_190_2	16	11	18	14	14	11.3	10.3	11.5	17.3	5	**32.9**	22.1	34.9	34.9	32.0	40.7

Besides the above mentioned differences, we observed that differences in throughput times between the municipalities has been affected also by other factors. For example, in Municipality-4, we noticed a jump in metric active cases over time illustrated in Figure 12. These jumps are created by resources which close a various number of building permits at same time. For example, resource 560752 has closed on 17.04.2013 in a total of 76 cases. The throughput times of these cases were various from 27 days to 635 days. The number of such jumps in active cases over time in Municipality-4 is numerous and we assume that it unnecessarily prolongs the throughput time. It could also mean that the closure of building permits by system, considering that a completed application was sent already sooner but figured still to be an unfinished building permit. This has to be checked by the process owner to verify this, because without more detailed information about the process, we cannot deduce it to be a problem.

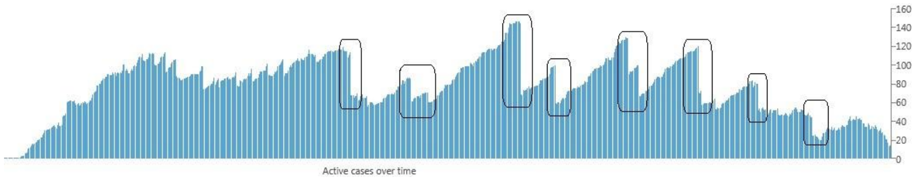

Fig. 12. Active cases over time for Municipality-4.

6

7 Control flow

During our investigation, we discovered that there are very few common characteristics in the processing of building permits between municipalities. For example, in the view of attributes "(case) last_phase" the most frequent values were seen here across different municipalities. The same applies to activities that represent steps in the processing of building permits. In this part of the work, we have identified differences in the processing of building permits. Once again, we have chosen the attribute "action_code" as a representative of activities. First, we identify what are the fundamental differences in flow control between municipalities through the analysis of activities used in processing of building permits. Table 13 shows the disparity in the use of codes in the processing of building permits.

There were a total number of 500 codes from which 275 are used in all municipalities. Differences in the use of the system for processing building permits is significant since only 55% of codes are used in all five municipalities. Municipality-2 had the most unique codes at an remarkable 409. The least amount of unique codes were seen used Municipality-4 at 355. During the control flow analysis, we found that there were codes specific to the individual municipality (e.g. 15_NGV_050 code is only used in Municipality-2), or were specific only for certain municipalities (e.g. 10_OLO_100 code is only used in Municipality-2/3/4). There were also steps in the process that have no value in the attribute "action_code" but did show having only labels. For example, "OLO received documents", which is only used in Municipality-2/4/5. We have identified codes that were used more frequently in specific municipalities, as seen through code 16_LGSV_010. Municipality-2 scored 243 times, Municipality-3 scored 372 times, Municipality-4 scored 421 times, while Municipality-1 scored 2 times and Municipality-3 scored 4 times.

Table 13. Disparity in use of codes among building permits.

	Count
Codes used by all municipalities	275
Codes used only by 4 municipalities	74
Codes used only in 3 municipalities	36
Codes used only in 2 municipalities	42
Codes used only in Municipality-1	14
Codes used only in Municipality-2	28
Codes used only in Municipality-3	12
Codes used only in Municipality-4	1
Codes used only in Municipality-5	18

7.1 Similarity in modified variants

Subsequently, we tried to identify control flows that would be common to all municipalities. We were trying to observe if there were any same cases between municipalities which were expressing the same set of activities. When identifying cases with the same set of activities between municipalities, we have not been given the chronological order of events in cases (such as default variants), but we were only looking for consensus on using the same activities in the approval of building permits. This modification was used because as we have mentioned in the work, there were observed timestamps which were inconsistent. To identify common cases we have created an algorithm that produces a modified variant in the individual municipalities and follow the same variant which connects through all the municipalities. Table 14 shows the number of found modified variants in every municipality.

Table 14. Count of modified variants found in each municipality.

	No. Modified Variants	No. Original Variants	No. of Cases
M1	703	1100	1199
M2	505	755	832
M3	590	1199	1409
M4	474	911	1053
M5	451	1003	1156

Total count of modified variants in all five municipalities were 2600. Subsequently, we were looking for modified variants among which of them were the highest in penetration. We identified only one modified variant, which appeared in all five municipalities. This variant has been defined by the set of 42 activities[2] and it covers 39 cases of all municipalities. Most cases in this variant originated from Municipality-5 with 27 out of the 39, 8 cases in Municipality-2 and 2 cases in Municipality-4. The lowest number of cases in this variant were seen in Municipality-1 and Municipality-3. Thereby, 6 modified variants of penetration were seen through four municipalities, 12 modified variants among three municipalities, and 77 through two municipalities, in which 30% of common penetrations were identified for Municipality-2 and Municipality-5.This analysis showed that in 34% of all penetrations in modified variants were in Municipality-2 and Municipality-5. We should consider that the smallest difference between the control flows is between the Municipality-2 and Municipality-5. We identified the most numerous modified variant for individual municipalities. Municipality-1 contained a maximum count of 34 cases in a modified variant, which has been defined by a set of 38 events. For Municipality-2, there were 47 cases in a modified variant with 61 events. For Municipality-3, we have identified that there were three of the most numerous modified variants. Two of which had a count of 46 and 44 events. Both instances had 82 cases and one of them with a count of 42 events in which was 81 cases. Municipality-4 had the highest count of 61 cases with 42 events. Municipality-5 had 61 cases with 46 events. The analysis shows that Municipality-3 had the most compact control flow because in the first three most numerous modified variants there were a total of 17.4% cases located. Results following were Municipality-4 with 14%, Municipality-5 with a 13%, Municipality-2 with 12%, and Municipality-1 with 7%.

[2] 01_HOOFD_010, 01_HOOFD_030_2, 01_HOOFD_065_2, 01_HOOFD_190_2, 01_HOOFD_015, 01_HOOFD_020, 01_HOOFD_030_1, 01_HOOFD_040, 01_HOOFD_060, 01_HOOFD_065_1, 06_VD_010, 01_HOOFD_100, 01_HOOFD_130, 01_HOOFD_050, 01_HOOFD_110, 05_EIND_010, 01_HOOFD_120, 01_HOOFD_180, 01_HOOFD_190_1, 01_HOOFD_200, 01_HOOFD_205, 01_HOOFD_270, 01_HOOFD_250, 01_HOOFD_260, 09_AH_I_010, 01_HOOFD_370, 01_HOOFD_375, 01_HOOFD_380, 01_HOOFD_430, 01_HOOFD_470, 01_HOOFD_480, 01_HOOFD_490_1, 01_HOOFD_490_2, 01_HOOFD_500, 01_HOOFD_510_1, 01_HOOFD_520, 01_HOOFD_490_3, 01_HOOFD_510_2, 01_HOOFD_530, 01_BB_540, 01_BB_770, 01_HOOFD_790

7.2 Composite events

Another step through which we described control flow and differences between municipalities were seen in the creation of composite events. By this method, we expected a substantial reduction in the number of activities and the creation of a more general control flow model for each municipality. In this method we could see instances of merged activities in processes regarding procedures in the phase for specific subprocesses or a main process. A new activity for a specific group was created once we found a group of events in the case where several procedures of same specific phase followed in subprocesses or a main process. This new activity has been represented in a subprocess or a main process with information regarding the phase in which they been seen to occur. New activity names were generated from original names so that from '01_HOOFD_xxx' we scraped off the last two digits and were left only with the first digit as an indicator of a phase within a process. In creation of this activity, we wanted at the same time to record the time interval necessary for the execution of subprocesses. The first timestamp was extracted from the first event as its start position and the end timestamp was extracted from the last event in the observed group. For the creation of composite events, we have noticed resources and monitoring resources which were included in procedures. Furthermore, we also recorded how many events or procedures had been included in the new composite event. Also, every case level attribute has been recorded. Figure 13 illustrates the procedure for creating composite events.

Activity	Complete timestamp		Activity	Start	End
01_HOOFD_490_2	19.12.2011 15:39:05				
01_HOOFD_491	19.12.2011 15:39:05	→	01_HOOFD_4	19.12.2011 12:39:05	19.12.2011 12:39:07
01_HOOFD_495	19.12.2011 15:39:07				
01_HOOFD_510_1	19.12.2011 15:39:08				
01_HOOFD_510_2	19.12.2011 15:39:08	→	01_HOOFD_5	19.12.2011 12:39:08	4.1.2012 0:00:00
01_HOOFD_520	19.12.2011 15:39:11				
01_HOOFD_530	4.1.2012 0:00:00				

Fig. 13. Illustration of merging composite events from part of case 5166463 in Municipality-5.

A new log was created and imported into Minit. In Table 15 we compared facilities of an original log, which is below seen in columns marked with “O” and logs with composite events in columns marked with “C”.

We have gained a new process flow from the viewpoint of the main processes and subprocesses. In such a modified process, there was a better identity of the mainstream seen more clearly. We have created process maps of all the municipalities, which contained 30% of the most frequent activities and only the most dominant process flows among those activities. We have chosen 30% the most frequent activities because we wanted to show in addition to the main process, the most frequent subprocesses. Figure 14 shows the process map of Municipality-4, its’ activities and edges are reduced in respect to the above mentioned characteristics and you can also see them displayed in a metric case count. Process maps with the same settings and same displayed metric of other municipalities are contained in Appendix Figures 25-29.

Table 15. List of municipalities and their activity records, events, cases, and variant differences.

	Activities		Events		Cases		Variants	
Municipality	**O**	**C**	**O**	**C**	**O**	**C**	**O**	**C**
1	398	**37**	52 217	**19 723**	1 199	**1 199**	1 100	**772**
2	409	**43**	44 354	**18 118**	832	**832**	755	**645**
3	383	**40**	59 681	**23 588**	1 409	**1 409**	1 199	**847**
4	355	**42**	47 293	**19 923**	1 053	**1 053**	911	**725**
5	386	**40**	59 083	**24 936**	1 156	**1 156**	1 003	**836**

We obtained a view of a process where we were able to identify the flow of the main processes taking place in a series of activities 01_HOOFD_0 → 01_HOOFD_1 → 01_HOOFD_2 → 01_HOOFD_3 → 01_HOOFD_4 → 01_HOOFD_5 → 01_BB_5 → 01_BB_7 → 01_HOOFD_8. Part of the process from 01_HOOFD_5 is not obtained in the appendix for Municipality-3. This aspect was visible with a higher count of visualized activities with less frequencies. Municipality-3 is thus different from the other four municipalities because of the majority of approximately 71% of all building permits were closed in activity 01_HOOFD_5.

Then we identified the characteristics of the subprocesses. Subprocesses have two main characteristics. The first is that, in certain activities of a main process flow continue to subprocess activity and then cycle back to the previous main process activity. Behavior can be observed in the process map seen on Figure 14. Here, we see that the start of subprocess 09_AWB45_0 activity from the main process 01_HOOFD_2 activity. After the operations in a subprocess, the flow continues back into the main process 01_HOOFD_2 activity. The second characteristic feature of the behavior of subprocesses is when some of the activity from the main processes follow subprocesses which take any action based on certain characteristics flow that proceed to another main process activity. The flow does not return back as the first type of behavior. When observing this phenomena more in depth, we discovered that this behavior is dependent mainly on procedures of which event attributes labeled as "question" takes the value "true" or "false". The value in the attribute "question" was influencing at the same time with the number of procedures in certain phases of subprocesses. The second characteristic is demonstrated in the process map on Figure 14. After the activity of the main processes 01_HOOFD_0, activity of sub-process 05_EIND_0 follow by which a case can be finished or a flow can carry onto an activity of subprocess 06_VD_0 where the case can again be finished or flow can carry on further to activities of the main process 01_HOOFD_1.

When a detailed analysis of the process maps of all municipalities took place, it became visible that the main process flow is common but in every municipality the subprocesses are carried out from different phases of the main process. If we compare the difference in predecessor and successor activities for activity 01_HOOFD_1, we can see the differences between the municipalities. In Municipality-1, predecessor activities 01_HOOFD_0 from the main process and two subprocess activities 06_VD_0 and 08_AWB45_0 are seen. Successors from the main process 01_HOOFD_2 and subprocess activity 08_AWB45_0 were also made apparent. In Municipality-2 predecessor activities 01_HOOFD_0 from main process and two subprocess activities

06_VD_0 and 07_OPS_0, had successors with activity from the main process 01_HOOFD_2 and subprocess activities 08_AWB45_0, 06_VD_0 and 07_OPS_0. In Municipality-3 predecessor activities 01_HOOFD_0 from the main process and subprocess activity 08_AWB45_0 were noted with successors from the main process 01_HOOFD_2 and subprocess activity 08_AWB45_0. In Municipality-4 predecessor activities 01_HOOFD_0 from the main process and two subprocess activities 06_VD_0 and 07_OPS_0, along with successor activity from the main process activity 01_HOOFD_2. Lastly, Municipality-5 predecessor activities 01_HOOFD_0 from the main process and two subprocess activities 06_VD_0 and 07_OPS_0 with successors of the main process activity 01_HOOFD_2 and subprocess activity 08_AWB45_0.

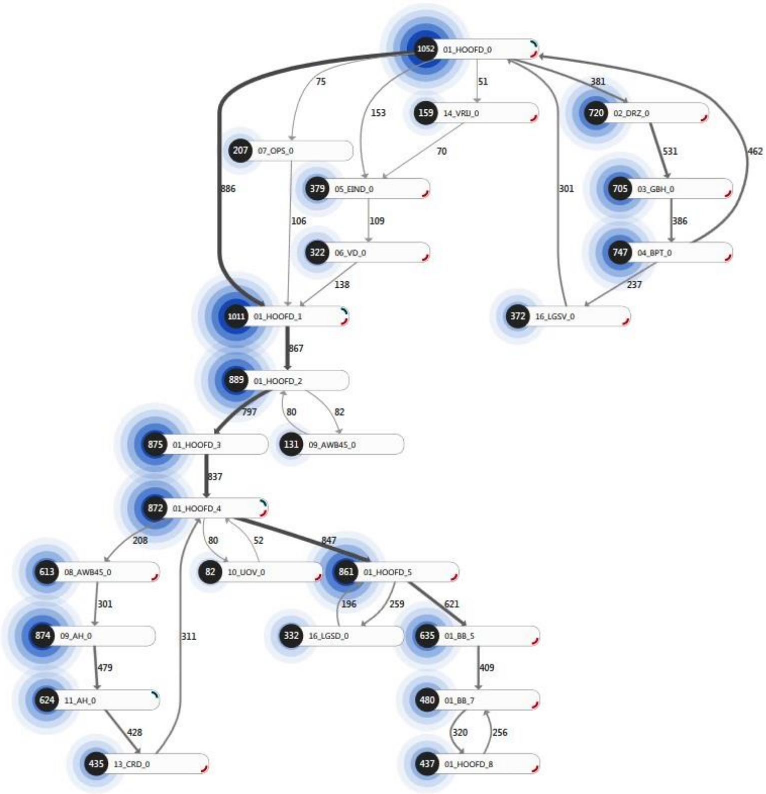

Fig. 14. Figure shows 30% most frequent activities and only the most dominant paths in composite process for Municipality-4.

Showcasing the differences of subprocess activities in a flow is great to observe the other main activities of a process. In Municipality-3 we found out that when this municipality is compared with others, it is has the smallest complexity seen in a process map. We can mostly observe that the first characteristic seen is in a property subprocesess. However, it can also be affected by the type of building permits, where it was not necessary to perform some procedures. When comparing the number of cases located in the first 5 of the most numerous variants, we found that out of them, Municipality-4 has 13% of cases seen, followed by Municipality-1 with 12%, and Municipality-3 with 11%. In Municipality-2 the first 5 variants represented 8 % in cases and Municipality-5 had only 7 % cases. If we compare the process map of Municipality-5 with other municipalities, we can confirm that even 30% of the most frequent activities show exactly the most complex flow of this municipality. This comparison is not the same for previous parts of this paper, where Municipality-3 had the most compact flow. However, omitting the procedures in phases of subprocesses, we have removed many exceptional behaviors. According to the frequency of cases in variants, we divided municipalities into groups according to flow similarities. The first group is consisted from Municipality-1/3/4. The second group is consisted from Municipality-2 and Municipality-5.

4

8 Outsourced procedures

This part of paper describes the effect it would have to outsource some of the procedures in the individual municipalities in regards to the organizational structures. To answer this question, we made observations based on two variations. First, we tried to identify the impact on the organizational structure in individual municipalities and then we researched the subsequent impact on the role that we had identified in the individual municipalities.

When searching for impacts on organizational structures, we need to mention that outsourcing some of these procedures will only affect individual employees in the organizational chart. Organizational structure will not be influenced. It may occur that, the employee who acts as a resource as well as a monitoring resource will fall from a level of monitoring resource. For example, procedures that an employee monitored would not be present in the future in the process, an employee would solely remain at the resource level.

Subsequently, we tried to identify what effect it would have to outsource some of the procedures in resource roles for individual municipalities, which we identified in section 4.2 of this paper. As an example, we have chosen procedures that were represented by subprocesses 10_UOV and 08_AWB45. If these procedures would not be present in the future process, it would influence the role of the resources in all municipalities. In Municipality-1, we see that it would affect identified roles: Role 1, Role 2, Role 4, and Role 5. These identified roles would disappear and there would only conclude completely new roles created, or the resources identified in these roles would be shifted to another unaffected roles. In Municipality-2 would be affected Role 2 and Role 7. In this case, resource 560519, which was identified in Role 7 could change their role to Role 10. In Municipality-3, Role 1, 5, 8, 9 would be affected. In Municipality-4 Role 2 would be affected, and Municipality-5 Role 1 would be affected. This observation can be repeated for a different combination of procedures, but at this point, the process owner needed to identify procedures which he or she wanted to outsource in the future.

9 Changes in the processes after relocation

For monitoring the changes in the processes after employees moved into the same location, we have decided that we will observe differences in processing building permits through the use of codes, log attributes, and between resources.

First, we made a survey used by individual codes, for which each code contains information on the date when it was used in the system the first and last time in individual municipalities. By the survey we have identified which municipality has seen the greatest change in the process, e.g. how many new codes were added during the reporting period, and how many codes are no longer used. The time interval in which we measured the number of codes that have been traced for the first time in the log, started from 1.6.2014 until the date of the last event captured in the logs. In the identification codes, which are no longer used, we chose the time interval from 1.6.2014 until the end of 2014. This time, intervals have been limited at the end of 2014, therefore, we do not include these in our observation of codes on pending building permits (the vast majority of building permits were approved up to 12 weeks). Table 16 displays the listing number of codes, which have started and stopped being used in the time interval for individual municipalities.

We observed that new codes were mostly introduced in Municipality-2 and Municipality-5. The same applies for the codes which are no longer used. Most new codes from 2015 were introduced in Municipality-2 (9 new codes) and Municipality-5 (6 new codes). In Municipality-4, no observed new code has been seen since early 2015. Interesting enough was, Municipality-4 was not starting to use any new code from 23.12014 until 07.14.2014 and then later from the period of 15.07.2014 to 26.11.2014, we noticed 22 new codes.

Table 16. Municipalities and count of their used codes defined by dates.

	Number of new codes (since 1.6.2014)	**Number of unused codes (1.6.2014-31.12.2014)**
M1	13	50
M2	**29**	**82**
M3	11	69
M4	22	71
M5	**32**	**73**

Changes were also observed between employees in Municipality-2 and Municipality-5. Employee number 560530 worked in Municipality-2 from 11.10.2010 to 06.27.2014 and we also identified this employee among employees in Municipality-5, where the date of his first appearance was seen in the log from 14.05.2014 and the date of last appearance was seen in 03.03.2015. We also identified the same for employee number 560532 from Municipality-2. This employee from 05.12.2014 also worked in both municipalities simultaneously. Such a transfers of employees was observed in the log several times (e.g., employee 560752 from Municipality-4 also worked in Municipality-5 from 15.3.2012 to 28.3.2013) but the two mentioned above are the most recent.

Changes were also observed in the case attribute “(case) Includes_subCases”. All municipalities apply cases in which a start time occurs after 1.6.2014 and have an attribute with an empty value.

To our amuse, Municipality-2 had a full working week from 08.12.2014 to 12.12.2014 , which has not adopted any new building permit applications but the following week adopted 10 new building permit applications. This observation may indicate that the employees may have been moving from Municipality-2, but we cannot confirm this observation without more detailed information on processes.

Conclusion

To conclude this analysis we present an overview of results, which were based on five process logs of building permit applications of the five Dutch municipalities.

Resulting from the analysis of people involved in the various stages of the process, we have gained a list of 13 identified roles across all municipalities. One common role was found to be common among all municipalities. The largest number of roles has been identified in Municipality-1, which also has the overall largest number of employees. This diversity of roles reflects the system usage differences across the municipalities.

A common organizational structure has been disclosed for all municipalities. This structure served as a basis for organizational chart creations per municipality. Core relations between the employees have been identified and have helped us in understanding the responsibility flow between them. A possible improvement point has been suggested for Municipality-5’s organization structure, where the 3rd hierarchical level representing the resources was not clearly observed. Organizational charts for all municipalities have been introduced and they showed the diversity of roles by attributes among the employees. We have also identified the responsibility distribution diversity for employees involved in the process, based on the created relations. A possible responsibility distribution improvement point has been identified in Municipality-1, which might serve as an example for all other municipalities. The most allocated resources have been identified and visualized in social networks of their respective municipalities. We have identified Resource 560752, who processed 88% of all building permits in Municipality-4, as well as Resource 560781 in the same municipality, who pushed their work to Resource 560752 in 50% of all building permits, which in the same time were the highest observable values overall.

Our initial exploration of throughput times leads us to define four methods for explaining differences between municipalities. Firstly, we have identified differences between municipalities based on attribute “(case) parts”. The observed values in the attribute obtained, was that Municipality-3 had the shortest throughput time. Next, we have identified the time intervals, in which the building permits have been processed in specific municipalities, as well as the number of process steps for the longest throughput times observed per municipality and for all municipalities together. We observe that Municipality-3 has resolved 60.3% building permits within seven weeks. The largest average in throughput times for all municipalities was 433 days in cases where a total of 83 steps took place to complete building permits. The median value seen was 50 events per case for all municipalities. Municipality-5 obtained the fastest average time upon completion with 50 days. Using the time to complete metric, we uncovered activities which are causing throughput times to grow. Municipality-2 was identified as having the most amount of codes (41 in number) with an average time to complete being longer than 50 days. The longest time to complete recorded was 602 days. This was seen utilizing code, 01_HOOFD_140 in Municipality-1.

Differences as well as common attributes in used activities have been described. Overall, we have identified 500 unique activities, 55% out of which are used by all municipalities. In the modified process variants that were identified, we found that Municipality-3 has the most compact control flow with 17.4% cases in 3 of the most numerous modified variants. More general control flow was modeled by creating composite events. We have introduced two subprocess characteristics, which helped us reveal the flow difference. Municipality-3 completed 71% of its building permits in the 5th stage of the main process, as opposed to other municipalities, which closed building permits mostly in 8th stage of the main process. Two groups have been created based on the similarity of control flow, main process and created variants. The first group included Municipality-1/3/4 and Municipality-2/5, which has been included in the second group. During analysis, Municipality-2 and Municipality-5 had the most common characteristics, thus we are assuming a certain connection between them.

Thus concluding that, future outsourcing of procedures, will not have an effect on organizational structures but will have an effect on identified resource roles

We have described the process differences that have been observed starting from this particular date, 1.6.2014. We have observed the changes in used activities, employees, attributes and in the number of incoming building permit requests. As a result, we cannot tell for certain, which resources were physically moved into the same location of a municipality, but we can presume that employees of Municipality-2 did in fact, physically move.

Overall we believe, that our analysis brought a better insight into the provided processes and that we helped the municipalities answer the stated questions.

References

1. Jessica W. Sun, Joseph Barjis, Alexander Verbraeck, Marijn Janssen, Jacco Kort, “Advances in Enterprise Engineering III: Capturing Complex Business Processes Interdependencies Using Modeling and Simulation in a Multi-actor Environment”, Springer (2009)

2. W.M.P. van der Aalst, “Process Mining: Discovery, Conformance and Enhancement of Business Processes”, Springer (2011)

3. doi:10.4121/uuid:a0addfda-2044-4541-a450-fdcc9fe16d17

4. doi:10.4121/uuid:63a8435a-077d-4ece-97cd-2c76d394d99c

5. doi:10.4121/uuid:ed445cdd-27d5-4d77-a1f7-59fe7360cfbe

6. doi:10.4121/uuid:679b11cf-47cd-459e-a6de-9ca614e25985

7. doi:10.4121/uuid:b32c6fe5-f212-4286-9774-58dd53511cf8

8. IEEE Task Force on Process Mining. “Business Process Intelligence Challenge (BPIC).” 11th International Workshop on Business Process Intelligence 2015, http://www.win.tue.nl/bpi/2015/challenge

Appendix

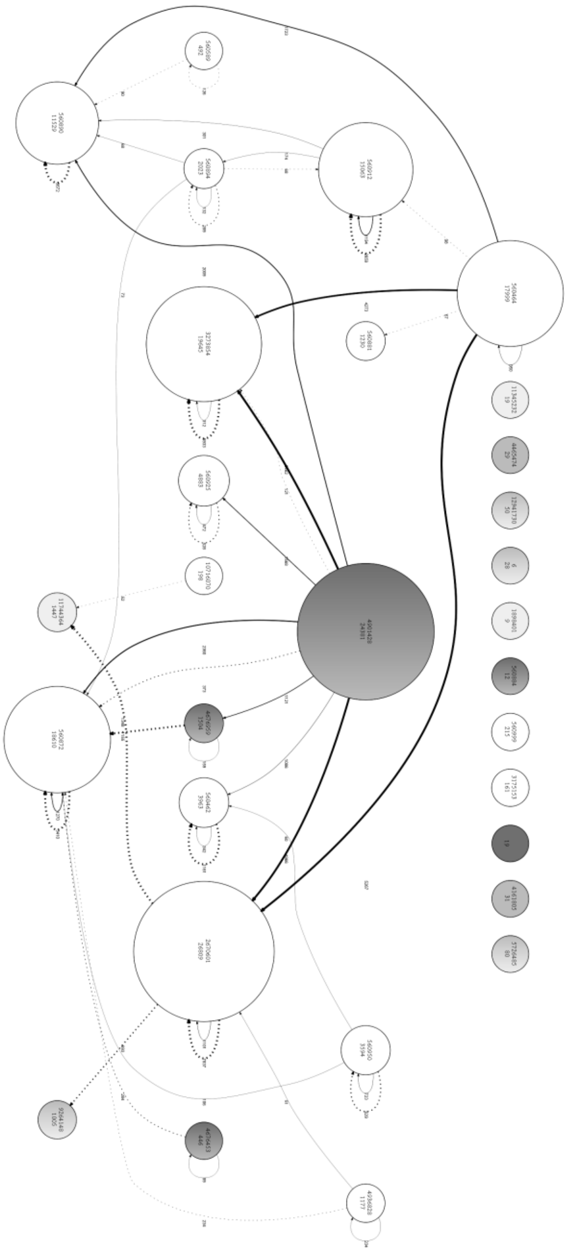

Fig. 15. Organizational chart for Municipaliyt-1 with limitation on average relation event frequency and filtered relations with less frequency than a value of 50 at both levels.

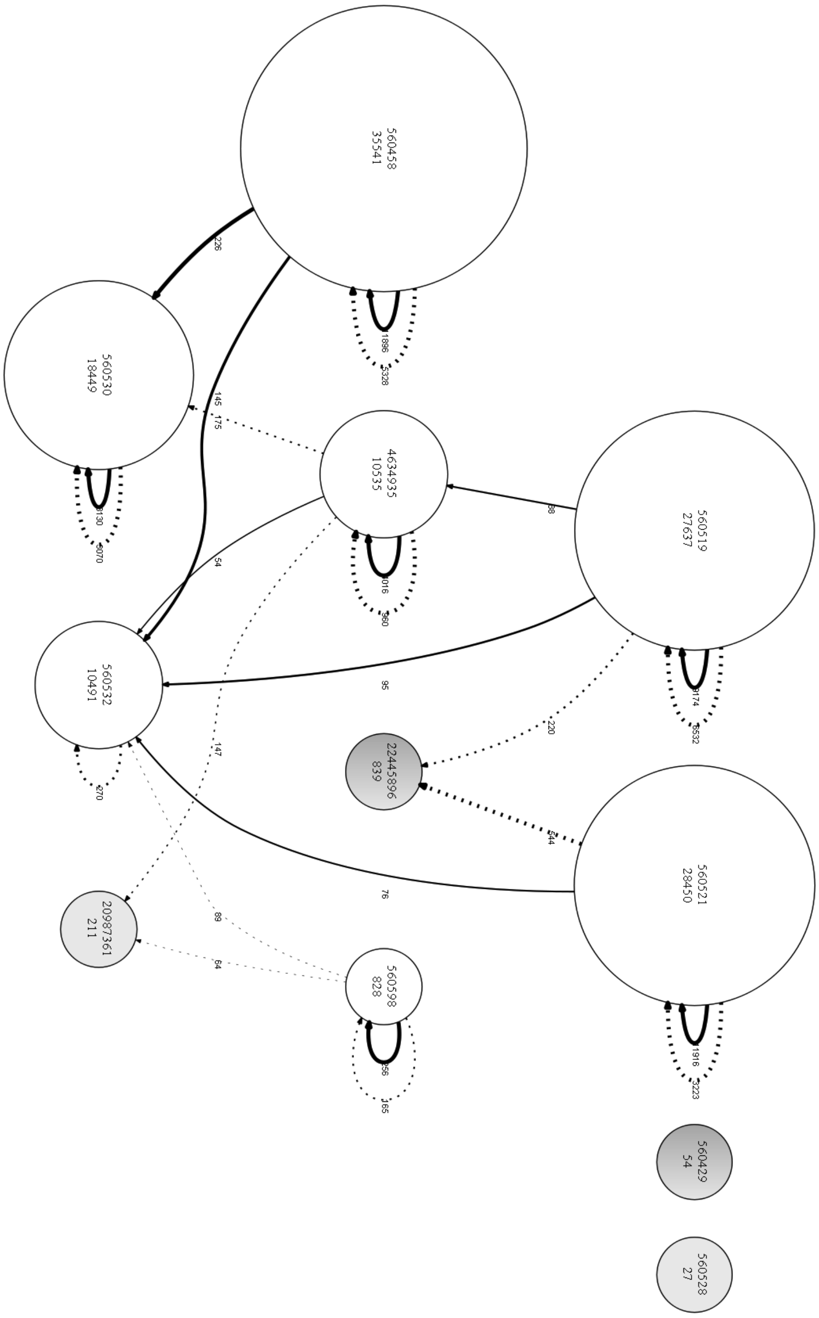

Fig. 16. Organizational chart for Municipaliyt-2 with limitation on average relation event frequency and filtered relations with less frequency than a value of 50 at both levels.

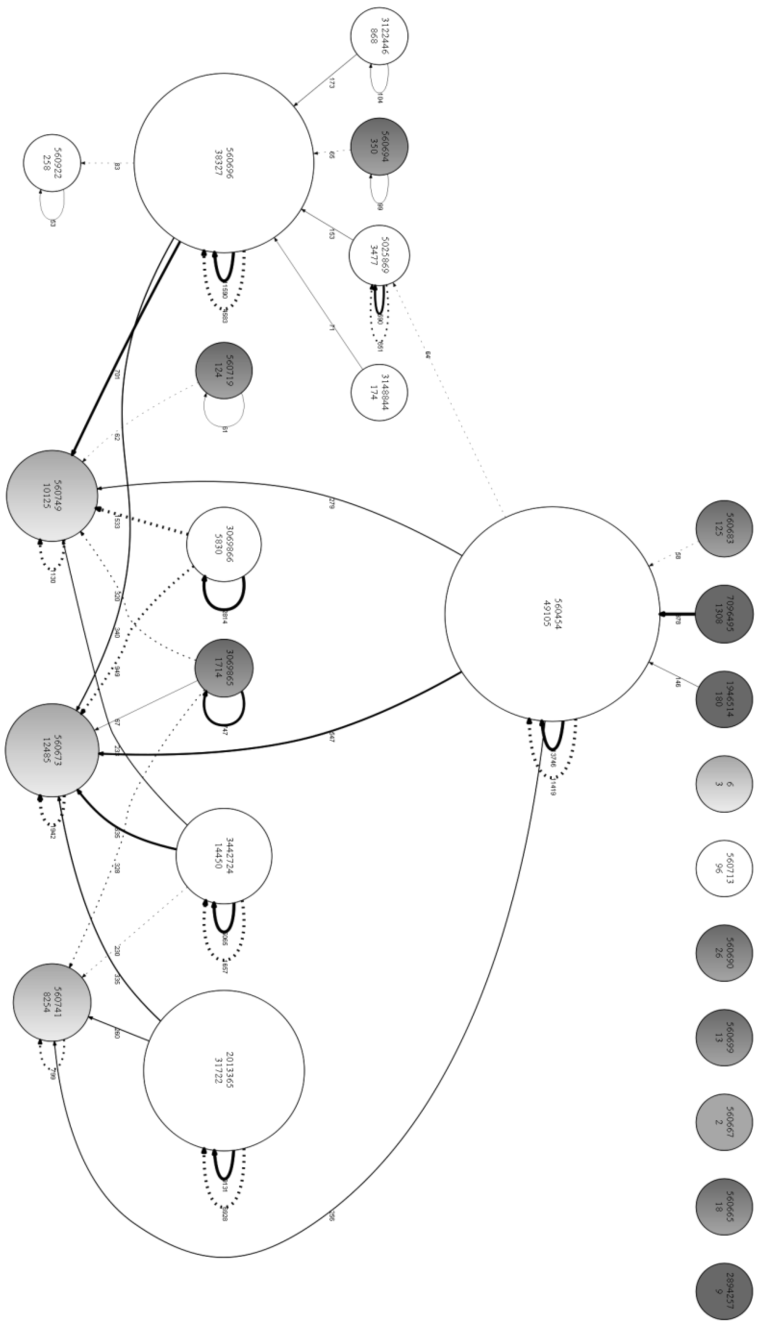

Fig. 17. Organizational chart for Municipaliyt-3 with limitation on average relation event frequency and filtered relations with less frequency than a value of 50 at both levels.

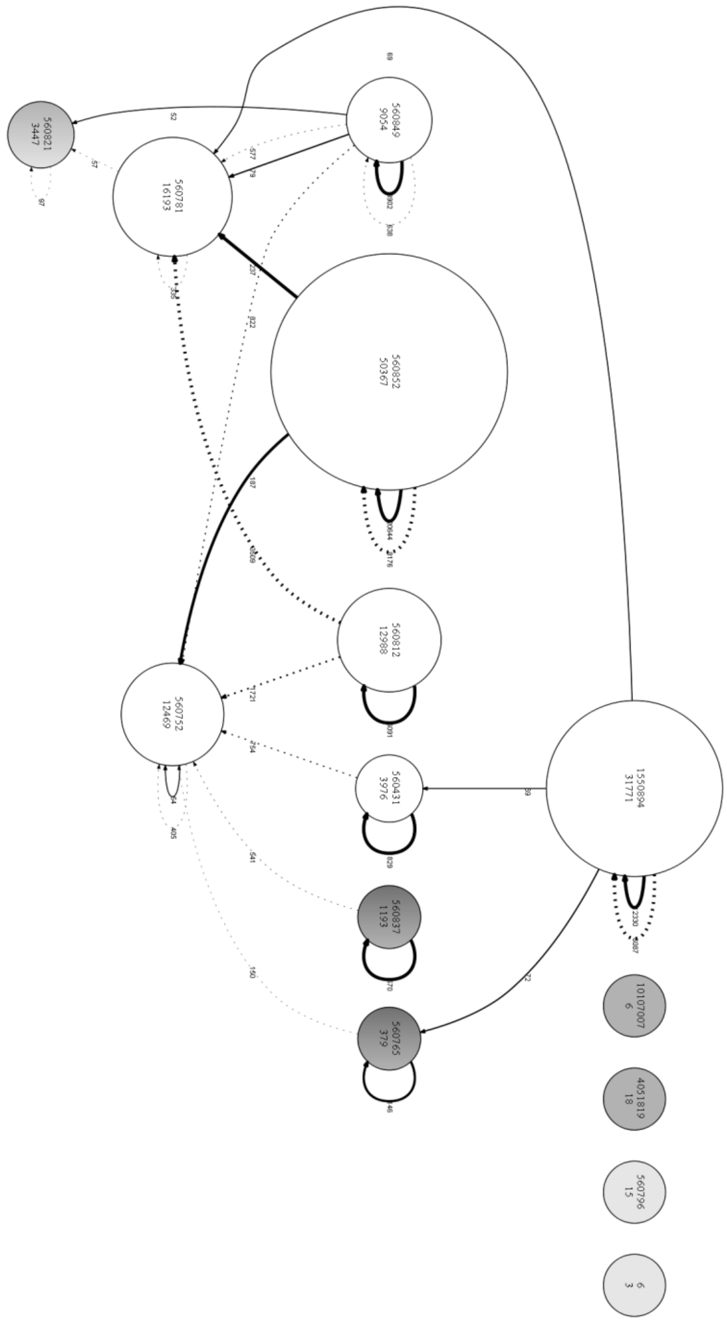

Fig. 18. Organizational chart for Municipaliyt-4 with limitation on average relation event frequency and filtered relations with less frequency than a value of 50 at both levels.

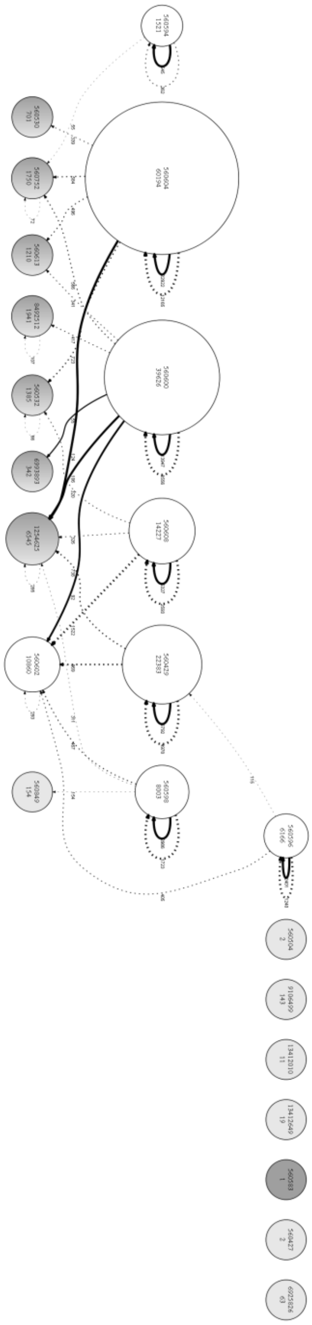

Fig. 19. Organizational chart for Municipaliyt-5 with limitation on average relation event frequency and filtered relations with less frequency than a value of 50 at both levels.

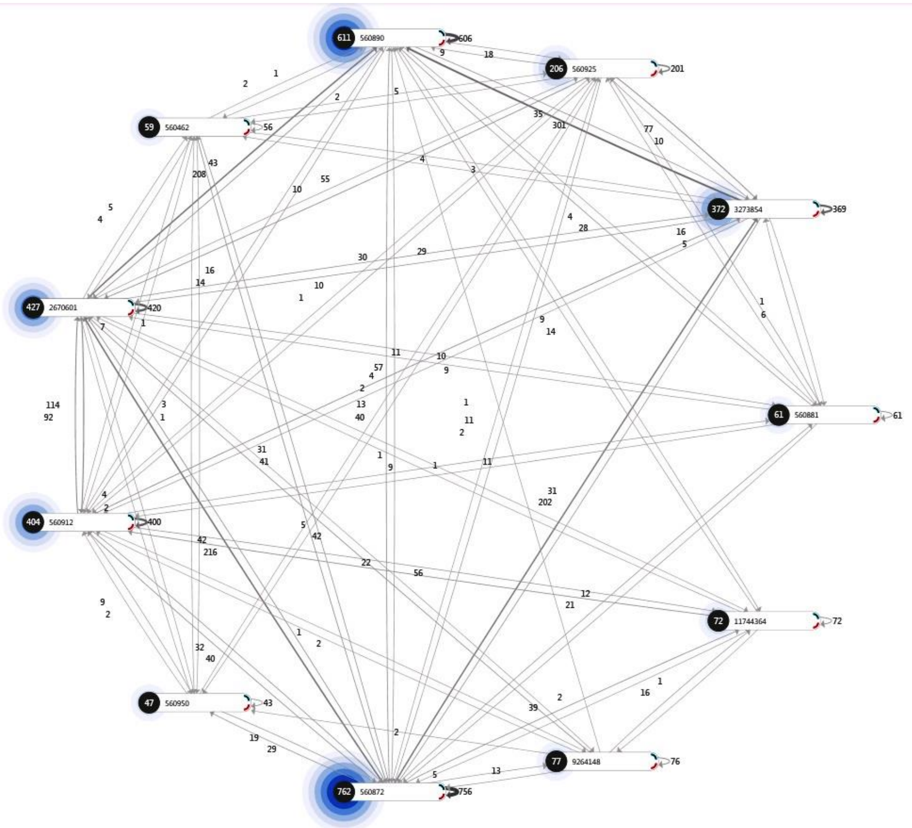

Fig. 20. Social network of Municipality-1, resources with relative event frequency greater than 1 % are displayed.

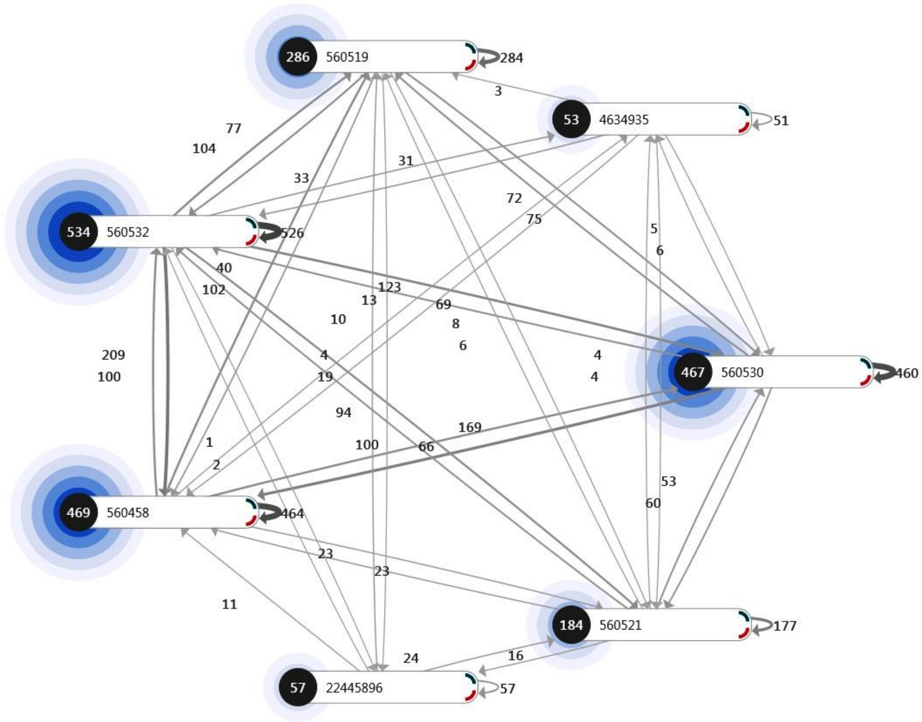

Fig. 21. Social network of Municipality-2, resources with relative event frequency greater than 1 % are displayed.

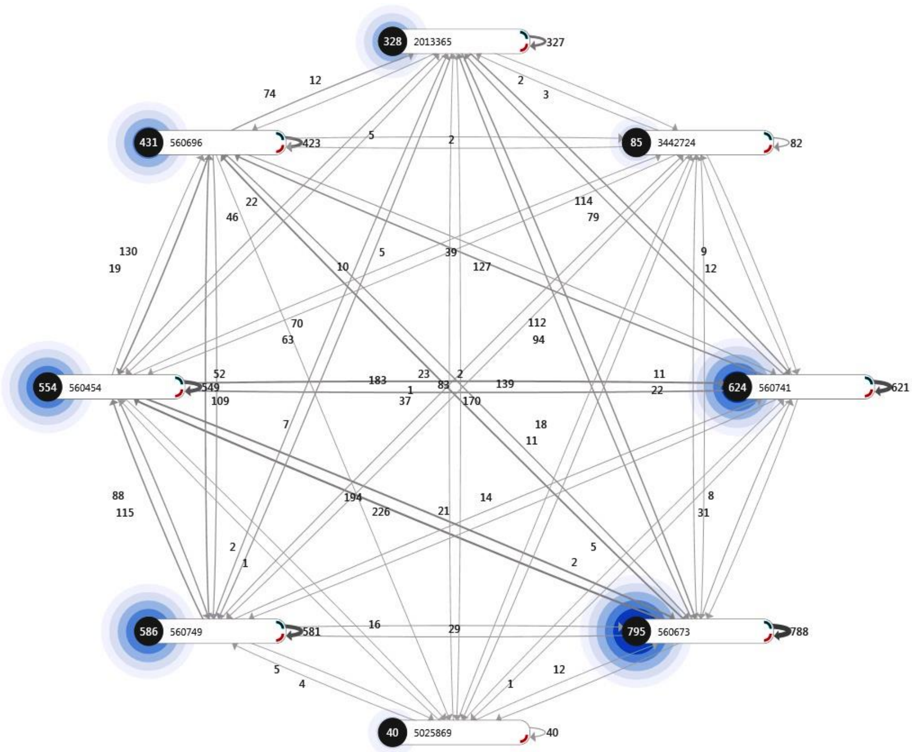

Fig. 22. Social network of Municipality-3, resources with relative event frequency greater than 1 % are displayed.

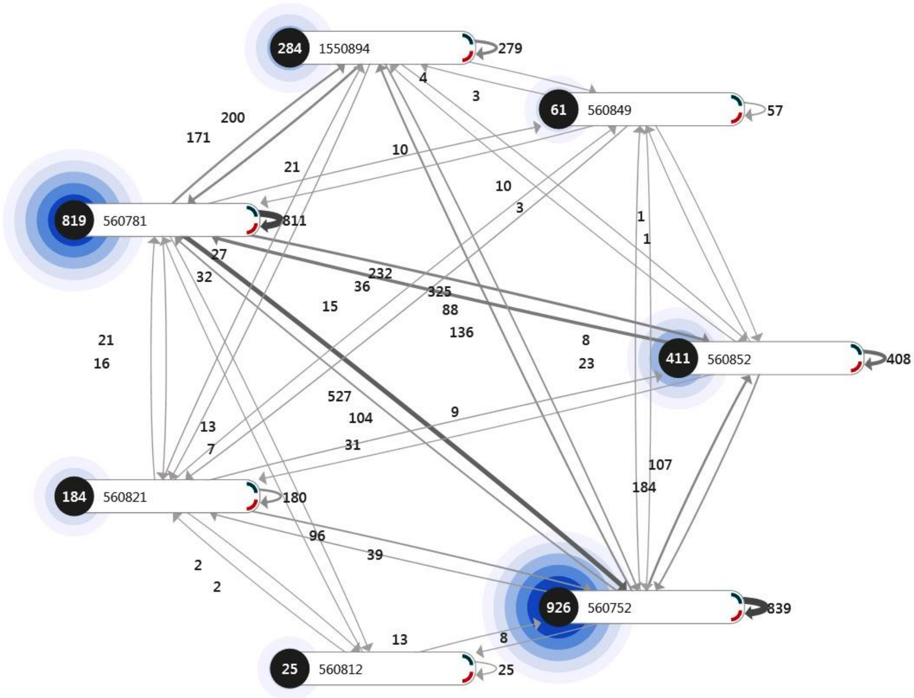

Fig. 23. Social network of Municipality-4, resources with relative event frequency greater than 1 % are displayed.

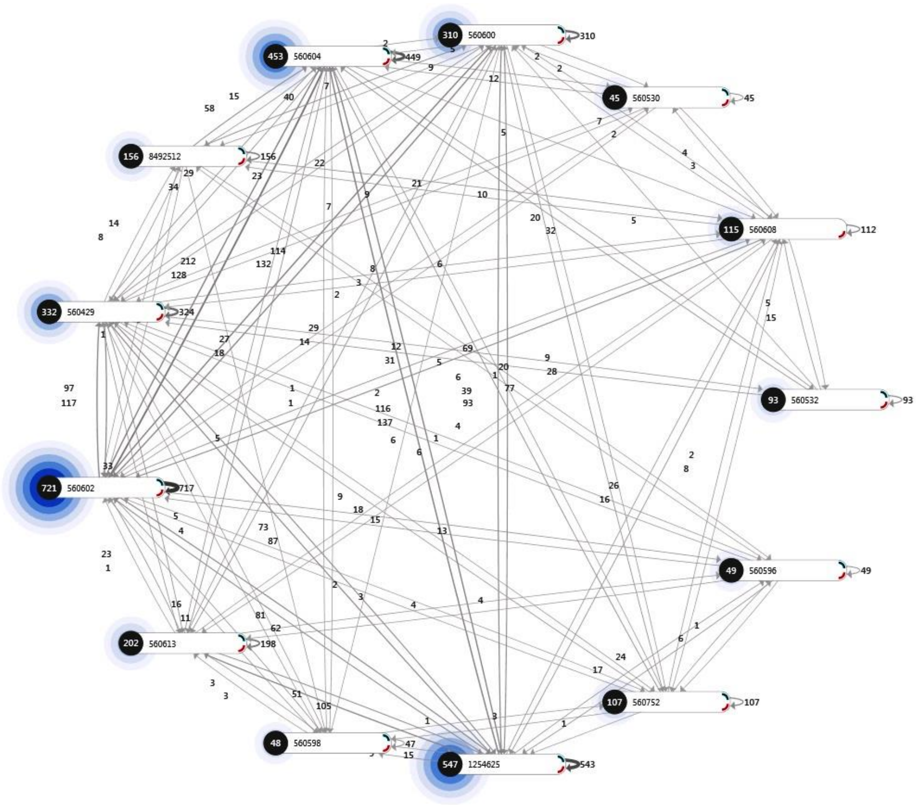

Fig. 24. Social network of Municipality-5, resources with relative event frequency greater than 1 % are displayed.

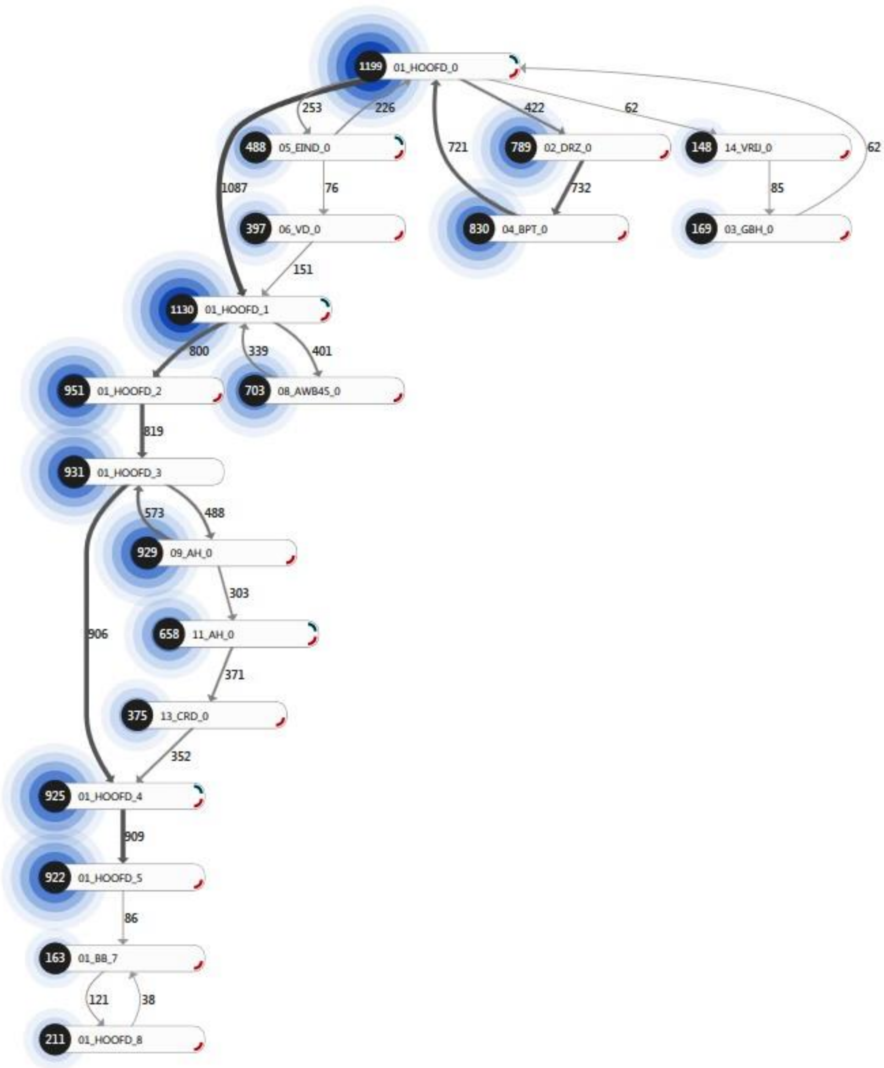

Fig. 25. Figure shows 30% most frequent activities and only the most dominant paths in composite process for Municipality-1.

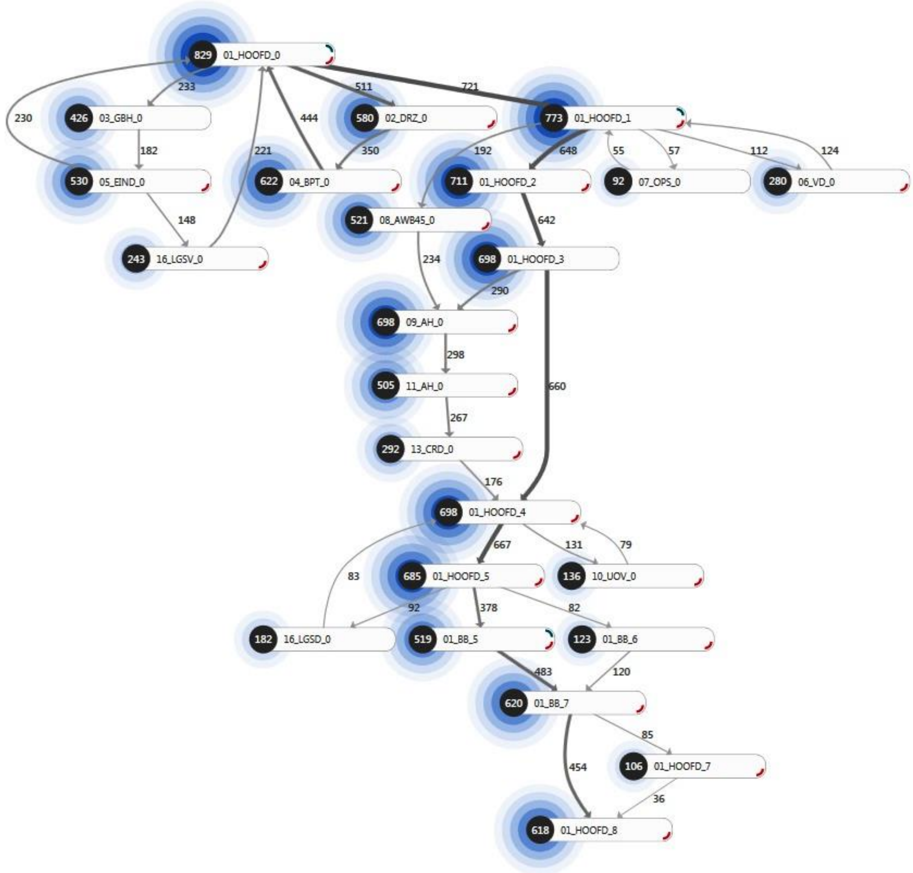

Fig. 26. Figure shows 30% most frequent activities and only the most dominant paths in composite process for Municipality-2.

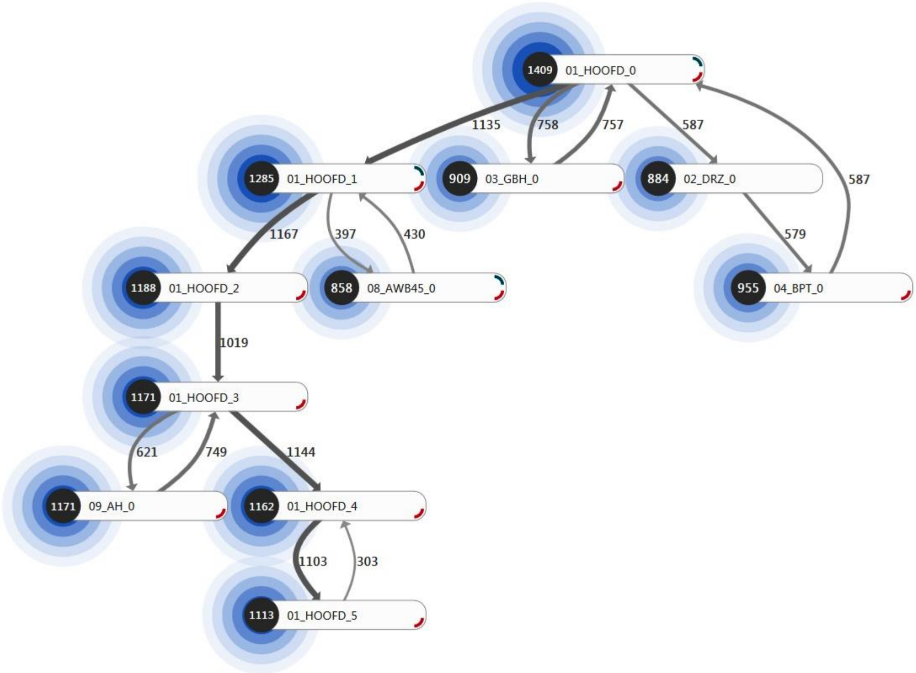

Fig. 27. Figure shows 30% most frequent activities and only the most dominant paths in composite process for Municipality-3.

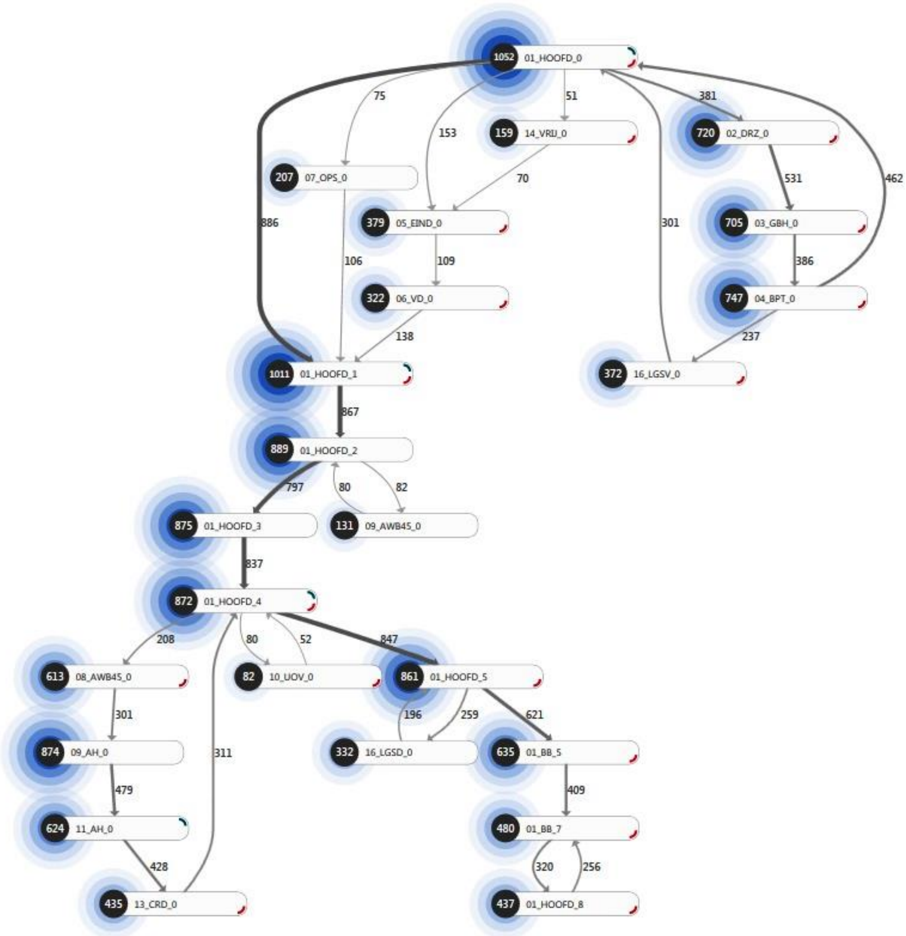

Fig. 28. Figure shows 30% most frequent activities and only the most dominant paths in composite process for Municipality-4.

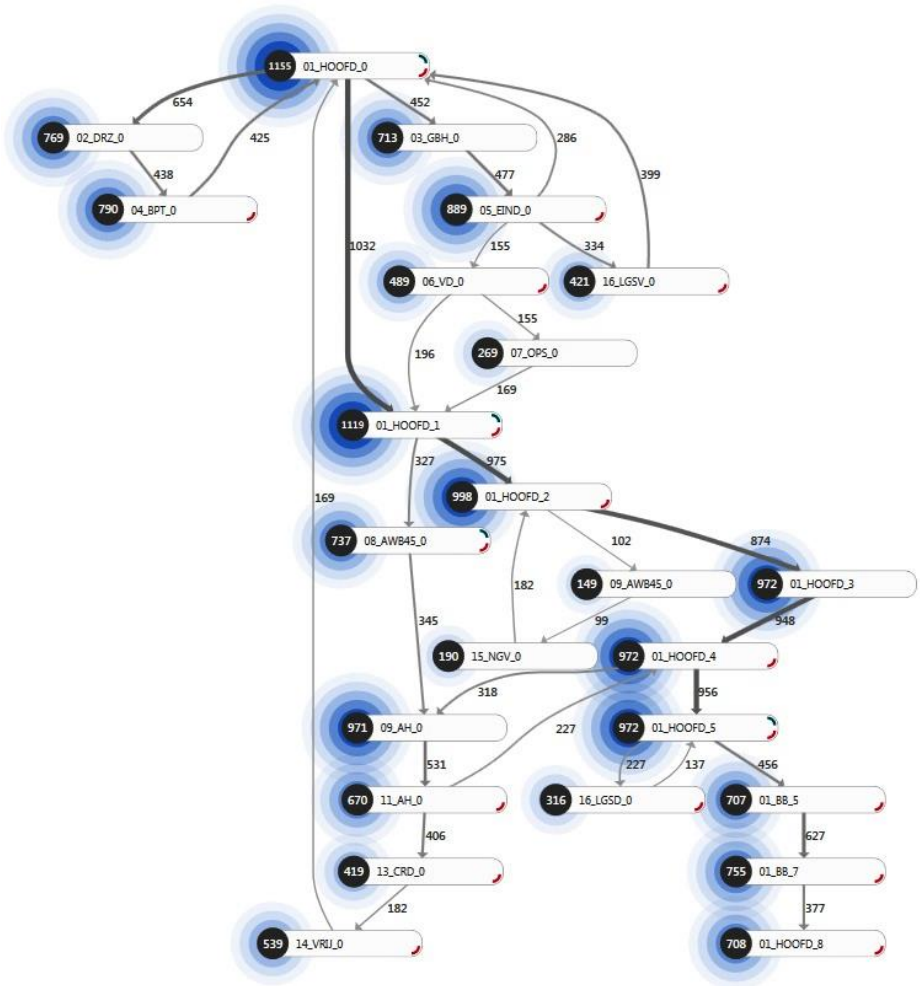

Fig. 29. Figure shows 30% most frequent activities and only the most dominant paths in composite process for Municipality-5.