A Predictive Model for the Impact of Changes on the Workload of Rabobank Group ICT's Service Desk and IT Operations

BPI Challenge 2014

Marcus Dees[1], Femke van den End[2]

[1] Uitvoeringsinstituut Werknemersverzekeringen (UWV), The Netherlands
[2] Consultrend, The Netherlands

Abstract. Every company has to deal with an increase in the number of changes on their IT systems. Due to security and performance issues it is necessary to keep the IT systems up to date. Having to do this lays a big strain on the IT organisation. Also the Rabobank Group ICt department has to optimize their processes to keep everything up and running. The planning of the capacity of the IT department to implement all necessary changes, is becoming more important. Knowing when which capacity is needed creates the slack the department needs to deal with unplanned work. This paper presents a predictive model with which this can be done. It is also shown how the size and duration of each change can be predicted. This is the input that is needed op optimize the planning.

1 Introduction

Every organisation has to deal with an increase in the velocity in which its IT environment is changing. New versions of software and improved hardware have to be implemented with a growing frequency. Keeping up is important to prevent security vulnerabilities, maintaining performance of all IT systems and making it possible to implement new technologies.

It is the job of the IT department to realise all these changes for the organisation. Next to that the IT department has to maintain a service desk where customers can ask questions and report problems they encounter. A contact with the service desk is referred to as an interaction. The IT department is also responsible for investigating and solving incidents that are triggered by issues reported to the service desk or by their own monitoring systems. The IT department must keep up the minimum service levels that the organisation needs to perform their primary duty. It is not enough to be reactive by answering questions and solving incidents as quickly as possible. It is critical to be proactive and to prevent major problems from happening.

The actions of the IT department are limited by the number of resources that are available. If all capacity is devoted to interaction and incident handling, then there is no time left to implement changes. On the other hand there is a limit to the number of changes that can be done on a system. There is a maximum number of resources that can be working on implementing changes. The question is how the IT department can keep up with all necessary changes while maintaining the service levels to the organization (and not using up too much resources on interaction and incident handling).

Rabobank Group ICT is an example of an IT department that has to deal with this question. Rabobank Group ICT has implemented the ITIL-processes and uses the Change-process for implementing planned changes. The Service Desk handles all interactions, i.e. telephone calls and e-mails of customers regarding issues they encounter when working with the Rabobank IT systems. The IT Operations department handles all incidents that are triggered by an interaction or that are detected by an automatic mechanism. The Business Change Management department plans and executes all planned and unplanned changes to the IT systems of Rabobank.

The BPI Challenge 2014 focuses on the question what impact a change has on the workload of the Service Desk and IT Operations. When the impact of change can be predicted then the planning of change implementations can be done with improved utilization of all the available resources and with less impact on the Service Desk and IT Operations. This makes it possible to handle more changes with the same amount of resources, hence keeping up with all necessary changes without the need to hire additional staff. Rabobank Group ICT would like to have a predictive model for the impact of changes, which can be made operation in a BI environment.

The main research question of the Rabobank is divided into sub-questions.

1. **Identification of Impact-patterns** Rabobank Group ICT expects there to be a correlation between the implementation of a change and the workload in the Service Desk (SD) and/or IT Operations (ITO), i.e. increased/decreased volume of Closed Interactions and/or increased/decreased volume of Closed Incidents. Rabobank Group ICT is interested in identifying any patterns that may be visible in the log for various service components to which a configuration item is related, in order to predict the workload at the SD and/or ITO after future changes.
2. **Parameters for every Impact-pattern** In order to be able to use the results of prior changes to predict the workload for the Service Desk directly after the implementation of future changes, the following parameters are of interest for Rabobank Group ICT for every impact-pattern investigated in sub question 1, *What is the average period to return to a steady state* and *What is the average increase/decrease of Closed Interactions once a new steady state is reached?*
3. **Change in Average Steps to Resolution** Project managers are expected to deliver the same or better service levels after each change implementation, Rabobank Group ICT is looking for confirmation that this challenge is indeed being met for all or many Service Components.

Finally, also an appeal is made to our creativity. We are challenged to surprise Rabobank Group ICT with new insights on the provided data to help change

Interaction	
CI Name (aff)	Confiuration Item (CI) where a disruption of an ICT Service is noticed, this is what we call the "Affected CI". A Service Desk Agent always uses questions in a Knowledge Document (identified by a KM number) to find the correct CI in theConfiguration Item Database (CMDB)
CI Type (aff)	Every CI in the CMDB is related to a Type.
CI Subtype (aff)	Every CI in the CMDB is related to a Subtype, which is related to a CI Type.
Service Comp WBS (aff)	Every CI in the CMDB is related to 1 Service Component, in order to identify which Product Manager is responsible for the CI. A Service Component is equal to a product in the Bill of Material and is part of one or more Services.
Interaction ID	The unique ID of an Interaction-record in HP Service Manager 9.
Status	Status of the Interaction-record.
Impact	Impact of the Service disruption to the customer.
Urgency	Indication of how urgent the customer needs a solution.
Priority	Impact and Urgency lead to a Priority for the Assignment Group to resolve the Service disruption.
Category	In order to select and compare similar Interactions in HP Service Manager 9, all records are Categorized. The categorization is derived from the Knowledge Document.
KM number	A Knowledge Document contains default values for the Interaction and a set of questions for a Service Desk Agent to derive which Configuration Item is disrupted and to determine Impact and Urgency for the customer.
Open Time (First Touch)	Date and time the Interaction record was opened in HP Service Manager 9.
Close Time	Date and time the Interaction record is closed in HP Service Manager 9.
Closure Code	Short code to classify the type of Service disruption.
First Call Resolution	Flag which indicates if the Service Desk Agent was able to provide the customer with a workaround for the Service disruption (Y) or if it was necessary to assign the record to a specialist by creating an Incident record.
Handle Time (secs)	Time registered to resolve the Interaction, based on a SLA Clock mechanism in HP Service Manager 9.
Related Incident	Record-number of the Incident created from this Interaction.

Incident	
CI Name (aff)	Confiuration Item (CI) where a disruption of an ICT Service is noticed, this is what we call the "Affected CI". When a Service Desk Agent decides to create an Incident from an Interaction, the Affected CI is copied from the Interaction-record into the Incident-record.
CI Type (aff)	Every CI in the CMDB is related to a Type.
CI Subtype (aff)	Every CI in the CMDB is related to a Subtype, which is related to a CI Type.
Service Component WBS (aff)	Every CI in the CMDB is related to 1 Service Component, in order to identify which Product Manager is responsible for the CI. A Service Component is equal to a product in the Bill of Material and is part of one or more Services.
Incident ID	The unique ID of an Incident-record in HP Service Manager 9.
Status	Status of the Incident-record.
Impact	Impact of the Service disruption to the customer.
Urgency	Indication of how urgent the customer needs a solution.
Priority	Impact and Urgency lead to a Priority for the Assignment Group to resolve the Service disruption.
Category	In order to select and compare similar Interactions in HP Service Manager 9, all records are Categorized. The categorization is derived from the Knowledge Document.
KM number	A Knowledge Document contains default values for the Interaction and a set of questions for a Service Desk Agent to derive which Configuration Item is disrupted and to determine Impact and Urgency for the customer.
Open Time	Date and time the Incident record was opened in HP Service Manager 9.
Reopen Time	Date and time the Incident record was reopened in HP Service Manager 9. This option is used when an Incident record was closed and within a short period of time it is discovered that the resolution is not effective for the customer.
Resolved Time	Date and time the Service disruption is resolved.
Close Time	Date and time the Incident record is closed in HP Service Manager 9.
Handle Time (Hours)	Time registered to resolve the Incident, based on a SLA Clock mechanism in HP Service Manager 9.
Closure Code	Short code to classify the type of Service disruption.
Alert Status	Alert status of the Incident-record, during it's lifecycle, based on defined Service Levels in HP Service Manager 9.
# Reassignments	Number of Incident Activities with Actvity Type"Reassignment".
# Related Interactions	Number of related Interactions to this Incident.
Related Interaction	Record-number if only one Interaction is related to this Incident.
# Related Incidents	Number of similar Incidents, related to this record, which will be used as parent record for ligging the Incident.
# Related Changes	Number of related Changes to this Incident.
Related Change	Record-number if only one Change is related to this Incident.
CI Name (CBy)	Confiuration Item (CI) which caused the disruption of an ICT Service, this is what we call the "CausedBy CI". When an Operator resolves an Incident from an Incident, the CausedBy CI must be registered before closing the Incident-record.
CI Type (CBy)	Every CI in the CMDB is related to a Type.
CI Subtype (CBy)	Every CI in the CMDB is related to a Subtype, which is related to a CI Type.
ServiceComp WBS (CBy)	Every CI in the CMDB is related to 1 Service Component, in order to identify which Product Manager is responsible for the CI. A Service Component is equal to a product in the Bill of Material.

Incident Activity	
Incident ID	The unique ID of an Incident-record in HP Service Manager 9.
DateStamp	Date and time when this specific Incident Acivity started.
IncidentActivity_Number	Unique ID for an Incident Activity.
IncidentActivity_Type	Short code to identify which type of Incident Activity took place.
Interaction ID	The unique ID of an Interaction-record in HP Service Manager 9.
Assignment Group	The team responsible for theis Incident Activity.
KM number	A Knowledge Document contains default values for the Interaction and a set of questions for a Service Desk Agent to derive which Configuration Item is disrupted and to determine Impact and Urgency for the customer.

Change	
CI Name (aff)	Confiuration Item (CI) which will be affected by this change. A change can be related to more than one Configuration Item.
CI Type (aff)	Every CI in the CMDB is related to a Type.
CI Subtype (aff)	Every CI in the CMDB is related to a Subtype, which is related to a CI Type.
Service Component WBS (aff)	Every CI in the CMDB is related to 1 Service Component, in order to identify which Product Manager is responsible for the CI. A Service Component is equal to a product in the Bill of Material and is part of one or more Services.
Change ID	The unique ID of a Change-record in HP Service Manager 9.
Change Type	In order to select and compare similar Changes in HP Service Manager 9, all records are categorized by Change Type.
Risk Assessment	Impact of change: Major Business Change, Business Change or Minor Change.
Emergency Change	Indication if the change is an emergency fix.
CAB-approval needed	Indication is the changes needs approval by the Change Advisory Board, before implementation.
Planned Start	Date and time the change implementation is planned to start.
Planned End	Date and time the change implementation is planned to end.
Scheduled Downtime Start	Date and time the Service Downtime is scheduled to start.
Scheduled Downtime End	Date and time the Service Downtime is scheduled to end.
Actual Start	Date and time the change implementation is actually started.
Actual End	Date and time the change implementation actually ended.
Requested End Date	Date and time before the change requestor wants the change to be implemented.
Change record Open Time	Date and time the Change record was opened in HP Service Manager 9.
Change record Close Time	Date and time the Change record is closed in HP Service Manager 9.
Originated from	Indication if the change originated from for instance Problem research, or is a quick fix for an Incident.
# Related Interactions	Number of Interactions caused by this change.
# Related Incidents	Number of Incidents caused by this change.

Fig. 1. The four tables with their fields and a description per field

implementation teams to continuously improve their Standard Operation Procedures.

The rest of this paper is structured as follows. The next section is about the preparation of the data that is needed for the analysis. Section 3 describes how a predictive model is derived and what the model looks like. The sub-questions are also answered in this section. In section 4 we have gathered all our other findings. Section 5 concludes the paper.

2 Data Preparation

To be able to create a predictive model the data that is available needs to be transformed into the required input format. This section describes how the supplied data is transformed.

2.1 Description of the Available Data

Four tables with data on interactions, incidents, incident activities and changes are supplied in comma separated value (csv) files. The data and the fields are described in Figure 1.

The data is imported into Microsoft SQL Server 2012, with each file creating a separate table. In the incidents csv empty records and columns were encountered which are all removed. Based on the tables the data model in Figure 2 is derived. The data for interactions, incidents and changes also have a direct relation to a Service Component (SC). In the data model this relation exists via Configuration Item (CI).

The relation between a CI and a SC is a many-to-one relation. In the data a CI can be linked to multiple SC's, leading to the conclusion that the relationship is actually many-to-many. A CI can be moved to an other SC, because of a

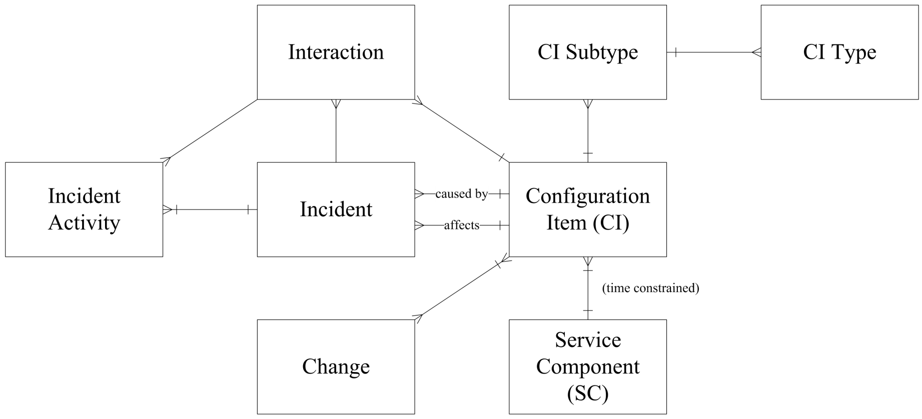

Fig. 2. Data model with all relevant objects and their relations

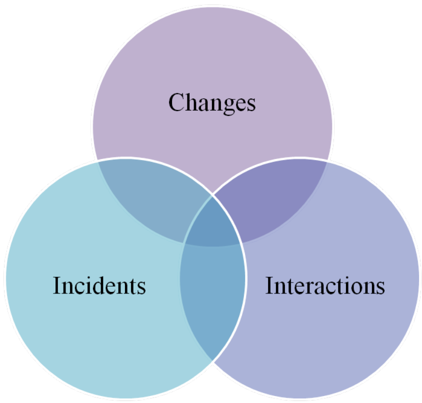

Fig. 3. Venn diagram showing the relation between changes, interactions and incidents

reorganization. Thus, at every moment in time a CI is always linked to precisely one SC. The many-to-one relation holds when time is taken into account.

2.2 Selecting the Data for Analysis

Not all of the available data is useful for our analysis. When a change is done for Configuration Items or Service Components that do not occur in the interactions or incidents datasets, then this change can not be used for building a model. Every change that is used in the model building phase must have at least one interaction or one incident. Figure 3 shows the different intersections of the changes, interactions and incidents that exist in the data. Only changes for CI's or SC's in the intersection with the interaction or the incidents table are used. The rest of the changes is discarded.

2.3 Data Related Decisions and Assumptions

To create an analysis dataset the following main decisions and assumptions were made.

- The smallest time period we use is a day. All interactions, incidents and changes that take place on a day are handled as if they occur in a single moment on that day.
- All weekends are removed. All interactions and incidents that take place in a weekend are moved to the Friday prior to the weekend. Interaction and incident hardly occur during a weekend because most of the Rabobank employees are not at work. When a change is implemented during the weekend then only from Monday the effects can be seen. So moving a change from a weekend day to the previous Friday can be done without losing any information.
- By removing all data from weekends we end up with time series without any artificial disturbance. All our time distance measures are presented in working days.
- For interactions the attribute *Open Time (FirstTouch)* was used as the moment the event occurs. For incidents this moment is based on the attribute *Open Time*. For changes the attribute *Actual End* is used. Since a change takes affect as soon as it's implementation ends. All these time related attributes are transformed into an uniform time format defined as *yyyy-mm-dd*.
- Sometimes the attribute *CI Name (aff)* has the value *#N/B*. This is interpreted as an empty value and the whole record is discarded.

3 Predictive Model

When a correlation exists between the implementation of a change and the workload in the Service Desk and/or IT Operations then it is possible to create a predictive model. This implication works also the other way around. If we can create a predictive model that accurately predicts the impact of a change implementation on the workload of the Service Desk and/or IT Operations then there is a correlation between them. To do this we need a dependent variable that expresses the relation between a change and the impact on the workload.

A change implementation affects one or more parts of one or more IT systems of the Rabobank. The smallest unit of a IT system is called a Configuration Item (CI). Examples of Configuration Items are server based applications, database software and a laptop. A Configuration Item is part of a set of Configuration Items which is called a Service Component (SC). Examples of Service Components are the Unix servers, the Storage Area Network (SAN) and desktop applications. A change may affect one or more Service Components. One of the questions that must be answered is at what level, CI or SC, the predictive model will be working.

3.1 Definitions

The first step towards a model is to define workload for the Service Desk and IT Operations. Workload for the Service Desk is defined as the number of closed interactions per day. For IT Operations workload is defined as the number of incidents that are registered on a day. Since incidents can be of a very different nature, the number of activities that are needed to solve an incident may vary widely. The definition of workload for IT Operations should also incorporate the number of activities that are needed to solve an incident. Unfortunately this is not known when an incident is registered. But if we know the average number of activities per incident and this number is stable, then we can use the number of incidents as an indication of the workload of IT Operations.

The next step is to define *impact on workload* of a change implementation. Since workload is defined as the volume of interactions for the Service Desk and the volume of incidents for IT Operations, when these volumes increase or decrease *after* the implementation of a change, then we assume that the increase or decrease is correlated with the implemented change.

On one day multiple changes may be implemented for the same CI and SC. We assume that all the changes on one day may have an effect on the workload starting not earlier then the day after the changes have been implemented. So we treat all the changes on one day as if only one change is implemented. We also assume that the effect on the workload of a change for a specific CI or SC is only shown by interactions and incidents also registered on that specific CI or SC. We make an exception for incidents because next to the CI on which the incident was registered, there is also a CI that caused the incident. When this caused by CI is present, then we replace the affected CI by the caused CI.

Based on these definitions and the observation that the level on which predictions should be made is yet to be chosen, we conclude that four models need to be build.

1. Predict workload of the Service Desk based on the number of interactions at the CI level
2. Predict workload of the Service Desk based on the number of interactions at the SC level
3. Predict workload of IT Operations based on the number of incidents at the CI level
4. Predict workload of IT Operations based on the number of incidents at the SC level

3.2 Building the Models

In this paragraph the dependent and independent variables are defined. Also the methods used to create the models are discussed.

The models we are building should make clear how an increase or decrease in the volume of interactions and incidents is related to a change implementation. In Figures 4 and 5 these volumes and the number of changes per day are plotted

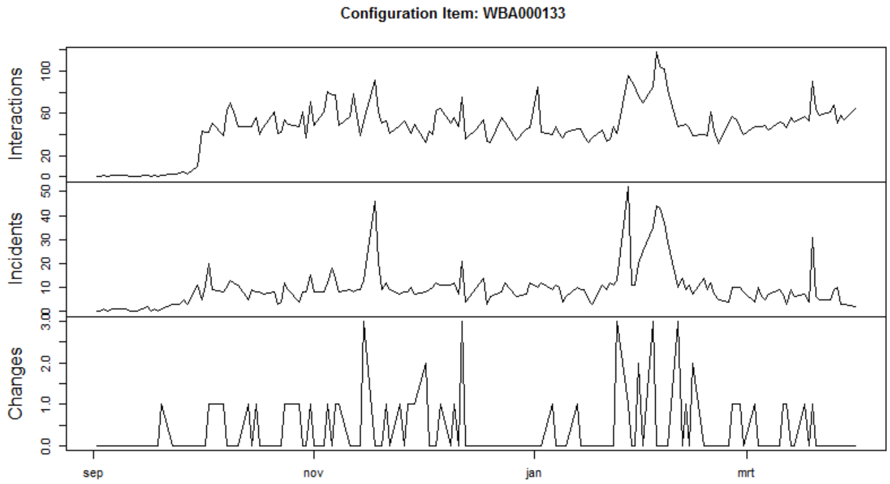

Fig. 4. Time series of the volume of interactions, incidents and changes for Configuration Item WBA000133

for a Configuration Item and a Service Component. Both figures show that a spike in the Changes time series sometimes coincide with spikes in the time series of the Interactions and the Incidents. It seems likely that the changes on that day are responsible for these spikes.

In general the time series show alternating periods with more and less variation and with a higher or lower average level. When a change implementation coincides with a change in the variation pattern or change in the level then there is a correlation between these two events. To find out which of the change implementations coincide with a change in the time series pattern, first the changes in the time series pattern need to be extracted. To achieve this we use the package *changepoint* [1] from R [2], a free software environment for statistical computing and graphics. With this package multiple change points can be detected within a time series using different techniques. We are interested in changes in variation and the mean of the time series. We assume that the arrival distribution of interactions and incidents follows the Possion distribution. Finally we chose to use the *PELT* method with the penalty parameter set to *Schwarz Information Criterion (SIC)*. Using these parameters in the function *multiple.meanvar.poisson* we determined the change points for the interaction and the incidents time series from the example. Figure 6 shows the change points and the periods where the algorithm determined that the variation and mean are alike.

Next we derive a new property for each change implementation on a day. We calculate the distance between the day of the change and the first change point that occurs after the change day. When the distance is at most 2 days, we

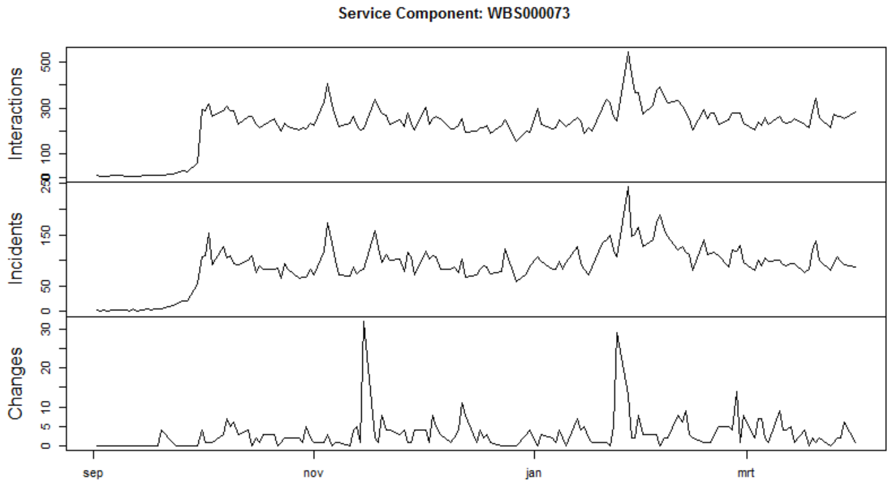

Fig. 5. Time series of the volume of interactions, incidents and changes for Service Component WBS000073

define that there exists a correlation between the change implementation and the change point. This property is called the Correlation Indicator. For each Configuration Item and for each Service Component the change points and this property are derived. We use the Correlation Indicator as the dependent variable in our model. We want to build a model that accurately predicts based on properties of the change implementation whether the Correlation Indicator is 0 or 1. If such a model can be build it is also possible to predict, based on properties of a proposed change implementation, whether this change implementation will have any effect on the volume of interactions and incidents. At this point it is not known what that impact will be, only that there is going to be some impact or not.

Since all changes that occur on one day for a CI or SC are grouped and treated as if one change happened, and these changes can be different it, is not possible to use the properties of each change as independent variables for our model. The properties of all the changes on a day need also to be grouped into one characteristics vector per change on a day. Tables 7 and 8 illustrate how the properties of multiple changes on a day are grouped into one characteristic vector per change on a day. Every unique value of a property becomes a variable. Then the number of times that a value occurs is counted and put into the column with the associated property value. The elements of this vector are used as independent variables.

The *dependent* variable is a binary variable. This makes it possible to use several model building techniques like Classification Tree, Logistic Regression,

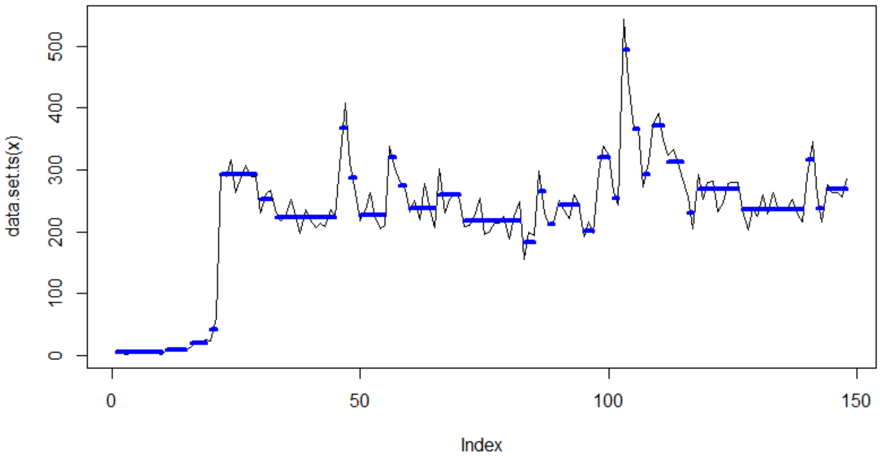

Fig. 6. Time series of the volume of interactions for Service Component WBS000073 including the change points

Service Component WBS aff	Day	Change ID	CI Name aff	CI Type aff	CI Subtype aff	Change Type	Risk Assessment	Emergency Change	CAB approval needed
WBS000073	10-10-2013	C00002829	SBA000360	application	Server Based Application	Standard Activity Type 02	Minor Change	N	N
WBS000073	10-10-2013	C00002866	SBA000360	application	Server Based Application	Standard Activity Type 03	Minor Change	N	N
WBS000073	10-10-2013	C00002868	SBA000360	application	Server Based Application	Standard Activity Type 02	Minor Change	N	N
WBS000073	10-10-2013	C00002868	SUB000458	subapplication	Web Based Application	Standard Activity Type 02	Minor Change	N	N
WBS000073	10-10-2013	C00002866	SUB000458	subapplication	Web Based Application	Standard Activity Type 03	Minor Change	N	N
WBS000073	10-10-2013	C00002829	SUB000458	subapplication	Web Based Application	Standard Activity Type 02	Minor Change	N	N

Fig. 7. Example records of changes that occur on one day for the Service Component WBS000073, before grouping

Random Forest and Boosting. We want to determine which of these techniques builds the best model. We select the best model based on the percentage of cases that are classified correctly. We use the R package *Rattle* to build all the models. All the techniques are part of this package.

Some extra attention is needed for the *False Negatives*. This is the situation where the model predicts that there is not going to be an impact, but there actually is an impact. In this case no adjustment to the planning would be done and there will be an impact on the workload. We would also would like to minimize this category.

For some of the CI and SC, where changes have been implemented for, only a small number of interactions and incidents are present in the data. Making a prediction for these CI and SC is not useful because the impact will be, based on historic behaviour, not very big. It is also a lot harder to fit a model on this scarce data with high variation. Based on Pareto charts a cut off is determined

Service Component WBS aff	Day	# Unique CI	# CIT application	# CIT sub application	# CIST SBA	# CIST WBA	# CT SAT	# Risk Minor	# Emergency Change	# CAB approval needed
WBS000073	10-10-2013	3	3	3	3	3	6	6	0	0

Fig. 8. Example records of changes that occur on one day for the Service Component WBS000073, after grouping. Only some of the columns are shown. Abbreviations are used to keep the column names compact: CIT=CI Type, CIST=CI SubType, SBA=Service Based Application, WBA=Web Based Application, SAT=Standard Activity 02 and Standard Activity 03

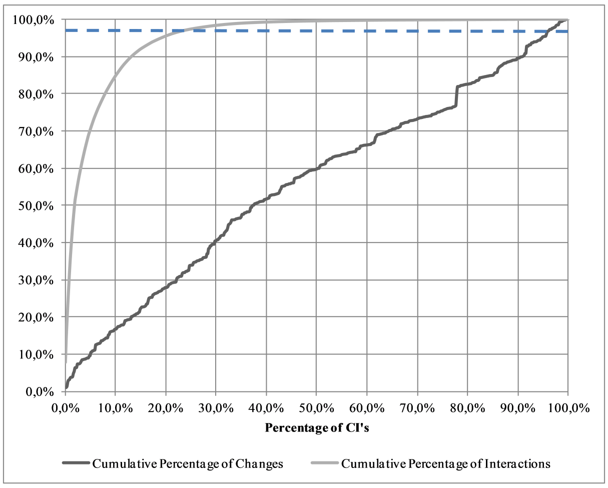

Fig. 9. Pareto chart for interactions on the CI level. The dashed line indicates the cut off at 98% of the total number of interactions

so 98% of all interactions and 98% of all incidents are captured by the model. The other CI and SC are removed from the dataset before the models are build. Figure 9 shows the Pareto chart for interactions on the CI level.

3.3 Training and Testing the Models

All the definitions and assumptions that are needed to create the models are in place. The first model build is to predict impact on the workload of the Service Desk where the workload is based on the number of interactions at the CI level. Table 10 shows how the different model building methods used have classified the cases.

The table shows that the Decision Tree model has the lowest overall error of 11%. The model will predict correctly for 89% of all the changes that affect a CI whether a change point should be expected in the number of interactions on that CI. When there is no change point in the test data (Actual=0) then the model predicts this with an error of 1%. When there is a change point in the test data (Actual=1) then the model predicts this with an error of 89%. The

Method	Predict=0 Actual=0	Predict=0 Actual=1	Predict=1 Actual=0	Predict=1 Actual=1	Total N Testset	Overall error	Actual=0 Error	Actual=1 Error	Predict=0 Error	Predict=1 Error
Decision Tree	392	49	4	6	451	11,8%	1%	89%	11%	40%
Ada Boost	392	53	4	2	451	12,6%	1%	96%	12%	67%
Random Forest	186	43	3	5	237	19,4%	2%	90%	19%	38%
Support Vector Machine	169	38	20	10	237	24,5%	11%	79%	18%	67%
Logistic Regression	189	46	0	2	237	19,4%	0%	96%	20%	0%

Fig. 10. Classification results for interactions on CI level

Method	Predict=0 Actual=0	Predict=0 Actual=1	Predict=1 Actual=0	Predict=1 Actual=1	Total N Testset	Overall error	Actual=0 Error	Actual=1 Error	Predict=0 Error	Predict=1 Error
Decision Tree	358	37	38	18	451	16,6%	10%	67%	9%	68%
Random Forest	167	36	27	11	241	26,1%	14%	77%	18%	71%

Fig. 11. Classification results for interactions on CI level with a balanced approach

overall error of 11% is a bit misleading because this is the weighted average of both errors. This result arises because the

This result arises because the ratio between changes with a change point and without a change point is not balanced. In the training data the percentage of changes correlated with a change point is 12%. With this kind of imbalance the model will fit the class with the highest occurrence better than the class with the lower occurrence. The solution to this problem is to add a prior to the dependent variable classes, a weight to the cases, oversampling of the minority class or down-sampling of the majority class. In this way the class with the lower occurrence is given a bigger influence in the model training phase, creating a model that better fits also this class.

Figure 11 shows the results of a balanced approach for the Decision Tree and the Random Forest methods for the same training data. This approach does not lead to a good model either because the overall error is higher and the error when predicting a change is also higher. Only the error when there is an actual change is lower. For the models on the SC level and for incidents we get the same kind of results. The confusion matrices are shown in figures 12, 13 and 14. Using the available independent variables and the selected model techniques we are not able to build a very accurate model. This does not mean that building a more accurate model is impossible.

In [3] an approach is discussed to deal with this for the Random Forest method.

3.4 Size and Duration of the Predicted Impact

When we are able to create a model to predict, based on the characteristics of a change, whether a change point will occur, the next thing we want to determine is what the impact of the change will look like. After each change it is possible that it takes some time for the organisation to adjust to the new situation. During this period more interactions and incidents may occur than before the change was implemented. The questions we want to answer are *What is the average*

Method	Predict=0 Actual=0	Predict=0 Actual=1	Predict=1 Actual=0	Predict=1 Actual=1	Total N Testset	Overall error	Actual=0 Error	Actual=1 Error	Predict=0 Error	Predict=1 Error
Decision Tree	555	26	14	6	601	6,7%	2%	81%	4%	70%
Ada Boost	563	30	6	2	601	6,0%	1%	94%	5%	75%
Random Forest	140	14	8	3	165	13,3%	5%	82%	9%	73%
Support Vector Machine	112	11	36	6	165	28,5%	24%	65%	9%	86%
Logistic Regression	148	17	0	0	165	10,3%	0%	100%	10%	-

Fig. 12. Classification results for incidents on CI level

Method	Predict=0 Actual=0	Predict=0 Actual=1	Predict=1 Actual=0	Predict=1 Actual=1	Total N Testset	Overall error	Actual=0 Error	Actual=1 Error	Predict=0 Error	Predict=1 Error
Decision Tree	730	122	14	17	883	15,4%	2%	88%	14%	45%
Ada Boost	734	127	10	12	883	15,5%	1%	91%	15%	45%
Random Forest	452	121	4	1	578	21,6%	1%	99%	21%	80%
Support Vector Machine	442	114	34	8	598	24,7%	7%	93%	21%	81%
Logistic Regression	446	116	10	6	578	21,8%	2%	95%	21%	63%

Fig. 13. Classification results for interactions on SC level

period to return to a steady state and *What is the average increase/decrease of Closed Interactions once a new steady state is reached?*

To answer these question first some definitions are needed. The average level of interactions or incidents prior to the change is the prior steady state. After the change implementation it may take some time for the average level of interactions and incidents to return to this prior steady state level or even better to become lower than this level. The duration of the impact is defined as the number of working days it takes to reach a level that is lower than the prior steady state level.

Since we use the *changepoint* method to determine the periods of the time series when the average and variation are alike, these stable periods can also be used to define the steady states that make up the time series, as shown in Figure 6. When we are looking for the first moment the level is lower than the prior steady state level, then we only need to find the first stable period with a level lower than the prior steady state level. The number of working days between the actual change implementation and the start of the first stable period with a lower level is defined as the duration of the impact of the change. The size of the impact is defined as the number of extra interactions and incidents that are registered on top of the prior steady state level, that is the expected number of interactions and incidents if nothing would have changed. If for example the steady state level is 10 interactions per day and after the change the average is 50 interactions per day, than the size is 50-10=40 interactions per day. These differences are summed for all days during the duration of the impact, thus leading to the total size of the impact of a change implementation.

For every change implementation that is linked to a change point the duration and the size are determined. When there is no lower level the duration is set to ∞. This means there is no point in the data where the level becomes lower than the prior steady state level. For the interactions time series on the CI level, there are in total 3869 changes present in the data. For 201 of these changes a change point can be found. In 64 of these changes correlated with a change

Method	Predict=0 Actual=0	Predict=0 Actual=1	Predict=1 Actual=0	Predict=1 Actual=1	Total N Testset	Overall error	Actual=0 Error	Actual=1 Error	Predict=0 Error	Predict=1 Error
Decision Tree	493	69	6	1	569	13,2%	1%	99%	12%	86%
Ada Boost	196	69	3	1	269	26,8%	2%	99%	26%	75%
Random Forest	316	58	2	0	376	16,0%	1%	100%	16%	100%
Support Vector Machine	304	57	14	1	376	18,9%	4%	98%	16%	93%
Logistic Regression	316	56	2	2	376	15,4%	1%	97%	15%	50%

Fig. 14. Classification results for incidents on SC level

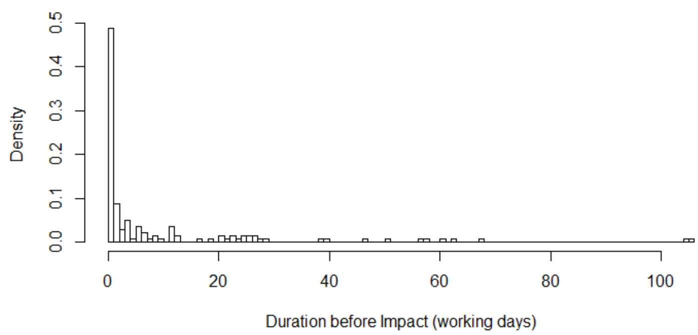

Fig. 15. Distribution of the duration until the level is lower than the prior steady level (Interactions on CI level. The change for SC=WBS000167,CI=WBA000008 on ChangeDate=2013-10-11 takes 106 workings and 154 extra interactions before an improved level is reached.)

point no improvement is present in the data. The final 137 changes lead to an improvement. The distributions of the number of days until an improved level is reached and the number of extra interactions that are encountered before the improved level is reached are shown in the figures 15 and 16.

Both figures show that the most common duration is zero workings days and zero extra interactions. This occurs for 50% of the changes with a change point and lead to an improvement. The other half of the changes also lead to an improvement, although not immediately after the change. The range of both the duration and the size of the impact is quite big. For duration the range is between 2 and 106 working days. The range for the size of the impact is between 4 and 3140 extra interactions. Taking a closer look to the distributions shown in figures 15 and 16, it becomes clear that there are approximately 20 outliers in both the impact duration distribution and impact size distribution. The changes that belong to the lather are the changes that influence the workload of the Service Desk the most.

With this approach similar results are reached for interactions on the SC level and incidents on both the CI and the SC level. The resulting distributions are shown in the Box Plots in figures 17, 18 and 19.

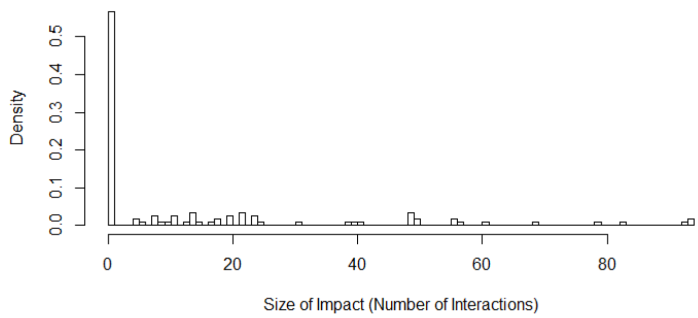

Fig. 16. Distribution of the size of the impact, the number of extra interactions compared to the situation that the prior steady level would have continued (Interactions on CI level). The change for SC=WBS000092, CI=DTA000616 on ChangeDate=2013-10-03 takes 63 working days and 3140 extra interactions before an improved level is reached.

(a) Box plot Impact Duration (b) Box plot Impact Size

Fig. 17. Box plots of incidents on CI level

The next step should be to try to build a model to predict which characteristics of a changes have a higher chance of turning into the types of outliers we just identified. To build such a model *anomaly detection* methods can be used, like discussed in [4].

3.5 Impact on Service Levels of Service Components

The final question is whether can be confirmed that project managers deliver the same or better service levels after each change implementation for all or many Service Components. The previous paragraph mentioned that some of the changes never lead to an improvement in the average level of interactions or incidents for a CI or a SC. Table 20 summarizes this for all categories.

(a) Box plot Impact Duration (b) Box plot Impact Size

Fig. 18. Box plots of interactions on SC level

(a) Box plot Impact Duration (b) Box plot Impact Size

Fig. 19. Box plots of incidents on SC level

This table is based on changes that are correlated with a change point. For the other changes no effect on the number of interactions or incidents could be found. Since these changes do not lead to a detectable effect, these may be discarded when determining the improvement in service level. The table shows that in general, when not looking at specific CI's or SC's, between 59% and 73% of the changes lead to an improvement.

Table 21 shows per category the number of unique items. For the category *CI Interactions* there are 79 unique CI's. For 22 of these CI's all the changes that are done for these CI's never lead to an improvement, for (also) 22 of the CI's more than 0% but at most 90% of the changes lead to an improvement and for 35 of these CI's more than 90% of their changes lead to an improvement. In short, changes for 27,8% of the CI's never lead to an improvement and more than 90% of the changes for 44,3% of the CI's lead to an improvement. So, 44,3% of the CI related managers do a good job improving the number of interactions. In the same manner can be concluded that 28,8% of the CI related managers do a good job improving the number of incidents. On the SC level, which is the main question, 37,4% of the SC managers are meeting their target for improving the number of interactions and 32,3% for improving the number of incidents. Again, with these percentages has to be taken into account that these are based

Category	# Changes	# Changes with Improvement	% Changes with Improvement
CI Interactions	201	137	68,2%
CI Incidents	140	83	59,3%
SC Interactions	472	346	73,3%
SC Incidents	472	310	65,7%

Fig. 20. The number of changes correlated with a change point and the number and percentage of these changes that lead to an improved level

Category	# Unique Items	# Items where 0% of Changes are Improvement	# Items where >0% and ≤90% of Changes are Improvement	# Items where >90% of Changes are Improvement	% Items where 0% of Changes are Improvement	% Items where >0% and ≤90% of Changes are Improvement	% Items where >90% of Changes are Improvement
CI Interactions	79	22	22	35	27,8%	27,8%	44,3%
CI Incidents	52	24	13	15	46,2%	25,0%	28,8%
SC Interactions	99	26	36	37	26,3%	36,4%	37,4%
SC Incidents	99	30	37	32	30,3%	37,4%	32,3%

Fig. 21. For every category the number of unique items. Per category the number of items that fall in the different improvement classes

on the number of changes that are correlated with a change point. The rest of the changes are discarded in this calculation.

4 Other Results

In this section we present miscellaneous findings that we gathered from investigating the available dataset. Because of the fragmented nature of these results they each are presented in their own subsection. There is no direct relation with the other results of this paper. We still would like to share them because they might give some interesting insights.

4.1 Incident Handling Paths

To understand the complexity of how incidents are being handled, the most common paths are analysed. Since a lot of incident activities occur on exactly the same time, and we do not have enough business understanding to make a selection, these activities are treated as an atomic event by concatenating their descriptions.

The 20188 of unique paths seem a relatively large number compared to the 46606 unique incidents that these paths are based on. A Pareto analysis shows that 10% of incidents are responsible for 60% of the paths. Because of the large number of unique paths this means that 60% of the incidents follow one of 2000 paths. The other 40% of the incidents have more than 18000 unique paths. We conclude that there not seem to be a standardized incident handling process. Table 22 shows the 20 most common paths.

Ranking	Path Frequency	First Incident Activity	Second Incident Activity	Third Incident Activity	Fourth Incident Activity
1	3319	Open	Caused By CI - Closed		
2	2503	Open	Assignment - Status Change	Caused By CI - Closed	
3	2010	Open	Assignment	Caused By CI - Closed	
4	927	Open	Assignment - Status Change	Mail to Customer	Closed
5	829	Open	Assignment - Status Change	Caused By CI - Closed - Quality Indicator Fixed	
6	557	Open	Assignment - Status Change	Caused By CI - Closed - Quality Indicator	
7	551	Open	Caused By CI - Closed - Quality Indicator Fixed		
8	473	Open	Assignment - Status Change	Mail to Customer	Caused By CI - Closed
9	433	Open	Assignment - Operator Update - Status Change	Caused By CI - Closed	
10	396	Open	Assignment	Caused By CI - Closed - Quality Indicator Fixed	
11	365	Open	Assignment - Status Change	Mail to Customer - Quality Indicator Set	Closed - Quality Indicator Fixed
12	296	Open	Reassignment - Update	Assignment	Caused By CI - Closed
13	293	Open	Assignment - External update - Status Change	External update	Closed
14	263	Open	Assignment - Pending vendor - Status Change	Closed	
15	206	Open	Caused By CI - Closed - Quality Indicator		
16	205	Open	External Vendor Assignment - Pending vendor - Status Change	Caused By CI - Closed	
17	201	Open	Assignment - Status Change	Closed - Quality Indicator	
18	191	Open	Assignment	Operator Update	Caused By CI - Closed
19	187	Open	Update	Closed	
20	170	Open	Assignment - Status Change	Operator Update	Caused By CI - Closed
>20	32231				
Total	**46606**				

Fig. 22. Top 20 unique incident handling paths

4.2 Incident Handling Teams

There are 242 teams active in solving incidents. Not every team plays an equal role in solving the incidents. About 20% of the teams perform more then 84% of all activities. The number of times a team change occurs is 106821 for all incidents. A team change is the handing over of an incident to an other team. The next incident activity is handled by a different team. Team 8 is the team that occurs the most in team changes. Team 8 is involved in almost 34% of all team changes. It could be worthwhile to investigate whether the number of times work is handed over can be reduced. Reducing the handing over of work will lead to a shorter throughput time.

Another observation we made is that the average amount of time an incident activity (events that happen on the same time are concatenated) takes is 1448 minutes. The average amount of handling time an incident takes is more than doubled when there is contact with customers (the incident activities *Communication with the customer* and *Update from customer* occur). For these incidents the average amount of handling time is 3705 minutes.

5 Conclusion

The questions raised by Rabobank Group ICT are questions that are relevant for most organizations with a Service Desk and an IT Operations department. It would be nice to be able to predict what the impact of a change will be on the workload of those departments. The capacity that is needed to keep up the required service levels can be planned better when the impact of changes is known in advance.

Creating an accurate model has proven to be difficult. Several techniques have been used and the Decision Tree gives the best results for all the categories. Whether the CI level or the SC level should be used to predict the number of interactions or the number of incidents, is not conclusively determined. Models can be build on both levels and the models on the CI level does not outperform the models on the SC level.

The next step is to determine impact patterns of changes. The duration and the size of the impact is calculated for the four categories. The variation in duration and size is large. Some changes lead to an immediate improvement, while others never lead to an improved level of interactions or incidents. The outliers have the biggest effect on the workload. With *Anomaly detection* methods these outliers can be predicted.

It seems that not all changes lead to improved service levels for CI's and SC's. Changes for only a small portion of the CI's and SC's lead to continual improvement of the service levels of these CI's and SC's. Service Managers do not seem to be delivering the constant improvement of service that is expected from them. Finally we have some recommendations for further research on this topic.

- When a change can be linked to a change point, then the impact of the change is sometimes mixed with the impact of a next change. After a change a new change may occur that can be more directly linked to a reaching a lower level. During the time that a lower level is not yet reached new changes may occur and be linked to a an other change point. It is not always clear whether the lower level is a direct result of the first change or a next change. In the approach we used we did not take this into account.
- For the next versions of the models it would be interesting to take the interference between the change for one CI on an other CI into account. Changing one CI might also effect an other CI.

References

1. Killick, R., Eckley, I.: changepoint: An r package for changepoint analysis. Journal of Statistical Software (To appear) (2013)
2. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. (2014)
3. Chen, C., Liaw, A., Breiman, L.: Using random forest to learn imbalanced data. Statistics Technical Reports (2004)
4. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. ACM Computing Surveys (CSUR) **41**(3) (2009)