An Analysis of Cargo Release Services with Process Mining: A Case Study in a Logistics Company

Sagit Kedem-Yemini  
Logistics Department  
Sapir Academic College  
Sderot, Israel  
sagity@mail.sapir.ac.il

Naor S. Mamon  
Logistics Department  
Sapir Academic College  
Sderot, Israel  
naormamon@gmail.com

Gal Mashiah  
Logistics Department  
Sapir Academic College  
Sderot, Israel  
gal_mashiah@walla.co.il

Abstract

Most logistics organizations have some kind of process-oriented information system that keeps track of business events. Process Mining uses event logs extracted from these systems in order to discover, analyze, diagnose and improve processes, organizational, social and data structures. While there has been much research on Process Mining in the last decade, the number of studies actually demonstrating the applicability of these techniques has been limited, particularly in the field of logistics. Consequently, there is a need for real-life case studies suggesting methodologies appropriate for Process Mining analysis, especially combined with IE&M well-established methods, and displaying the benefits of its application in real-life environments. In this paper we present a methodological framework for a multi-faceted analysis of real-life event logs based on Process Mining. As such, we demonstrate the usefulness and flexibility of Process Mining techniques combined with traditional IE&M methods in order to analyze and improve the organizational process in a case study that is centered on the actual cargo release process of a large Israeli logistics company. Our analysis shows that Process Mining techniques constitute an ideal means to tackle organizational challenges by suggesting process improvements and creating a companywide process awareness.

Keywords

Process Mining, Cargo Handling, Service Standards, Process Conformance, Logistics Operations
1. Introduction

During the last decades, information systems (IS) have developed from simple systems with limited functionality to complex, integrated architectures. As a result, it has become harder to understand and monitor how these systems impact the execution of everyday processes in organizations. Process Mining (W. M. P. van der Aalst, 2011) offers a solution based on the extraction, analysis, diagnosis and visualization of the data recorded by an IS during process execution. Although in the past, major contributions to the Process Mining literature were predominantly technical in nature, these techniques have proved their usefulness in practice as well. Nevertheless, application-oriented studies have only received modest attention (De Weerdt et al., 2013). This study demonstrates the benefits and challenges of applying Process Mining techniques by means of a multi-faceted analysis of business processes within the cargo release process of a major logistics company. Process Mining goes beyond the capabilities of traditional business intelligence tools (Golfarelli, Rizzi, and Cella, 2004) with respect to process analysis. Accordingly, it can be considered a proficient means of helping organizations understand their actual way of working, thereby serving as a foundation for process improvement. This is mainly due to the fact that the cornerstone of Process Mining is real data that demonstrates how business operations are actually carried out in an organization, an approach that is significantly different from other techniques of process improvement, such as interviews with key stakeholders. In order to examine the performance of this approach, we developed a methodological framework based on existing literature, describing the Process Mining study as implemented in a logistics services company. While this framework is similar to earlier research (Bozkaya, Gabriels, and Werf, 2009; Rebuge and Ferreira, 2012; De Weerdt et al., 2013), it is also unique in many ways, as it examines warehouse locations rather than activities. This enabled us to formulate different analyses and combine them with traditional IE&M methods. Consequently, this study clarifies the benefits as well as the challenges of conducting a real-life Process Mining study. In this paper, for example, we focus on service standards, substantiating that other data collection methods are unreliable.

In the following section, we will outline how the field has developed in the last decades and why the application of Process Mining techniques in services organizations faces distinctive challenges, as well as introduce a real-life case study. In section 3 we will elaborate on the research methodology. Section 4 is dedicated to the case study, in section 5 we report on the results. Finally, we present our conclusions and recommendations for future implementation.

2. Process Mining

As Process Mining is a relatively new discipline, some background is required to evaluate the research. Dumas et al. (2005) discuss one of the most influential trends of the past decades: the shift from data-orientation to process orientation. The changing context stimulated the progression of process-aware information systems (PAISs) which can be found along the entire value chain: ERP (Enterprise Resource Planning), WfM (Workflow Management), CRM (Customer Resource Management), case handling, B2B (business To business) and SCM (Supply Chain Management) systems.

Process Mining goes beyond the capabilities of traditional business intelligence tools (Golfarelli, Rizzi, and Cella, 2004) with respect to process analysis. Accordingly, as De Weerdt et al. (2013) claim, it can be considered as a proficient means of assisting organizations to understand their actual manner of working, thereby serving as a foundation for process improvement. This is mainly due to the fact that the cornerstone of Process Mining is real data that demonstrates how business operations are carried out in an organization de facto. This is significantly different from other approaches to process improvement, such as interviews with key stakeholders or artifact collection.

Process Mining (W. M. P. van der Aalst et al., 2007) is a relatively new discipline, based upon model-driven approaches and data mining. It proposes to provide methods, techniques and tools for the construction of models that adapt to concrete situations, examining system execution traces (i.e., logs). Although some Process Mining techniques have been proposed and a few tools are available, their usage still requires expertise in formal modeling and analysis. Therefore, they cannot be considered straightforward solutions.

2.1 Process Mining Tools

In the early nineties, with the emergence of management techniques such as process re-engineering, the modeling of processes began to gain prominence. Most of these systems keep track of the actual execution of the business processes by logging large amounts of data that form the input for process analysis techniques (De Weerdt et al., 2013).
Nowadays, Process Mining is being mentioned time and again in the literature as the most appropriate methodology for modeling processes based on the data recorded in an information system (W. M. P. van der Aalst, 2011). Process Mining is typically used to uncover the process model, determine its compliance with the standard model and identify the potential for process enhancement. In Table 1 we examine some of the current Process Mining tools which perform actual analysis.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Producer</th>
<th>Description</th>
<th>Reference</th>
<th>Web Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROM</td>
<td>Non-Commercial Tool</td>
<td>Aims to help practitioners as well as academicians to implement Process Mining techniques based on various algorithms such as genetic algorithm, heuristic miner, et cetera. This tool enables evaluators to demonstrate business processes and analyze the model in great detail. Requires expertise in the field.</td>
<td>(Verbeek H.M.W., Buijs J.C.A.M., van Dongen B.F., 2011)</td>
<td><a href="http://www.promtools.org">http://www.promtools.org</a></td>
</tr>
<tr>
<td>DISCO</td>
<td>Fluxicon</td>
<td>Technology that can automatically create smart flow diagrams of log files. High level of visualization, including process simulation. Filtering ability, variants analysis and statistics. A product that is suitable for the use of managers. Supports a well-developed academic initiative.</td>
<td>(W. van der. Aalst, 2016)</td>
<td><a href="http://www.fluxicon.com">www.fluxicon.com</a></td>
</tr>
<tr>
<td>Celonis Process Mining</td>
<td>Celonis GMBH</td>
<td>Claims to combine Process Mining with Machine Learning and Artificial Intelligence to achieve highly intelligent and fully automated insights from data logs.</td>
<td>(W. van der. Aalst, 2016)</td>
<td><a href="http://www.celonis.com">www.celonis.com</a></td>
</tr>
<tr>
<td>Minit 3</td>
<td>Gradient ECM</td>
<td>Automatically analyzes business processes and highlights paths and variants. It focuses on banks, insurance and manufacturing.</td>
<td></td>
<td><a href="http://www.minitlabs.com">www.minitlabs.com</a></td>
</tr>
<tr>
<td>Lana</td>
<td>Lana Labs GmbH</td>
<td>Process Mining automated, pulse visualization. Open API, ability to export process to dashboards or R analysis.</td>
<td></td>
<td><a href="http://www.lanalabs.com">www.lanalabs.com</a></td>
</tr>
</tbody>
</table>

### 2.2 Challenges on Performing Process Mining

In order to actually utilize all the advantages of Process Mining, efforts have been made to create project methodologies that are tailored toward supporting Process Mining projects, as methodologies such as CRISP-DM and SEMMA are very high-level and provide little guidance for Process Mining specific activities (W. M. P. van der Aalst, 2011). In the past, two well-known Process Mining methodologies prevailed: Process Diagnostics Method (PDM) (Bozkaya, Gabriels, and Werf, 2009), which has also been adapted for healthcare environments (Rebuge and Ferreira, 2012), and the L* life-cycle model (W. M. P. van der Aalst, 2011). PDM is designed to quickly provide a broad overview of a process, while L* covers many different aspects of Process Mining and focuses on broader topics, including process improvement and operational support. Unfortunately, these methodologies are not suitable for every project (van Eck et al., 2015). Some of their main problems include a limited scope of PDM, which covers only a small number of Process Mining techniques and emphasizes avoiding the use of domain knowledge during the analysis (Bozkaya, Gabriels, and Werf, 2009), thus deeming it less applicable for larger, more complex projects (Suriadi et al., 2013). L* covers more techniques, but was primarily designed for the analysis of processes and aims at discovering a single integrated process model. Neither L* nor PDM explicitly encourage iterative analysis, which proved vital for both our own case study as well as the case study performed by Suriadi et al. (2013). Van Eck et al. (2015) presented
a new methodology, known as PM2: Process Mining Project Methodology. PM2 is designed to support projects targeted at improving process performance or compliance with rules and regulations. It covers a wide range of Process Mining and other analysis techniques, and is suitable for the analysis of both structured and unstructured processes. The research presented in this paper utilizes the PM2 methodology and tools.

3. Research Methodology

In order to maximize the benefits generated by the research analysis, our team performed data collection through traditional methods (interviews, observations and document collection), followed by the collection of data from the WMS system. The data was analyzed first by forecasts, process simulation, Lean and Pareto analysis, and then shifted to Process Mining by using the PM2 framework. This methodology consists of six stages: (1) planning, (2) extraction, during which initial research questions are defined and event data are extracted. After the first two stages, one or more analysis iterations are performed. Each analysis iteration executes the following stages one or more times: (3) data processing, (4) mining & analysis, and (5) evaluation. An analysis iteration focuses on answering a specific research question by applying Process Mining related activities and evaluating the discovered process models and other findings. Such an iteration may take anywhere from minutes to days to complete, depending mainly on the complexity of the mining and analysis. If the findings are satisfactory, they can be used for (6) process improvement and support. The tool selected for the Process Mining was DISCO (vendor Fluxicon), as Sapir Academic College belongs to its academic initiative and is a tool adapted for managers. Therefore, all the illustrations presented in this paper were generated by DISCO. Additional explorations were done with PROM.

Three types of Process Mining methods were employed (W. van der. Aalst, 2016):

1. Discovery – This technique takes an event log and produces a process model without using any apriori information.
2. Conformance - The existing process model described in interviews was compared to the event log of the same process. Conformance checking was used to check if reality, as recorded in the log, conforms to the model.
3. Enhancement - Intended to generate added value to Maman's Group by changing or extending the a-priori model, demonstrating the process' improvements.

4. Case Study

As the purpose of this paper is to demonstrate the usefulness of Process Mining analyses in practice, it is best demonstrated by a case study in the logistics services industry. This industry is of main interest for Process Mining since it includes many computed business processes which are suitable for event log analysis, yet not reported as worthwhile. The case at hand involves a large Israeli Logistics Company, the Maman Group, which uses the WMS system to provide services.

4.1 Maman Group

Maman Cargo Terminals & Handling Ltd. is the leading provider of logistics services in Israel, offering comprehensive services to government bodies and the foremost companies in the market. Maman's cargo operations, based at the cargo terminals located at Ben Gurion International Airport, provide a full range of cargo handling services for all international air cargo imported or exported from Israel. Although as of 2008 the Maman terminal is no longer the sole service provider in this field, it handles a cargo capacity of up to 300,000 tons per year, which is about 56% of all the cargo that entered Israel via air in 2017. Macro-economic data for 2018 indicates an increase of 8% in the amount of air cargo in Israel. The terminal is open around the clock, seven days a week. It employs about 450 people in three daily shifts, and serves all airlines and courier companies operating at BGIA, all in a one-stop shop including the airlines' representative offices, customs, customs agents, cargo agents, transport companies, economic and relevant governmental bodies. The terminal is also comprised of various types of storage options that meet the needs of the incoming and outgoing cargo, a crane system, pallet systems, and free shelving.

Our research focused on the cargo release process, where three central entities are involved in the work process: the customer, customer services, and the cargo release department. The process begins with the arrival of the customer at
the customer service center with a batch form and a release form (Get Pass). Once the forms are approved by customer services and scanning barcodes, the customer receives an operational permit to enter the premises with the truck and approach a ramp. The customer service clerk determines the ramp according to the type of cargo to be released. When the truck reaches the release ramp, the customer hands the forms to the release officer, who sends the forklift operators to pick up the cargo. When the batch and all cargoes have been collected, the release officer checks the cargo and scans its barcodes, and the forklift operators move the cargo to the appropriate truck.

4.2 Research Objectives

Our research goal was to discover and validate the actual process model of cargo release and check if the company performed according to the service standard of 90 minutes, as promised to customers. If not, one of the objectives was to suggest methods that would enable the Maman company to achieve this goal. In order to examine this, the research team planned to utilize combined methodology – traditional IE&M methods and new data science methods, particularly Process Mining.

5. Results

5.1 Traditional Methods

In order to determine the levels of effort (LOE) for the next year, the team first employed observations and forecasting techniques, which enabled us to predict next year’s demands. The results presented here are based upon 228 observations ($\alpha=0.05$) and upon all the data collected in 2017 (January to October) in the WMS system. For the task of gathering the prediction data for Jan-Oct 2017, the team used three different forecast methods (simple average, moving average with $k=4$ and exponential smoothing), which were compared with 4 measures (NAPE, MSE, MAD, ME) as per Nahmias (2009). The results are presented in Figure 1, and Table 2 shows the error measure values. According to the results, the best method to predict cargo releases was a simple average, and the forecast is for more than 10,000 cargo releases per month for 2018.

Figure 1. Forecasting demands – anticipated cargo releases.
5.2 Data Log

Creating a real model from a log data file using Process Mining requires a focus on the desired processes and the ability to filter "noise" activities. Our research focused on main release activities, and therefore the log file was filtered by main locations, main customs codes and all carriers but two (that had a different release agreement). This process excluded activities that affect performance indicators but are irrelevant to the research (such as handling corpses, different types of hazards, food etc.).

Figure 2 shows some statistics on the main log after filtering, summarizing the events overtime (based on the filtered log). It reflects the main issue-related cargo handling release time – a mean duration of 10.8 days and a median duration of 2 hours (compared to the 90-minute service standard) and reflects 72% of the original log cases. After converting the log to a process map we can present a fuzzy logic diagram (W. van der. Aalst, 2016) of the cargo release process.

<table>
<thead>
<tr>
<th>ME</th>
<th>MAD</th>
<th>MSE</th>
<th>NAPE</th>
<th>Prediction Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.357</td>
<td>1.579</td>
<td>2,825,893</td>
<td>14.53</td>
<td>Moving average $K=4$</td>
</tr>
<tr>
<td>-0.2</td>
<td>987</td>
<td>1,448,862</td>
<td>6.63</td>
<td>Simple average method</td>
</tr>
<tr>
<td>1,137</td>
<td>1,137</td>
<td>1,574,708</td>
<td>11.94</td>
<td>Exponential smoothing</td>
</tr>
</tbody>
</table>

5.3 Process Discovery

Regular Process Mining data load relates to "case ID" versus "activity". Usually, the "case ID" when applying this technique would be a form, such as a purchase order or an insurance policy (Mahendrawathi, Astuti, and Nastiti, 2015; Rebuge and Ferreira, 2012). DISCO utilizes Activity-Based Decomposition (W. van der. Aalst, 2016), and due to the research team's definition of warehouse location as "activity" and the "case ID" as the Bill of Lading, the process discovery map is quite different than usual. These definitions allow the process to begin when the customer arrives to pick up the goods, creating a flow in the process by moving the goods in the warehouse connected to the same Bill of Lading. Therefore, the representation in Figure 3 is innovative, as it shows a physical map of movements inside the warehouse in the form of a 'Spaghetti Diagram' created by the research team with DISCO. While a client receives a...
Bill of Lading that is comprised of many lines, which are in different locations in the terminal warehouse, the following diagram is presented in low resolution detailing in order to make it simple to understand. It shows clearly that the Automatic Storage Area is the most active area, where most of the warehouse locations exist.

It should be clarified here that Process Mining utilizes all the available data rather than just sample data. Figure 3 is based on over 205,000 records, with frequencies. This accomplished the first goal – determining the actual process as per the real data log.

Figure 3. Handling cargo layout with frequencies as built from the filtered log file.
5.4 Conformance

Conformance checking compared the release cargo process as described in interviews to the actual process as found in Process Mining. It revealed new facts, such as rework and attributes affecting it. Figures 4 summarize this. We discovered that the actual process is somewhat more complicated than expected by the managers, that cargo is moved around the crane area frequently and that some locations act as bottlenecks in the release process.

![Figure 4. Handling cargo release layout with mean time, based on the filtered log file.](image)

5.5 Enhancement

For the third goal, enhancement, we performed 3 steps – analysis, improvements and evaluation. First, for the analysis, we focused on the cases that took up to 9 hours, as other cases were defined as exceptions. This enabled the research team to perform statistics on over 120,000 events that are 45% cases of original log file. In figure 4 (mean time) the darker red indicates the locations in the warehouse that consumed more time for the release process to be implemented. These were identified as 'problematic locations', including, for example, the free storage area. Similar analysis was performed on median times. Figure 5 shows the map created from this filtered log file with mean time. The total release time at each location is indicated by the darkness of the item's background color.

The analytics on the data log file revealed valuable information regarding the work processes of the Maman group. Based on these analytics, the second step was improvements. Process research showed a map with the actual release process, including all the paths of the process, defined as the variants of the process. We found 4,594 variants of the original process, with 8 of the variants representing 67% of the cases. Analysis of the frequent variants identified 4 problematic locations, where the cargo handling release process lasted more than 90 minutes by median time. We found the release handling time for each of the 23 locations in the warehouse, and produced statistics for each agent and carrier (including rework, idle time and bottlenecks). This analysis of the handling process revealed that preplanning while unloading the planes and storing according to the Bill of Lading would dramatically improve the release time. In order to complete the enhancement goal, our team combined other methods, in particular 5S methodology and jobs scheduling (during the unloading process), so that single Bill of Lading locations could be improved, consequently reducing the time consumed by the cargo release process. The consequences of this improvement were estimated by using the Arena simulation model, presented in Figure 6. The model was run twice, once without changes (mean time 267 minutes to release, median time 112 minutes) and then with the suggested enhancement (mean time of 79 minutes).
Figure 5. Handling cargo release layout median time, focusing on the "free storage" area.

Figure 6. Arena simulation model for cargo release time after improvement estimation.
6. Conclusions

As the results of this research produced many conclusions, they will be divided into two main types: specific and general. Among the specific conclusions, four will be discussed following.

First, a major gap was discovered between the actual process and the described process, especially obvious in the time analysis. We found that most of the customers did not receive their cargo after 90 minutes. In fact, 64% of the customers received their goods after 4.5 hours (!). Based on these findings, we suggested that Maman change their storage procedures in order to store according to the Bill of Lading, as our team recorded major delays on long releases, resulting from the current location of different stations. We believe that the release time may be cut by approximately half.

Our second recommendation also relates to warehouse location. We recommend that 3 remote storage areas (such as Pharmaceuticals) be shifted to areas closer to the loading area. Analysis of the distances and the impact of the distance on release time resulted in a faster release time for each location.

The Third conclusion reached by our team was based on the fact that many releases required rework. For example, shipments were moved from one crane area to another 4,974 times. This usually meant that the cargo was in the wrong location, which caused both a delay in release time and a waste of resources and funds. We reviewed the implications of the relatively large amount of required rework with the Maman management.

Our fourth conclusion is related to quality. This analysis displayed the company exceptions in quality (error) as prohibited cargo mobility that occurred last year. Moving cargo to improper storage stations can cause quality issues. For example, putting pharmaceutical cargo on an export platform results in cooling issues that can damage the cargo, thus exposing the company to various risks. This issue was presented to the management so that they may improve procedures and avoid mishaps.

As for general conclusions, we would like to emphasize the fact that we performed Process Mining on a field which has yet to be examined, Logistics Services. By using existing tools, designed for other purposes, we succeeded in presenting a visual representation of the activities taking place in the warehouse, and combined traditional IE&M methods with methods new to data science, in order to create a holistic picture. We also encountered various dilemmas during our research, which will not be discussed here.

7. Discussion

This paper presents the application of Process Mining in a new field, cargo release processes. We propose an applicable framework that reveals actual processes, stressing conformance and possible quality issues that the company can improve upon. The framework of our research can be implemented regardless of the information systems in use, as it relies on actual data and statistics and produces clear maps of actual processes, including all their variants. We combined well known IE&M techniques (such as observations, forecasting, 5S and simulation) with a new data science technique (Process Mining) and used a variety of computerized tools (Arena, Disco, Excel, WMS System) in order to achieve the research goals. The Maman Group embraced the results and recommendations and immediately began to change their company’s work process both for received packages and cargo release.

Acknowledgements

The authors would like to thank Yuval Ben Abu and Ygal Shem-Tov for their contribution.

References


Biographies

**Sagit Kedem-Yemini** is an Industrial Engineer, proficient in information systems and currently holding two lecturing positions: a tenured lecturer position at Sapir Academic College (Logistics Department) and an adjunct lecturer at Ben Gurion University (both in IE&M and Faculty of Business and Management). Her teaching portfolio is broad, focusing on Enterprise Systems implementation (SAP and Oracle Applications) and derivatives of ERP data collection – from Business Analytics to Process Mining. Additionally, she has extensive experience in academic curriculum development, is head of her department’s teaching committee and serves as liaison to the graduation projects unit. Her research interests include Process Mining and its practical applications, ERP relates issues and DSS development. Since 2015 serves as member of JITCAR Editorial Review Board.

Prior her to academic career, Dr. Kedem-Yemini worked at a global Clean-Room Fab Build-Up Construction Management Company with major clients (such as Intel, Tower Semiconductors, and Teva Pharmaceuticals), where she held various positions, including Logistics Manager, Scheduling Manager and CIO (Chief Information Officer).

**Naor S. Mamon** is currently a senior in the Logistics Department at Sapir Academic College. This study was part of his final project for a Bachelor of Arts degree, and this is his first academic paper. His professional experience is in the areas of operations and logistics, and for several years he was responsible for the Operational Teams at The Cargo Terminal in Ben Gurion Airport.

**Gal Mashiah** is currently a senior in the Logistics Department in Sapir Academic College. This study was part of his final project for a Bachelor of Arts degree, and this is his first academic paper. His professional experience is in the areas of operations and logistics. Gal works in an international shipping company.