Case Study

PROCESS MINING:
PACKAGE DELIVERY
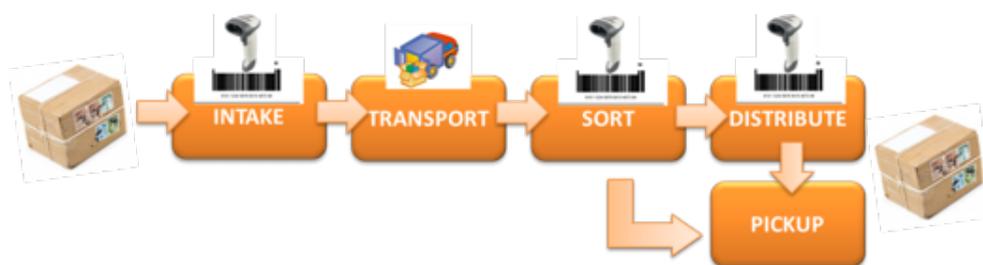
ae®

INTRODUCTION

A Package Delivery Company located in Belgium processes around 300,000 packages on a daily basis.

The package delivery process is fairly simple and consists of 4 basic steps as depicted in Figure 1. Each package is given a bar code which is scanned at different points throughout the process. The scanned barcodes, together with a large set of metadata (time and location of scanning, reason for scanning, etc.) is stored in a database and processed by a state machine. The processed information is subsequently used for reporting, tracking & tracing packages, invoicing, and handling of packages. In the future new scanning and tracking methods such as RFID (Radio frequency identification) and GPS should be added to the system.

The current setup is very complex. Over time, it has developed up to 7 different databases and a state machine configuration which has many relationships between different states, events and event reasons, some of which are not always clear.

This has led to a situation where the state machine has become a black box and nobody really understands its full functioning. In this context the link with the business process has become unclear and uncertainty about the correct configuration of the state machine has risen.

The goal of this analysis is to:
• Open the black box and gain insight into the functioning of the system.
• Evaluate the correctness of the configuration of the state machine.
• Investigate a number of abnormal flows that have been identified by business users.
• Link the different states and events back to the business process.

The results of the analysis should also help the company to take a decision with respect to future investments in the current system, i.e. within the context of the planned new functionalities (RFID & GPS).

THE EVENT LOG

For the analysis we created an event log that contained a limited subset of 18,000 cases (packages) starting on April 1 2012 with a maximum case duration of approximately 2 months. A number of meta data attributes were also included into the event log.

The following list gives a summary of the resulting event log:

• Period: April 1 2012 until June 1 2012.
• Number of cases: 18,000
• Number of events: 253,986
• Used attributes:
    » Bar code: unique identifier of a package (= case id)
    » State: the result of an event taking place (= activity)
    » Event: an event leads to a change in the state of the state machine
        (= activity for alternative view on the process)
    » Event reason: the reason for an event taking place
    » Direction: to sender or to addressee
    » Start & end time of an event: for every event two scans take place, one scan indicating the time
        the event started and one indicating the time the event ended
    » Product code: the service or product type the package belongs to
        (e.g. Express Delivery, Home Delivery, Delivery at Pick-up point, etc.)

Consider the following example of a Delivery event taking place. This event can result in one of two possible states: Delivery Successful or Delivery Unsuccessful. The possible event reasons for the latter state are Address Incorrect or Addressee not Home.

PROCESS MINING RESULTS

After some basic pre-processing of the event log, such as removing unmapped events and outliers, the following analyses were performed.

Basic Analysis

Basic statistics such as mean case duration and standard deviation and analysis of working time versus waiting time were calculated. Additionally different visualisations of the "process" were generated from different angles, such as state-based, event-based, event-reason-based, product-based (e.g. the process for Express Delivery packages and non-express delivery packages), etc. The basic analysis helped us to obtain a better understanding of the process and made it possible to make the right decisions regarding the setup of subsequent analyses.

Variant Analysis

The different variants of the process were analysed in detail. The first observation we made was the high variability in the process, i.e. a high number of variants, many of which had a very low frequency. Next the different variants were analysed in more detail to see if abnormal flows could be detected. In addition the most frequent variants were analysed in terms of their duration and a number of different criteria such as product type and country of origin (i.e. there are different steps for handling by customs for packages originating from outside Belgium).
The conclusion of the variant analysis was that there is a high variability in process execution, not all variants are "normal behavior" or some cannot even be explained. More analysis is needed.
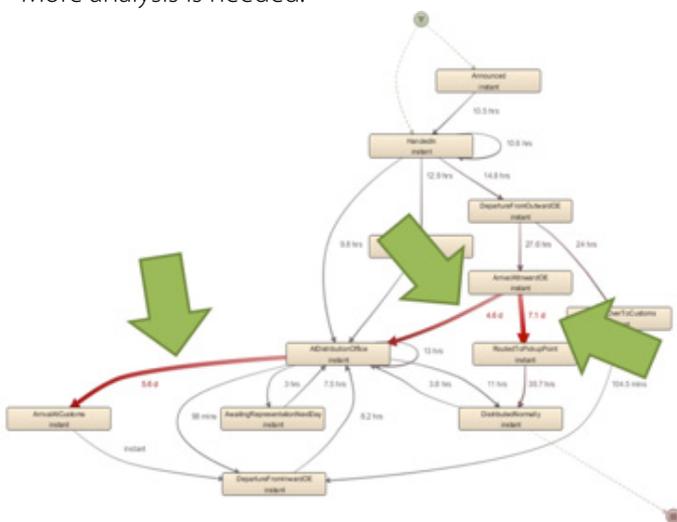
State – Event Analysis

This analysis was performed to evaluate the correctness of the configuration of the state machine. Filtering different combinations of states and events made it possible to analyse which events lead to which states and how the different states and events relate to each other. Consider the following example.
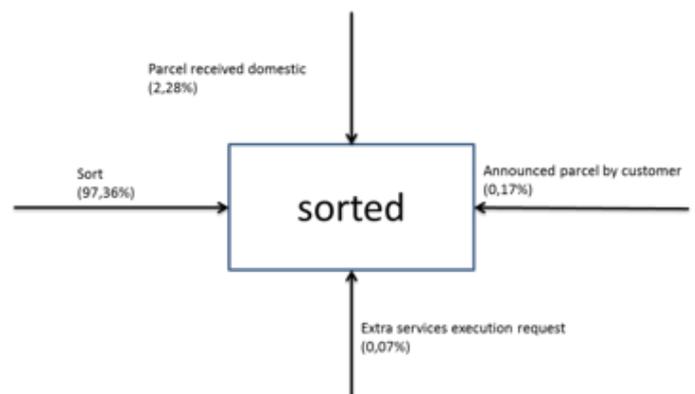
From Figure 2 it can be seen that in 97% of the cases the state 'sorted' is the result of the event 'Sort' (as can be expected). But there are 3 more events, with a much lower frequency, that can also lead to the same state 'sorted'. It was investigated which other states these 3 events may lead to, how these states relate to the state 'sorted' and in which cases and process variants they occur.
This enabled us to explain a number of the non-straightforward business rules that implicitly describe the relationships between states and events, although not all relationships could be explained and further analysis is needed.

More Analysis

We also performed a number of analyses to answer some specific questions. One example concerned service level agreements. E.g. special service levels apply to particular products, such as in the case of a Express Delivery product, to which a 24 hour delivery service level applies. A performance analysis was performed for these products, which are known to have problems with meeting the service levels, to investigate where most time was spent or "lost" during the process. An example of this can be seen in Figure 3.

**architects
for business
& ict**

ae®

ae nv / sa
Researchpark Haasrode 1250
Interleuvenlaan 27b
BE - 3001 Heverlee

info@ae.be
tel +32 16 39 30 60
fax +32 16 39 30 70

www.ae.be