Customer Journey Analytics: A Model for Creating Diagnostic Insights with Process Mining

Daan Weijs and Emiel Caron

Tilburg University, Department of Management, Warandelaan 2, PO Box 90153, Tilburg, The Netherlands

Keywords: Customer Journey, Touchpoints, Customer Journey Analytics, Data Models, Process Mining.

Abstract: The customer journey is becoming more complex due to digitization of business processes, broadening the gap between the proposed journey and the journey that is actually experienced by customers. Customer Journey Analytics (CJA) aims to detect and analyse pain points in the journey in order to improve the customer experience. This study proposes an extended version of the Customer Journey Mapping (CJM) model, to measure the impact of different types of touchpoints along the customer journey on customer experience, and to apply process mining to gain more insight in the gap between proposed and actual journeys. Moreover, this model is used to develop dedicated CJA based on process mining techniques. A case study on e-commerce applies the CJM-model in practice and shows how the combination of process mining techniques can answer the analysis questions that arise in customer journey management.

1 INTRODUCTION

The digitization of businesses increases the complexity of the customer journey. Customers now interact with a firm trough many touchpoints and they expect a seamless experience throughout their journey with the firm. Companies try to manage the customer journey to create an optimal experience, but in practice there are many deviations between the proposed journey and the journey that is actually experienced by customers. The large amounts of data that are being generated at touchpoints open up opportunities to find out how to provide more value to customers. This study aims to find out how to analyse and gain better insights into the impact of different types of touchpoints along the customer journey on customer experience, by producing *Customer Journey Analytics* (CJA) derived from process mining.

The *customer journey* encompasses all inter-actions a customer has with a firm during the end-to-end purchase process (Lemon and Verhoef, 2016). Customer experience is built over an extended period of time and includes many touchpoints between a customer and an organization (Zomerdijk and Voss, 2011). A *touchpoint* can be defined as "an instance of communication between a customer and a service provider" (Halvorsrud et al., 2016). All touchpoints

together constitute the customer journey (Zomerdijk and Voss, 2011) and companies carefully design their journeys to create an optimal experience. The journey has a static state which reflects the hypothetical journey of the service delivery process, i.e., the *expected journey*, and a dynamic state that represents the actual execution of the process (Halvorsrud et al., 2016), i.e., the *actual journey*. The shift from offline to online commerce increases the complexity of marketing communications, broadening the gap between the static and the dynamic state. This results in new types of management questions about the dynamic state of the journey, that support marketing decision-making. These questions revolve around determining the effectiveness and the impact of different marketing communications on business outcomes, such as revenue or customer experience.

The data that is generated at online touchpoints open up opportunities for the visual representation of the customer journey, i.e., maps, and applying analytics (Holmlund et al., 2020; Lemon and Verhoef, 2016), which helps to answer the questions that arise in customer journey management. Figure 1 presents an overview of the elements of an online customer journey. A firm creates different types of advertising touchpoints to attract visitors to its website. A website typically has a funnel structure that supports visitors with relevant information at every stage of the customer journey, moving from
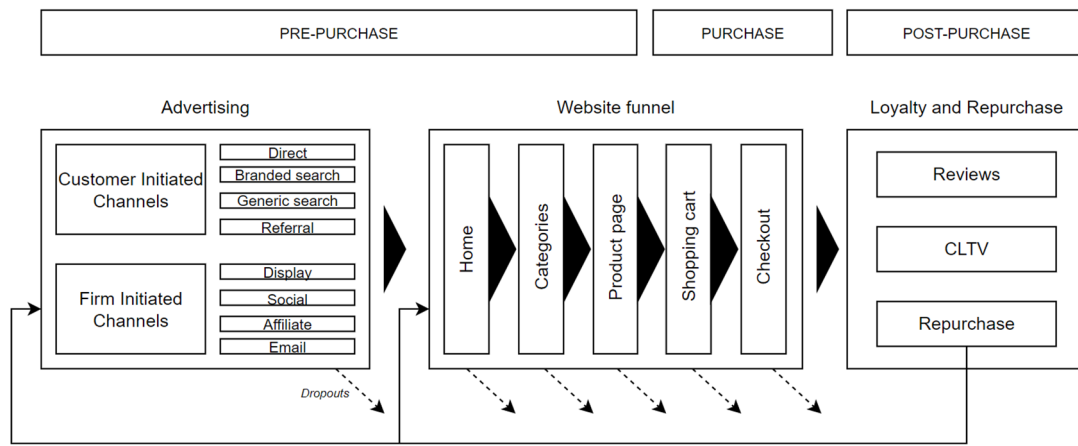
Figure 1: Elements of an Online Customer Journey, adapted from Anderl et al., 2016; de Haan et al., 2016; Lemon and Verhoef, 2016.
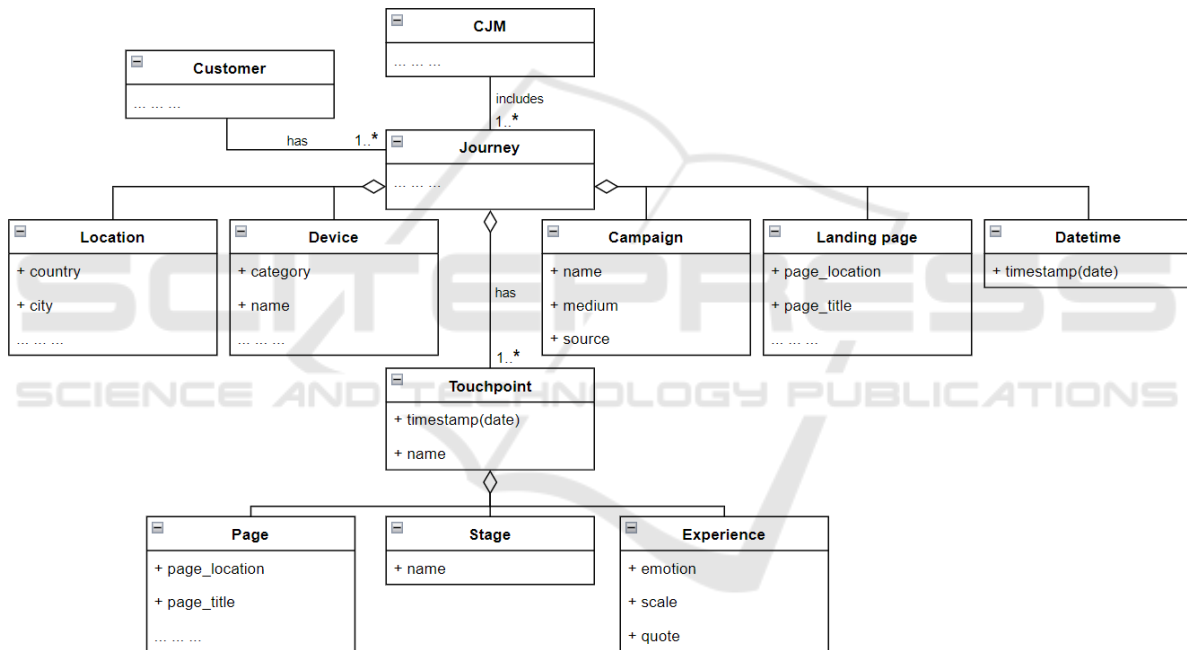

Figure 2: Components of the online Customer Journey Mapping (CJM) model, adapted from Bernard and Andritsos (2017).

general interest to more specific consideration of the actual purchase or download. The post-purchase phase includes reviews, customer lifetime value and repurchase rates. Every action that a visitor performs on the website can be logged. In addition, a website registers how the user landed on the website. Data generated at the advertising and website touchpoints is a rich source for analytical purposes. The goal of CJA is to enable marketers to detect and analyse pain points, allowing them to intervene and develop possible actions to improve customer experience.

The objective of this research is threefold: (1) to extend the *Customer Journey Mapping* (CJM) *model*

of Bernard and Andritsos (2017), to measure the impact of different types of touchpoints along the customer journey on customer experience, (2) to develop dedicated CJA based on specific process mining techniques, and (3) demonstrate the results of (1) and (2) on a real case study in e-commerce.

This paper is structured as follows. In Section 2 we propose the extended CJM-model and connect CJA with process mining techniques. After that, we apply the model in a case study to produce CJA for an e-commerce company. Finally, conclusions are drawn in Section 4.

2 PROCESS MINING FOR CJA

2.1 Extended CJM-model

Process mining aims to discover, monitor and improve business processes (Aalst, 2012). It can be used in many different contexts and settings where processes are executed. One of the main contributions of process mining is that it uses real event data that is generated during processes, closing the gap between assumed process models and the actual execution of the process. Bernard and Andritsos (2017) proposed a CJM-model to store CJMs as XML structures and they illustrated how process mining can be used to analyse CJMs. While the original model provides a structured starting point for mapping, the advertising elements and the website funnel structure (see Figure 1) have to be included as model components to create a more comprehensive overview of an online customer journey. In addition, a further mapping between process mining techniques and dedicated managerial CJA questions is required to discover the applicability of the techniques within this domain.

Figure 2 presents the complete CJM-model for online customer journeys in terms of an UML (Unified Modeling Language) class diagram. The CJM-model consists of multiple customer journeys. Each Journey is performed by a Customer, and a customer can perform one or many journeys. A customer journey might include six dimensions, captured by the aggregations and components of the Journey class. A Journey is started from a physical Location (1), performed on a specific Device (2), is started from a specific channel, i.e., Campaign, as defined by Anderl et al. (2016) (3), has a Landing page (4), and a Date (time) dimension (5), that captures when the journey started. In addition, it consists of multiple Touchpoints (6) that each have a timestamp and a descriptive name, a Page dimension where the touchpoint is encountered, a Stage (i.e., pre-purchase, purchase and post-purchase phases), as defined by Lemon and Verhoef (2016), and an Experience that can be measured with an emotion, scale or quote (Bernard and Andritsos, 2017). Here the model of Bernard and Adritsos (2017) is extended by adding the Journey – Touchpoint hierarchy (dimension 6) and dimensions 1 – 4.

The hierarchy of the online CJM-model is reflected in the import data to make meaningful analysis with process mining software. The requirements of an event log - a case id, timestamp and event - are met in the Touchpoint class. However, a single flattened event log, i.e., fact table, including all touchpoints is not enough to capture the structure

of the customer journey. The dimensions that characterize the journey are important and need to be included in the data model. This allows for slicing and dicing the data from multiple perspectives and enables marketers to detect and analyse pain points, for example, by comparing journeys from different campaigns or journeys that started from desktop versus mobile.

2.2 Mapping Process Mining with CJA

The CJM-model in Figure 2 shows that the customer journey and process mining can directly be related in terms of the required data structure. Moreover, we map classes of process mining techniques to specific CJA analysis questions. The techniques applied in process mining should be able to provide meaningful insights that managers can use to manage the customer journey. The goal of CJA is to enable marketers to detect and analyse pain points and opportunities in the customer journey in order to develop possible optimization efforts. CJA should focus on the gap between the proposed journey and the actual journey as experienced by the customer. Process mining includes three main classes of analysis: discovery, conformance checking and enhancement (Aalst, 2012). *Process discovery* uses event log data to create a process model of the actual execution of the process. *Conformance Checking* compares the proposed model with the actual model to check the process conformance. *Enhancement* aims to change or extend the process model, based on the insights from the other techniques. In addition to the types of analysis, process mining covers different perspectives (Aalst, 2016). The *control-flow* perspective focuses on the ordering of activities, i.e., the control flow. The *organizational* perspective focuses on the resources, i.e., actors, people, roles, departments, that are involved in the process and how their tasks are related. The *case* perspective includes the properties of the cases. It focuses on the individual characteristics of cases beyond the activities or resources. Finally, the *time* perspective focuses on the timing and frequency of events.

Process mining techniques have shown to be effective in analysing gaps between proposed and actual processes by discovering the actual process model based on event logs (Aalst, 2016). The objective of process mining is to discover, control and improve actual processes. These objectives fit the types of business questions that arise in customer journey management, where process discovery can be used to discover the actual customer journeys, conformance checking to analyse the gap between the

proposed and the actual journey, and enhancement could focus on new events and attributes to create richer analysis. In Table 1, CJA is aligned with process mining analysis. The table presents an overview of which mining techniques and perspectives are used to answer the different types of analysis questions that arise in customer journey

Table 1: Mapping process mining with CJA.

PM Analysis	Perspective	CJA questions
Discovery	Control-flow	What is the most common path on the website?
	Case	What are the average events per case for different journeys?
	Time	How quick do visitors drop out in the path to purchase? (Lemon and Verhoef, 2016)
Conformance	Control-flow	Where do visitors deviate from the customer journey map?
	Control-flow	Where in the purchase process do visitors drop out? (Lemon and Verhoef, 2016)
	Case	Was my marketing campaign successful?
	Time	Which types of blog content are engaging?
Enhancement	Control-flow	Which new events can be included to create richer analysis?
	Case	Can we cluster visitor behaviour to identify customer segments? (Lemon and Verhoef, 2016)
	Time	How does journey duration deviate for different product segments?

management. The alignment is based on careful review of the literature on process mining and CJA (Aalst, 2012; Aalst, 2016; Lemon and Verhoef, 2016) and validated by case studies (see e.g., Section 3). In this respect, relevant mining algorithms for CJA are: Directly-Follows Graphs, Heuristics miner, and Fuzzy miner, that take into account the frequency of events, and where the output of the algorithm is transformed into dedicated CJA.

2.3 Customer Journey Data

The website is the central pillar in online customer journeys (see Figure 1). Google Analytics[1] (GA) is a popular technology for tracking and analytics of websites. GA uses first-party cookies to track individual users as they navigate through a website. The latest version, GA4, tracks user behaviour in an event structure. Every action that a user performs on a website – e.g., clicking on a link, reading a blog page, or playing a video – is called a 'hit' and triggers an event, generating huge amounts of data. All user behaviour on a website is hierarchically ordered in three classes: A *user* has one or more website *sessions* that each contain one or more *events*. The class structure of GA4 corresponds with the extended CJM-model as explained in Section 2.1. A user corresponds to the customer, a session corresponds to a journey including its dimensions and an event corresponds to a touchpoint. The data, generated by GA4 can be mapped directly to the CJM-model of Figure 2. The end result of this mapping is an actual customer journey that can be visualized with process mining techniques.

3 CASE STUDY IN E-COMMERCE

The case study is conducted for the ABC company[2], an e-commerce company that sells products via their own website and via third-party e-tailers. On average their website receives more than 150 thousand visitors per month. Their goal is to guide visitors through the journey to the product page where they can buy the product or click through to a third-party e-tailer (i.e., lead generation). The ABC-company distinguishes touch-points in five categories related to their website funnel: homepage, blog-post, category-page, product-page, and checkout. An underlying data model and (import) data set have been created

---

[1] https://analytics.google.com

[2] The actual name of the company is not disclosed. https://www.celonis.com
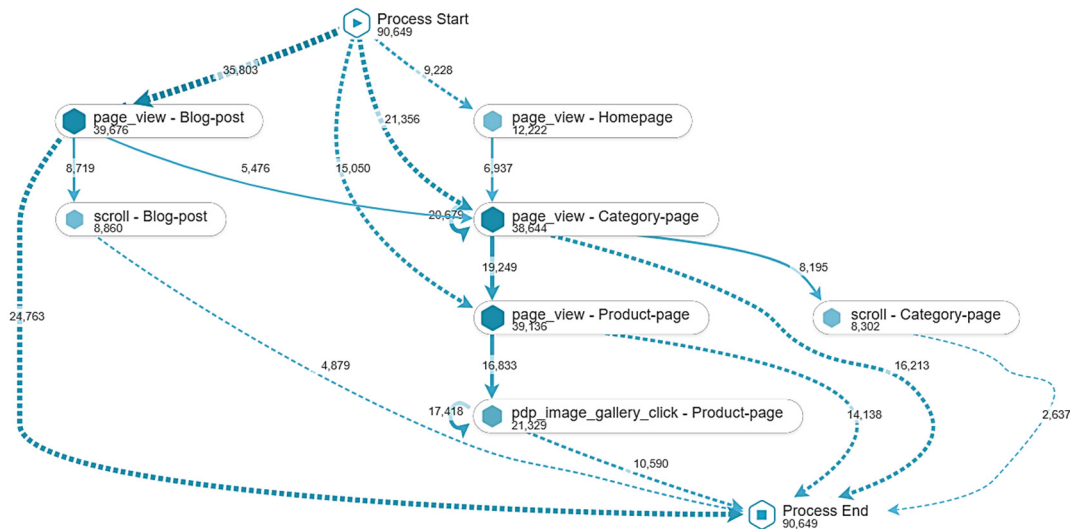
Figure 3: Discovered actual customer journeys for the ABC-company.

based on the CJM-model (see Figure 2) and applied in the case study. The data is loaded into Celonis process mining software and subsequently a part of the CJA questions from Table 1 are used to test and validate its usability with the software.

Figure 3 shows the discovered process model with 'fuzzy mining' and provides insights in the flow of visitors along the touchpoints in the customer journey. Most of the journeys start on a blog-post and the least start on the homepage. More than 50% of the journeys end directly after viewing a blog-post. Process discovery with a *control-flow* perspective shows the most common path on the website. The *time* perspective can be used to find out how quick visitors drop out in the process. Process discovery techniques help analyse the gap between the static and the dynamic state of the customer journey.

Conformance checking and process analytics techniques measure the performance of different customer journeys. Table 2 shows the *journey conformance* and *purchase probability* of four expected journeys, based on the website funnel depicted in Figure 1. Company ABC proposed four types of expected journeys, that can start at either one of the pages. Journey conformance uses the *control-flow* perspective and measures the percentage of journeys that conform to the proposed flow. It answers the questions where visitors deviate from the proposed journey and where in the purchase process visitors drop out, by showing the actual journeys that visitors have. The purchase probability measures the probability that a visitor makes a purchase, given that he followed the proposed journey, using conditional probability $P(A|B)$. These insights help a company to measure the impact of different types of touchpoints

along the customer journey towards customer experience. Table 2 shows the biggest gap between the expected and the actual journey for the Blog-Category-Product journey. Only 10.31% conforms to the proposed journey, meaning that about 90% follows a different path. This indicates some bottlenecks in the journey and since the purchase probability is relatively high, it is valuable to find areas of improvement. By diving deeper into these journeys and slicing the data on the 5 perspectives of a journey, the root causes behind the pain points can be identified. Two insights were a shifted ratio of mobile/desktop traffic for these journeys (i.e., the Device dimension) and a paid search campaign that led visitors to a wrong blog category (i.e., Campaign dimension).

Table 2: Conformance Checking and Process Analytics.

Expected Journey	Journey conformance	Purchase probability
Homepage → Category → Product	46.97%	2.34%
Blog → Category → Product	10.31%	2.44%
Category → Product	46.28%	2.99%
Product	-	1.39%

The product page contains five types of touchpoints that visitors can encounter. In addition to the product information, these pages include elements such as a video player or a button where visitors can

download a file. The interactions on these elements are measured by custom events. Process analytics provides insights in what behaviour leads to a purchase. The conditional probability P(A|B) of each of these touchpoints has been calculated towards a purchase or a lead generation. In this case, A is the purchase and B is the journey that the visitor followed. The impact on purchase, measures the probability that the visitor makes a purchase, given that (s)he followed the proposed journey. Table 3 shows the results. For example, for the review impression the probability of making a purchase is 2.86% given the fact that the visitor has checked the reviews of the product. The same calculation is done for lead generation, where the probability of clicking through to a third-party e-tailer is 13.59%. Both measures indicate the effect of a touchpoint. Table 3 shows the values for all journeys, but these values can be filtered on different product segments, campaigns or mobile versus desktop journeys to gain more insight into purchase behaviour under different circumstances.

Table 3: Process Analytics on the product page.

Touchpoint	Purchase probability	Lead generation probability
View image gallery	2.39%	7.05%
Select product options	3.81%	11.80%
Check the reviews	2.86%	13.59%
Play the product video	3.52%	10.41%
Download a file	1.49%	5.16%

Based on the findings of the process mining analysis, some suggestions for improvement can be made, for example, changing the blog categories for paid search campaigns and optimizing the website for mobile devices. The extended CJM-model, with the Journey – Touchpoint hierarchy and additional dimensions, allows for these analyses. The insights provided result from slicing the data on location, device, campaign and landing page characteristics. Due to space limitations, only brief examples are presented here. The case study shows that the combination of process mining techniques can answer the questions that arise in customer journey management. Moreover, another case study has been conducted in a business-to-business environment, showing similar results. We are now working on

extending the model's analytical capabilities to ensure its generalizability to other web applications.

4 CONCLUSIONS

This paper contributes to the CJA and process mining research by: (1) extending the CJM-model to measure the impact of different types of touchpoints on customer experience, by adding the Journey – Touchpoint hierarchy with several analytical dimensions that allow for a richer analysis of the customer journey, (2) developing dedicated CJA based on process mining, and (3) demonstrating the results in an e-commerce case study. It further closes the gap between CJA and the application of process mining techniques. The case study applies the extended CJM-model in practice and shows how the different process mining techniques can be used to create insights in the customer journey. The ability to slice the data from multiple perspectives allows for new insights that were not possible with the model of Bernard and Andritsos (2017). In addition, it shows the application of the theoretical model in practice, closing the gap between academia and managers. Further research could focus on applying the model in a more extensive case study to validate if the improvement efforts that result from the process mining analysis indeed help a firm to improve customer experience. In addition, a more complete evaluation in comparison with existing CJA techniques is envisaged.

REFERENCES

Aalst, W. V. D. (2012). Process mining. *Communications of the ACM*, 55:76–83.

Aalst, W. V. D. (2016). Process mining: Data science in action. Springer Berlin Heidelberg, 2 edition.

Anderl, E., Schumann, J. H., and Kunz, W. (2016). Helping firms reduce complexity in multichannel online data: A new taxonomy-based approach for customer journeys. *Journal of Retailing*, 92:185–203.

Bernard, G. and Andritsos, P. (2017). A process mining based model for customer journey mapping.

de Haan, E., Wiesel, T., and Pauwels, K. (2016). The effectiveness of different forms of online advertising for purchase conversion in a multiple-channel attribution framework. *International Journal of Research in Marketing*, 33:491–507.

Halvorsrud, R., Kvale, K., and Følstad, A. (2016). Improving service quality through customer journey analysis. *Journal of Service Theory and Practice*, 26:840–867.

Holmlund, M., Vaerenbergh, Y. V., Ciuchita, R., Ravald, A., Sarantopoulos, P., Ordenes, F. V., and Zaki, M. (2020). Customer experience management in the age of big data analytics: A strategic framework. *Journal of Business Research*, 116:356–365.

Lemon, K. N. and Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80:69–96.

Zomerdijk, L. G. and Voss, C. A. (2011). Nsd processes and practices in experiential services. *Product Development & Management Association*, 28:63–80.