Multichannel Marketing Attribution Using Markov Chains for E-Commerce

Lukáš Kakalejčík [1] | *Technical University of Košice, Košice, Slovakia*
Jozef Bucko [2] | *Technical University of Košice, Košice, Slovakia*
Paulo A. A. Resende [3] | *University of Brasília, Brasília, Brasil*

Abstract

There are plenty of online media that can be used by ecommerce companies in order to drive the revenue. However, use of this media is usually connected to investment into the selected media. From the company perspective, it is wise to evaluate the outcome of these investments in order to choose the best media mix possible. As customers do not usually buy during their first website visit, it is important to monitor their customer journey and assess the value to particular interactions. The objective of this paper is to analyze the data of selected companies using the Markov chains. The data about online customer journeys were analyzed. We found that the Markov model decreases the credit assigned to the channels favored by last-touch heuristic models and assigns more credit to the channels favored by first-touch or linear heuristic models. By using Markov order estimator GDL (Global Dependency Level), we also found that 4^{th} and 5^{th} order was the most suitable.

Keywords

Attribution modeling, multichannel attribution, e-commerce, marketing, order of Markov chain

JEL code

M31, M21, L1, C10

INTRODUCTION

In collective sports, such as football, trainers usually do not rely on single player to win the game. Moreover, the player himself is usually not ready to take a ball and score a goal without the help of any of his teammates. In marketing, it is very similar. Managers usually do not rely on a single marketing channel to deliver outcome they desire. Successful companies (and football teams, too) used to use several marketing channels (players) to deliver. The goal of the team managers is usually evaluating the most useful players; however, most of them understand that scoring a goal is not the only attribute that matters. Some players usually need to acquire a ball and create opportunity for other players to score.

The same applies in marketing. Some marketing channels are great to acquire a potential customers for the first time, while others are great at closing the deal. Approximately 96% of website visitors

[1] Faculty of Economics, Technical University of Košice, Němcovej 32, 040 01 Košice, Slovakia.
[2] Faculty of Economics, Technical University of Košice, Němcovej 32, 040 01 Košice, Slovakia. Corresponding author: e-mail: jozef.bucko@tuke.sk, phone: (+421)556023269, ORCID ID: 0000-0002-0791-1179.
[3] Department of Computer Science, University of Brasília, Brasília, Brasil.

are not ready to purchase a product online during their first website visit (Bulygo, 2012). On the contrary, since the first visit toward conversion (purchase), the visitors move through the process called the buyer journey. This process represents the sequence of the steps taken by customers during move through the phases of awareness, decision-making, and purchase (Roberge, 2015). The modeling of the buyer journey consists of mapping the customer's interaction with the brand aiming to improve these interactions. This process should result in an increase in sales and customer satisfaction (Wang et al., 2015). Through the progress of digital advertising and technological innovations, companies can track digital "footprints" of the customers on a granular level, bringing the knowledge about customer's behavior and measure the impact of displaying the particular marketing channels to the customers on conversions Ghose and Todri (2015), and Smarandache and Vladutescu (2014). Many researchers such as Peterson (2005), Constantin (2014), or Massara, Liu, Melara (2010) have tried to model consumers' behavior to predict their response. Attribution modeling can be considered another point of view on this particular topic.

Companies do not usually rely solely on a single marketing channel to acquire customers. As was mentioned, several marketing channels are used while working in cohesion to accomplish the company's goals. The value of importance should be assigned to each of these channels. Attribution modeling is a set of rules based on which the credit for conversion or purchase is assigned to the particular marketing channels (Clifton, 2015; Shao and Li, 2011). In research (Ferencová et al., 2015), were defined as a problem connected to the evaluation of the utility of marketing channels in the sales cycle. Despite of executed surveys for customers, it is often difficult to determine the channels they interacted with along their journeys. This issue can be solved by using attribution models where each customer touchpoint with the company can be evaluated. Szulc (2013) and Sterne (2017) claim that the use of attribution modeling helps optimize the allocation of marketing budget, support marketing budgeting, ensure more precise planning of marketing campaigns, ensure the accuracy of cost-per-acquisition calculation and help optimize payments to affiliate partners.

In currently available web analytics tools (such as Google Analytics), there are several heuristic models implemented to determine the merits of each marketing channels, for example in (Clifton, 2015; Kaushik, 2011; Shao and Li, 2011):

- last touch model (100% of the credit is assigned to the channel before the conversion),
- first touch model (100% of the credit is assigned to the first channel that customer got in interaction with),
- linear model (an equal amount of credit is assigned over all the channels that customer interacted with during the journey),
- time-decay model (the highest value is assigned to the last channel or campaign, and the assigned value decreases towards the first channel),
- position-based model (40% is assigned to the first and the last interaction, the rest of the credit is distributed evenly across the remaining channels),
- custom model (analyst itself assigns the value to the channels based on his own set of rules).

In Anderl et al. (2014; 2016), Barajas et al. (2016), and Bryl (2016) were reported that the use of heuristic attribution models is not proper for attribution purposes. Barajas (2016) claim that heuristic models assign a value to each displayed and converting channel, however, they ignore hypothetical reaction without a user being in touch with the advertisement. They stated that heuristic models are not data-driven. Anderl (2014) discusses that despite heuristic models are not accurate, the use of more sophisticated attribution approaches found its place in managerial practice. Bryl (2016) claims that heuristic models are not proper for the channel attribution because of their poor quantity, while their selection requires a managerial decision to choose the right one that will be suitable for the company's data.

There have been several studies that offered more data-driven approaches to the attribution to overcome the weaknesses of heuristic models. Yadagiri et al. (2015), Nissar and Yeung (2015) use Shapley value[4] in their non-parametric approach to attribution as a game theory-based model. In his thesis (Rentola, 2014), Rentola used two models: binary logistic regression to classify customers to converters and non-converters (purchasers/non-purchasers), as well as a logistic regression model with bootstrap aggregation. On the other hand, Shao and Li (2011) used bagged logistic regression and a probabilistic model in their study. In their study, (Li and Kannan, 2014) used a hierarchical Bayesian model. Geyik et al. (2015) developed their attribution algorithm MTA (Multi-Touch Attribution) to solve two problems: spending capability calculation for a sub-campaign and return-on-investment calculation for a sub-campaign (more in (Geyik et al., 2015). On the contrary, Wooff and Anderson (2015) offer an attribution mechanism based on the appropriate time-weighting of clicks using the sequential analysis. Hidden Markov Model was used in the studies conducted by Abhishek et al. (2012), and Wang et al. (2015).

We can see many approaches to the attribution, however, we incline to a Markov chain model proposed by Anderl et al. (2014; 2016) discussed in the following parts of our study. Anderl et al. (2014) and the following study by Anderl et al. (2016) use a higher order of the Markov chains model to attribute the value of the marketing channels. They propose that for practical reasons, the 3rd order is the most proficient when calculating the outcome of particular marketing channels. There was further reported that the Markov model meets the following criteria: objectivity, predictive accuracy, robustness, interpretability, versatility, and algorithmic efficiency. Anderl et al. (2016) also stated that heuristic models undervalue display advertising and pay-per-click campaigns, social media, and e-mail activities. On the other hand, Markov chains distribute the value of the channel more evenly. Based on these criteria, we selected Markov chains to be a suitable method for the analysis of our study. We also prefer this method because of the following reasons:

- The export of customer journeys consisting of marketing channels customers used to come to the website before the purchase is among the standard features of Google Analytics that is the most used web analytics tool (Clifton, 2015). This ensures our analysis might be executed broadly by a company of any size and budget.
- Attribution analysis using Markov chains can be easily executed in software The R Project (2019) with couple lines of code using the package ChannelAttribution (Altomare, 2016) and allows users to compare the results in the standard heuristic models. The data exported from Google Analytics almost exactly suit the structure supported by this package.

Based on the abovementioned claims, we choose Markov chains to be a suitable model for our analysis and therefore will be discussed in detail in the forthcoming section.

1 MARKOV CHAIN AND ITS USE FOR ATTRIBUTION MODELING

Formally, a sequence of random variables $\{X_t\}_{t=1}^{\infty}, X_t \in S := \{s_1, \ldots, s_m\}$, is a Markov chain of order r if, for all $(a_1, \ldots, a_{t+1}) \in S^{t+1}$, where S is a set of possible states of random variables of X_t. $P(X_{t+1} = a_{t+1} | X_1 = a_1, \ldots, X_t = a_t) = P(X_{t+1} = a_{t+1} | X_1 = a_1, \ldots, X_{t-r+1} = a_{t-r+1})$ and $r<t$ is the smallest integer to satisfy it. Essentially, this represents that the probabilities related to X_{t+1} depend only on the last r events, for all t.

In this context, S is referred by the state space, a particular sequence $(a_1, a_2, \ldots) \in S^{\infty}$ is called by a trajectory, the size of S is the length of state-space or number of states, represented by m, and the probabilities of $X_{t+1} = a_{t+1}$ considering that $(X_{t-r+1}, \ldots, X_t) = (a_{t-r+1}, \ldots, a_t)$ are called the transition probabilities represented by the notation[5] $p(a_{t+1} | a_{t-r+1}, \ldots, a_t) := P(X_{t+1} = a_{t+1} | X_{t-r+1} = a_{t-r+1}, \ldots, X_t = a_t)$.

[4] The Shapley value is a concept from coalitional game theory. Shapley values are used in cooperative games to fairly distribute the "payout" among the features (Molnar, 2020).

[5] Here, we consider that the Markov chain is stationary, i.e., the transition probabilities do not depend on t.

A particular state b is absorbing if the probabilities to leave the state are “0”, i.e., $p\left(c|a_{t-r+1},\ldots,b\right)=0,\forall c\neq b$ and, consequently, $p\left(b|a_{t-r+1},\ldots,b\right)=1$.

A Markov chain can be represented by an initial probability distribution for the first r steps and the m^{r+1} transition probabilities. When $r = 1$, it is possible to have a graphic representation for the Markov chain. For more details about Markov chains, we recommend Karlin and Taylor (1975).

Anderl (2014) proposes the use of Markov chains on channel attributions, considering the state space S as the states “Start” and “Conversion” combined with the set of marketing channels. In this case, the process $\{X_t\}$ represents the possible customer journeys through these channels. They suggest using a removal effect for attribution modeling. The removal effect is defined as the probability to achieve the conversion from the “*Start*” state if some of the states (s_i) are removed from the model. As the removal effect reflects the change in conversion rate if the given state s_i is removed, the value (or importance) of the given marketing channel can be determined. If N conversions are generated without the particular channel (compared to the number of conversions in the full model), the removed channel determines the change in the total number of conversions (Bryl, 2016). The Markov chain described in this section defines the methodical framework used in our analysis conducted in the following parts of the study.

2 OBJECTIVES AND METHODS

The main objective of this paper is to define the current state of multichannel attribution and, based on the literature, to analyze the customer journey data of selected companies by using the Markov chains. The main objective is decomposed into two partial objectives:

- to determine the current state of use of attribution modeling; to analyze the multichannel paths of a selected companies with the use of Markov chains,
- to propose the most appropriate order of the Markov chain in terms of predictive accuracy and computing efficiency.

The studies by Anderl et al. (2014; 2016) demonstrated that using the Markov chain to analyze customer journeys, it is appropriate to use its third order. In line with these studies, it could be assumed that by the research of the benefit of marketing channels used by e-commerce on the Slovak market, it is appropriate to use the Markov chain of the third order. Accordingly, we formulate the following hypothesis:

H1: We assume that the use of the third-order Markov chain is appropriate in the attribution modeling of the outcomes of marketing channels used by e-commerce stores operating on the Slovak market.

Approximately 20 companies operating on the Slovak market were approached the goal, while for our study, companies selling their products online through an e-commerce store were selected. From the list of the companies, four of them agreed to take part in the study, provided that their business name remains

Table 1 Share of positive answers to job search questions and item-response probabilities

	Company 1	Company 2	Company 3	Company 4
Scope of activity	Distribution of industrial electronic components for industrial production	Retail sale of sporting goods of a wide range	Retail sale of sporting goods with the focus on running and triathlon	Retail sale of food and nutritional supplements
Revenues for 2016 in ths. €	15 561	16 018	308	4 993
Number of employees	50–99	200–249	3–4	20–24
Tracking period of customer journeys	1.4.2016–31.8.2016	1.7.2016–30.6.2017	1.7.2016–30.6.2017	4.12.2016–4.12.2017

Source: Finstat and our own processing anonymous for work purposes. Table 1, however, shows the basic characteristics of the companies based on data from Finstat (2017) (website providing financial information about Slovak companies), which will give the reader an idea of the nature of the business focus and its size.

The data about the customer journeys of e-commerce customers of the analyzed companies were obtained from the analytics platform Google Analytics that businesses use to measure the performance of their websites (e-commerce). The data from the most common conversion journeys (customer journeys) were analyzed using heuristic models and the Markov chain. The data will be analyzed using The R Project for Statistical Computing (2019), using the web analytics platform Google Analytics (2018), and MS Excel from the Microsoft Office Suite (Microsoft, 2016).

3 DATA

The attribution modeling data used in our work were collected from four e-shops (businesses), one of which is focused on the sale of electronic components, two of which are focused on the sale of sportswear, and the last one is focused on the sale of nutritional supplements. For ease of use, we have chosen to export the customer journeys that ended up as a purchase, based on the visits sources from *Top Conversion Paths* report available in the analytics software Google Analytics, while within the individual customer journeys the following sources of visits (marketing channels) could occur:

- **Direct visit:** represents a situation where a user enters the URL of a web page directly into the browser window or visits a web page using a saved bookmark;[6]
- **Organic visit:** represents a situation where a user enters a key phrase in the search engine (Google, Bing, Yahoo, and others), and then clicks on search results to go to the business website;
- **Referral source:** represents a user visit by clicking on a link on another website (the Social network visits are usually not included);
- **Social networks:** represents a user visit by clicking a link on social networks (Facebook, Twitter, LinkedIn, and others);
- **Email:** represents a user visit by clicking on the link in an e-mail delivered to his mailbox;
- **Paid Search:** represents a user visit by clicking on paid search results (e.g., on the Google AdWords platform);
- **Display advertising:** represents a user visit by clicking an ad banner placed, for example, on the Google Display Network;
- **Other:** represents a user visit from a source that was not mentioned above;
- **N/A:** represents a user visit from a source that the analytical system for some reason has failed to identify.

An example of a customer journey looks as follows:

Paid > Direct > Social > Direct > Direct.

Such a customer journey tells that a user visited the website for the first time through a paid search, then came directly to the page by typing the URL into the browser, later clicked through social networks, and purchased on the website after two more visits when he directly entered the URL into a search engine. No higher data granularity was chosen for the comparability of the results with already existing studies. This means that Facebook, Instagram, or Twitter, for example, are considered as one source of visits marked as *Social networks*. Similarly, this applies to the *Banner ad* channel, where the ad platform or placement of the banner ad is ignored.

---

6 Visitors from mobile apps or offline advertising sources (TVs, billboards, flyers, etc.) may also be considered as direct visits for inappropriately selected or implemented visitor source tracking.

The use of the Markov model allows analysis of the customer journeys that have not ended up with purchases. However, Google Analytics cannot track them without any further technical implementation. As our goal is to provide a platform for performing attribution analysis that will be available to a wide range of businesses regardless of budget and technical skills, we will consider this limitation as a compromise between availability and accuracy of the analysis being conducted. The analysis of shopping paths that have not ended with purchases also has its weaknesses, as it is not possible to determine with certainty whether or not they will end up as a purchase in the future. Therefore, shopping paths (purchases) that are incomplete can mislead and influence the model results and performance.

When selecting companies for the purpose of this study, we conducted the following steps:

1. We identified Slovak ecommerce stores with the majority of online sales compared to offline sales.
2. We gathered contacts where was possible and initiated request for cooperation on this study.
3. From companies that agreed on participation, we analyzed the available data (tracking setup, sample size etc.) and selected the proper companies.
4. We signed non-disclosure agreement with selected companies and obtained the data.
5. As a result, we were able to work with data from four Slovak companies.

To analyze the e-commerce of the electronic component seller, the data included 284 034 customer journeys, with total revenue generated € 7 665 694. The description of customer journeys is summarized in Table 2. Most of the customer journeys will be analyzed regarding the company selling sports nutritional supplements (Company 4), while the highest sales being recorded at the same time. Nevertheless, the lowest average order value was measured for this business, which also results from the nature of the products sold. On the contrary, the highest average order value was recorded for Company 1, which specializes in the sale of electronic components. In this business, however, the highest average and median value of the number of customer interactions with the business website before the purchase was found. We can assume that the average order value is directly related to the length of the shopping path as customers are likely to decide longer to buy the products. However, the problem of the length of a shopping path may be noted for Company 4, where customers are deciding on average for a long time, and its customers only make low-value purchases. A high number of interactions can also result in a higher cost per customer acquisition if this customer journey involves many interactions with paid marketing channels. The lowest median value for the length of the shopping path can be observed in Company 3, which is a good indicator of the accuracy and persuasiveness of the marketing communications used due to the relatively high average order value. In this case, Company 3 can achieve a low cost-per-acquisition and, therefore, achieve a satisfactory return on investment of its marketing communications.

Table 2 Characteristics of the customer journeys entry data

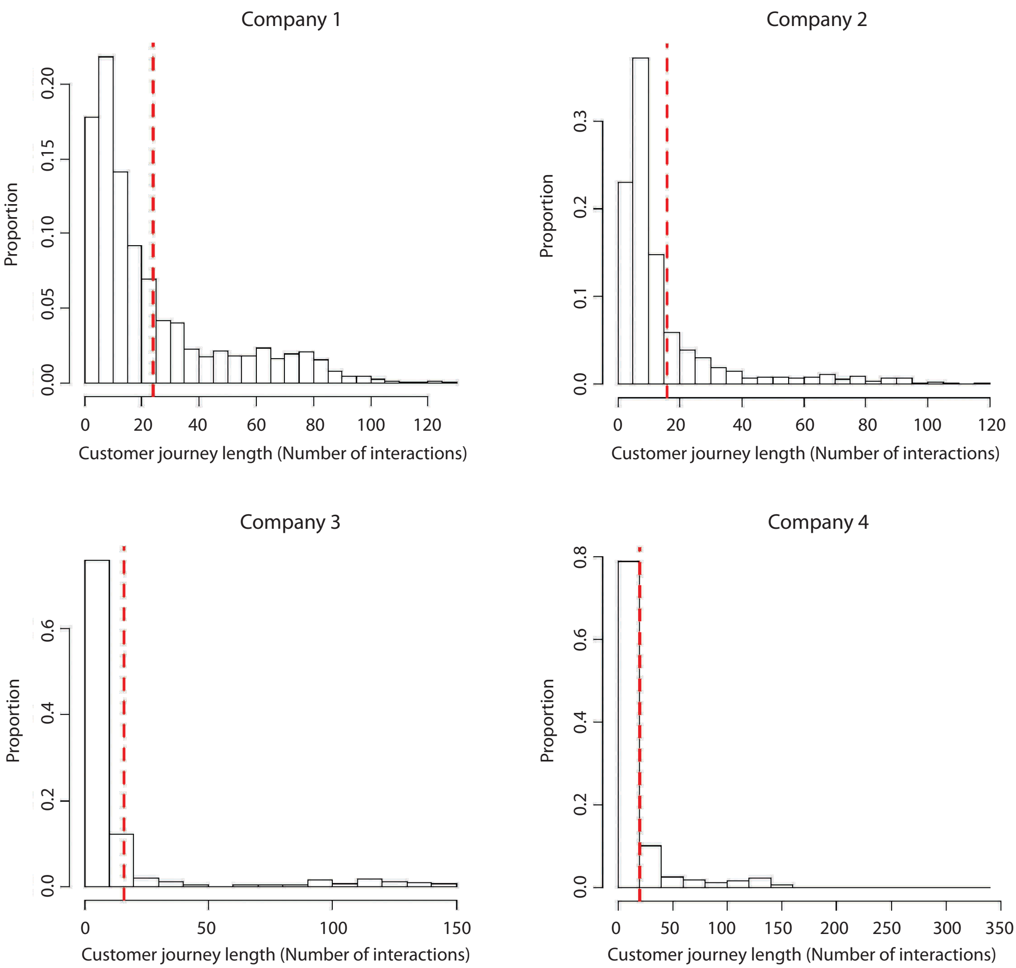
	Company 1	Company 2	Company 3	Company 4
Number of conversions	6 304	21 119	2 118	255 034
Total amount of purchases	1 579 778 €	976 515 €	219 720 €	4 889 682 €
Average order value	250.60 €	46.24 €	103.74 €	19.17 €
Customer journey duration (mean)	23.92	15.72	15.73	20-Feb
Customer journey duration (median)	14	9	6	10

Source: Finstat and our own processing

Figure 1 shows the histograms of the number of interactions in customer journeys per individual analyzed companies. When looking at the histograms, we can say that the distribution of a several customer interactions is positively skewed. This means that most customers choose to buy a product with fewer visits to the website (interactions). Since it was not possible to isolate new and returning customers, we assume that returning customers also have an impact on the skew of this distribution. Another effect may be the inability to monitor a user's activity on multiple devices, for example, when a user can perform product research on a mobile device and then purchase the product on a desktop computer. In this case, such behavior will be recorded as a single-interaction customer journey (if the user did not visit the website on the desktop before making the purchase once again).

In addition to the distribution of customer journeys lengths, it was intended to analyze how this length will affect the revenue generated by businesses.

Figure 1 Customer journey lenght (histogram)



Source: Our own processing

When analyzing the cumulative growth of individual business revenue, it is possible to see that the largest proportion of revenue is generated by customer journeys with fewer interactions. Except for Company 1, where almost 60% of revenue is generated in 20 or fewer interactions in the customer journeys, there is at least 80% of the revenue generated in fewer than 10 interactions. This means that indecisive customers participate in only a small part of total revenue. For business, therefore, a better strategy is to acquire new ("decisive") users than to persuade indecisive potential customers to buy. However, this statement does not apply if these indecisive customers buy repeatedly and decide on the next purchase with fewer visits to the website. As with analyzing the length of shopping paths, the growth of cumulative sales may be distorted by repeating customers and customers buying on multiple devices.

Table 3 Characteristics of the customer journeys entry data

	Company 1	Company 2	Company 3	Company 4
Number of conversions	3 219	16 187	2 118	193 352
Total revenue	463 263 €	736 805 €	219 720 €	3 445 730 €
Average order value	144 €	46 €	104 €	18 €

Source: Our own processing

For the purpose of this study, we were supposed to cover the behavior of new customers, not repeating ones. Therefore, we excluded customer journeys that contained only channel *Direct* from our analysis. This type of journeys might indicate the behavior of repeating customer who already knows the website and therefore visit the website directly. Such customer journeys might affect positively the impact of the channel *Direct* and, therefore, provide us with false assumptions. Table 3 shows the difference in number of conversions, revenue and average order value of analyzed customer journeys without excluded ones.

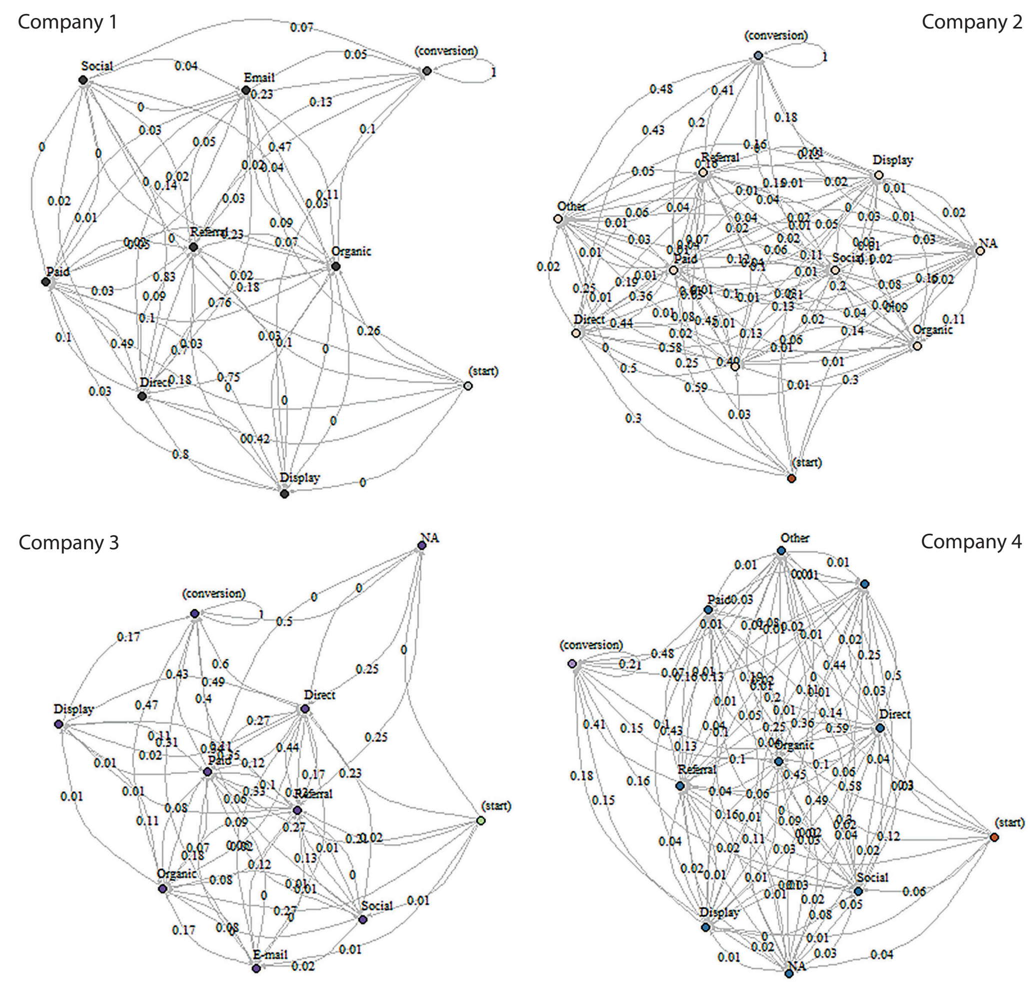
4 THE GENERAL MODEL, CHOICE OF THE APPROPRIATE ORDER OF THE MARKOV MODEL

The Markov model and its application in the customer journey analysis before purchasing products is the content of this part of the study. In the introduction, a basic Markov model of the first order will be applied and the subsequent analysis will focus on assessing the transition probability in the set of determined states defined in the previous sections of the study. This will make it possible to determine which marketing channels can help generate purchases most likely. Subsequently, using the GDL (Global Dependency Level) estimator, which has been mentioned as the most appropriate, an optimal order of the Markov model will be selected. In the last part of this section, the results of attribution modeling of the Markov model of the higher order will be compared with the results obtained using particular heuristic models (first interaction, last interaction, linear model).

In all four analyzed businesses, we were interested in whether there is a greater likelihood that the customer will purchase when using the selected marketing channels. For this reason, the initial step of the analysis was the generation of transition diagrams for each analyzed company. Figure 2 shows the customer journeys transition diagrams for all the analyzed businesses.

The individual points of the graphs represent specific states – marketing channels. It can be noticed that the transition diagram starts with the state *(start)* that represents the start of the customer journey, and ends with the state *(conversion)* that represents the conversion or transaction. Individual states are linked by nodes, each node containing information about the transition probability from a particular state to another particular state. The nodes between the two marketing channels m and n show two probabilities – the probability of transition from state m to state n and the probability of transition from state n to state m. The nodes that connect states *(start)* and *(conversion)* contain only one probability because no customer journey is heading to the state *(start)*. Likewise, no customer journey is heading from the state *(conversion)* towards the marketing channels used. This is a logical state because, after the performed transaction, it does not make sense to monitor what marketing channels the customer uses at the point of further interaction with the business. In the state *(conversion)*, you can see a loop with a pictured probability of 1. This loop originated from computational reasons during the implementation

Figure 2 Transition diagrams



Source: Our own processing

of all the customer journeys. Since all the customer journeys have to go from the state *(start)* to the state *(conversion)*, the loop serves to complete each further interaction until the moment when all the customer journeys from the data file get through the state *(conversion)*.

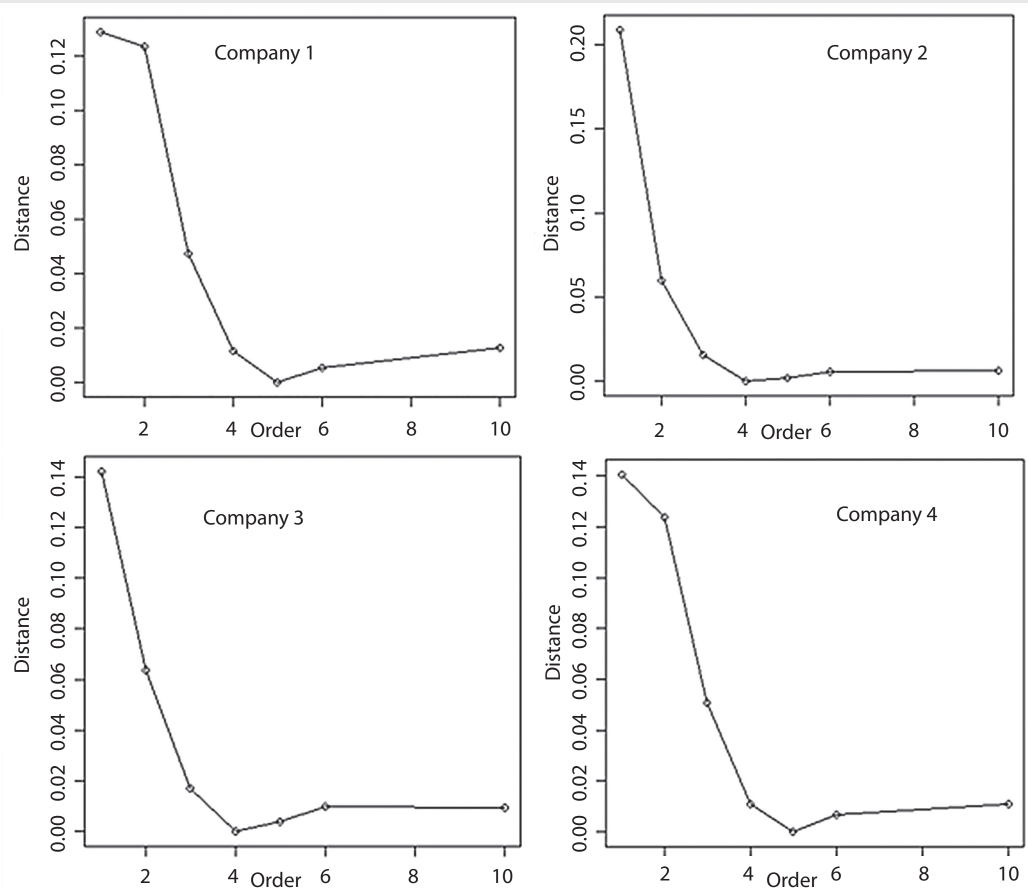
Looking at the heatmaps, it is clear that the analyzed businesses slightly differ in the used marketing channels, but most of them have been used by all the businesses. When analyzing the transition probability, two events can be observed in all four companies:

1. If we do not consider completing a purchase (conversion), with almost every marketing channel, the most likely next step is to visit the website from the *Direct Traffic* source. This means that the result of a positive brand experience when visiting from any source is that customers will either remember the page (and will write a URL directly to the browser on the next visit), or they will save the page as a bookmark to access it later. The positive experience is, therefore, a predictor of increased awareness of the website/brand. This observation is presented as Direct Traffic Effect in study by Kakalejčík et al. (2020).

2. When visiting a website whose source is labeled as *Direct Traffic*, in case of the most of the companies, there is the highest probability that customers will purchase the product. This means that after the customer brand awareness is built, the customer himself will make a purchase when visiting from the *Direct Traffic* source, with the highest probability. Based on this knowledge, a company should strive for the best possible experience during the first customer visits, resulting in a transaction later. In case of Company 3, *Direct Traffic* has the highest transition probability to close the sales, however, the rest of the channels in company's portfolio is similarly valuable.

Within the defined H1 working hypothesis, it was assumed that the first order of the Markov chain used does not represent an accurate representation of the behavior of customers who purchase products online. This supposition continues to work with the assumption that in some cases the customer historically recalls not just an immediate visit to a previous visit to an online merchant's shop during which the customer purchased. In many cases, a customer's memory may exceed the limit of one visit before to the purchase. For this reason, we consider the Markov model of the first order to be irrelevant, while our goal is to set the order that will respond to the customer's behavior the most. Anderl et al. (2016) in their study proved that the use of the third order is the most appropriate for the attribution problem. However, another estimator was used for this prediction.

Figure 3 Markov order (order evaluation)



Source: Our own processing

Figure 3 presents the results of the use of various Markov chain orders, with the criterion for the efficiency of use being chi-square distance when applying the GDL estimator. In this case, the appropriate order is the order in which the local minimum of chi-square distance is achieved. Referring to Figure 3, it is obvious that for Company 1 and 3, the most appropriate order is the fourth order. On the other hand, the fifth order seems to be the most appropriate in the case of Company 2 and 4. It can be seen that the right order when analyzing only new customer varies between 4^{th} and 5^{th}. For practical reasons, however, it can be argued that even with this company the use of the fourth order would fulfill the intended purpose. For the use of the fifth order, it would be necessary to obtain $(9 \times 8)^5$ parameters (almost 2 billions) for a given number of states, but in the case of the fourth order, it is only $(9 \times 8)^4$ parameters (almost 27 millions). The number of parameters directly affects the size of the sample needed for the analysis, as well as the calculation capacity to perform the analysis. Even the transition matrix diverges more when the fifth order is used. So, we concluded that the Markov model of the fourth order might be more appropriate attribution modeling method for e-shops operating on the Slovak market. Based on this finding, we can reject the defined H1 working hypothesis. The result obtained is in contradiction with the result achieved by Anderl et al. (2016).

When analyzing the selection of the appropriate order for the Markov chain, changes in the removal effect were also in the center of our attention. It is established that the higher the removal effect of a given marketing channel, the more important a marketing channel for the business, because excluding it from the marketing portfolio would greatly reduce the number of transactions (conversions) achieved. Table 4 shows the removal effects using the first order of the Markov model in terms of conversions (C), as well as in terms of revenue generated (R).

Table 4 Removal Effects (Markov model of the first order)

	Company 1		Company 2		Company 3		Company 4	
	C	R	C	R	C	R	C	R
Direct traffic	0.90	0.92	0.54	0.60	0.58	0.61	0.81	0.84
Organic search	0.51	0.51	0.51	0.50	0.40	0.37	0.51	0.50
Reference resources	0.19	0.19	0.09	0.36	0.406	0.40	0.16	0.17
Social networks	0.03	0.04	0.07	0.07	0.03	0.03	0.21	0.21
E-mail	0.29	0.30	-	-	0.02	0.02	0.07	0.07
Paid Search	0.37	0.35	0.48	0.50	0.38	0.36	0.50	0.50
Display	0.01	0.01	0.01	0.01	0.02	0.02	0.04	0.03
Other	-	-	-	-	-	-	0.04	0.04
N/A	-	-	0.19	0.20	0.01 <	0.01 <	0.14	0.14

Source: Our own processing

From Table 4 it is obvious that *Direct Traffic* is reaching the highest removal effects, both in terms of conversions and revenue. This knowledge directly supports the findings from the previous part of the analysis which concluded that visits from the *Direct Traffic* channel have the greatest chance of ending up with a purchase. Looking at *Organic Traffic* we can also see relatively high removal effects, which, moreover, are very similar to each of the analyzed businesses. From Table 4, it is also possible to deduct the considerable strength of *Organic search* and Paid *search* and its impact on the number of conversions and sales achieved.

As the testing of the H1 work hypothesis defined the fourth order (Company 1 and 3) and fifth order (Company 2 and 4) as the most effective order in attribution modeling, the analysis of the removal effects using this order was also the focus. The results and percentage differences compared to the first-order removal effects of conversions (C) and generated revenues (R) are shown in Table 5.

Table 5 Removal effects (Markov model of the fourth and fifth order)

	Comp.1 (4th order)		**Comp.2** (5th order)		**Comp.3** (4th order)		**Comp.4** (5th order)	
	C	R	C	R	C	R	C	R
Direct traffic	0.88	0.94	0.62	0.64	0.49	0.54	0.80	0.86
	(–2.7%)	(+2.0%)	(+10.9%)	(+6.4%)	(–15.1%)	(–12.6%)	(–1.4%)	(+1.8%)
Organic search	0.56	0.56	0.50	0.49	0.39	0.36	0.54	0.53
	(+9.2%)	(+9.5%)	(–1.5%)	(–3.1%)	(–1.1%)	(–2.2%)	(+7.5%)	(+6.3%)
Reference resources	0.15	0.19	0.20	0.21	0.31	0.39	0.15	0.18
	(–22.4%)	(–1.1%)	(115.7%)	(+78.6%)	(–15.9%)	(–1.4%)	(–7.9%)	(+7.7%)
Social networks	0.03	0.06	0.07	0.07	0.03	0.03	0.22	0.27
	(+1.8%)	(+71.2%)	(+12.7%)	(+3.9%)	(–6.6%)	(–5.5%)	(–13.5%)	(–0.9%)
E-mail	0.30	0.42	-	-	0.02	0.02	0.07	0.08
	(+3.7%)	(+40.7%)			(–2.6%)	(–9.1%)	(–5.2%)	(+20.2%)
Paid search	0.36	0.35	0.52	0.52	0.33	0.36	0.49	0.46
	(–4.2%)	(–0.2%)	(+7%)	(+2.8%)	(–12.4%)	(+1.3%)	(–0.8%)	(–3.9%)
Display	0.01 <	0.02	0.01<	0.01<	0.02	0.02	0.03	0.04
	(–3.6%)	(+73.8%)	(+6.6%)	(+16.3%)	(–24%)	(–14.8%)	(–13.4%)	(+3.8%)
Other	-	-	-	-	-	-	0.04	0.04
							(–7%)	(–0.2%)
N/A	-	-	0.22	0.23	0.01 <	0.01 <	0.11	0.11
			(+19.3%)	(+13.8%)	(+4.1%)	(+31.9%)	(–18.9%)	(–19.2%)

Source: Our own processing

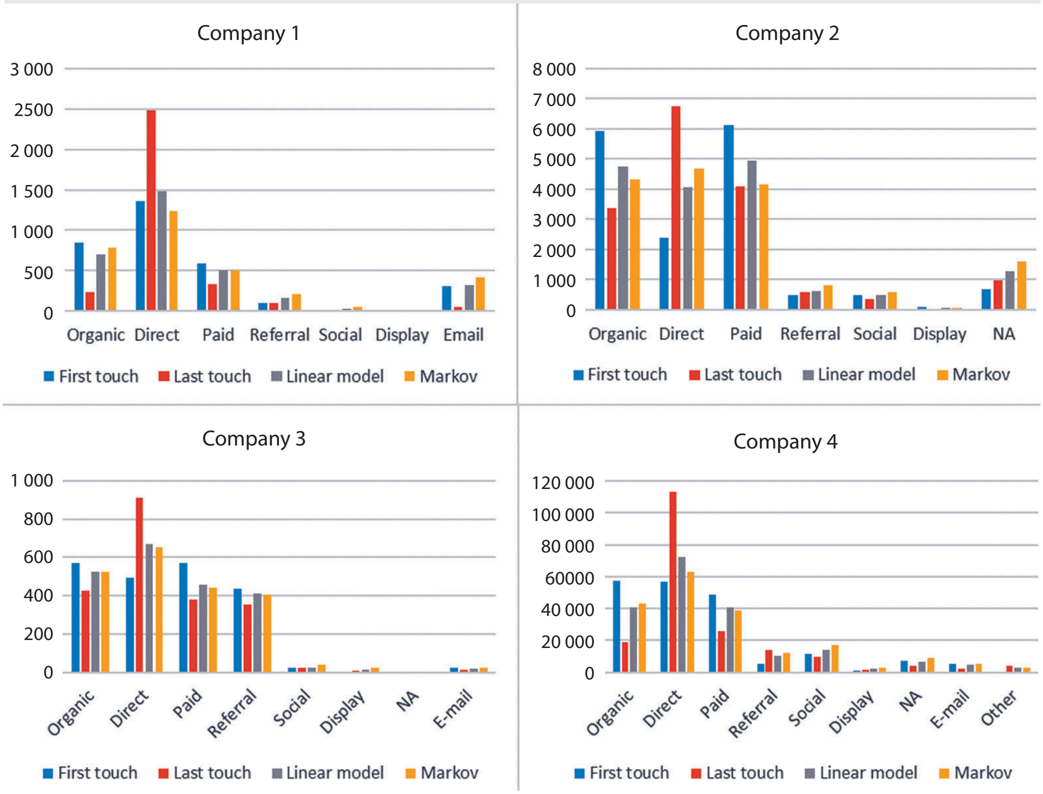
When using the higher order of the Markov chain, it can be noticed that the removal effect after a higher order application changes differently in case of generated conversions compared to the revenue generated. In some cases, the removal effect for the revenue generated increases while removal effect for conversions decreases. This means that, although the dependence of a company on a single marketing channel when selling is lower in terms of the number of conversions generated (as the transition matrix diverges more compared to the first order), the potential loss of revenue value when removing this channel is lower by less percent or, on the contrary, is higher. In other words, despite the lower number of lost conversions, the lost value of revenue would still be high. Therefore, despite the lower significance in the field of transactions obtained, the marketing channel is still of high importance in terms of its contribution to the overall economic result. The subject of analysis should, therefore, emphasize monitoring the removal effects of both attributes, as ignoring one of them may lead to an error in deciding whether or not to use the particular marketing channel.

The last part will focus on the comparison of the Markov model with heuristic models. Figure 4 represents a comparison of the Markov chain and heuristic models, focusing on the distribution of the number of conversions given by the models.

5 ATTRIBUTED NUMBER OF CONVERSIONS BY THE MARKOV MODEL AND HEURISTIC MODELS

As far as the *Direct Traffic* source is concerned, it can be noted that in each of these four cases this source is undervalued by the Markov model in comparison with the heuristic models: the last interaction and the linear model. Since the last interaction model is currently the most used and the linear model represents one of the few multi-touch heuristic models, this finding is significant. The impact of *Organic search* increased, at least compared to last interaction model. In case of Company 1, 2 and 4 *Organic search* also gained compared to linear model, too. In case of *Paid Search*, results are unclear. In case of Company 1, *Paid Search* has gained against each other models. However, in case of Company 2 and 4, it has gained only in comparison with last interaction model and linear model. When comparing the source of traffic *Social Networks*, despite the low number of conversions achieved, it is obvious that the Markov model attributes a higher value to this source than heuristic models. The same situation occurs when comparing the *E-mail* source.

Figure 4 Markov order (order evaluation)



Source: Our own processing

When analyzing the number of conversions of the *Display ad* source, the attributed results were very low. Except for Company 2, where the Markov model was undervalued compared to the first interaction model, the Markov model was assigned a higher number of conversions than all the heuristic models.

Referring to Figure 4, it is also possible to see that conversions mostly depend on 3 basic marketing sources – *Direct Traffic, Organic Search, and Paid Search* (except for referring resources at Company 4, which may have an established network of influencers,[7] and Company 3 which might get a lot of other websites provide links to their website). As a result, a company should strive to be placed as high as possible in the search results for business-related search queries. Also, it should invest in the paid search because it can achieve rapid victories that would not be possible to achieve by the search engine optimization (more in Halligan and Shah, 2014) in a short time. Last but not least, the business should focus on brand awareness as well as a positive user experience when users from other marketing sources are visiting the website. At this point, it is important to add that awareness can also be created by banner advertising, but if the user does not click on the banner but remembers only the brand name or URL of the website, the banner ad will not get a conversion credit (more in the discussion on this part of work).

6 LIMITATIONS AND FUTURE WORK

The results of this analysis can be influenced by factors that could not be taken into account during its implementation. The limitations are as follows:

- **The absence of a NULL state:** the analyzed customer journeys only contain the data about the users who purchased on the website. Clifton (2015), however, discusses that only a small percentage of users (a standard of 3%) purchases in the e-shop. Therefore, the non-inclusion of the NULL state in the analysis can be reflected in the results of the benefits of individual marketing channels. This may also be the reason why Anderl et al. (2016) list the third order of the Markov chain as the appropriate order while we list the fourth order.
- **The impossibility of separating new from returning customers:** we tried to remove repeating customers from our analysis, however, our approach was certainly not accurate. Some of the companies might have offline communications in place (out-of-home communications) and therefore, some of the customers might visit the website directly without previously purchasing on the website.
- **Customer journeys represented the interactions with a website:** the customer could also come into contact with the marketing communication of businesses elsewhere than on the company's website. For example, he could see an ad and not click on it, look at a social network page (e.g., Facebook), and not click on a website, etc. Such behavior was not included in customer journeys.
- **The ambiguity of the *Direct Traffic* source:** the *Direct Traffic* source could represent other marketing channels, e.g. a visit from a mobile app (Facebook, Messenger), a browser bookmark, or an offline ad such as TV, billboards, flyers, or catalogs. Also, each of the analyzed businesses has a brick-and-mortar store.
- **The impossibility of verifying the accuracy of the results obtained:** in contrast to Anderl et al. (2016) or Li and Kannan (2014), it was not possible to test the results of our study using the prediction model, as the Markov model does not allow it. As only the customer data were available, the predictive capabilities of this model might be limited and the model would probably be overfitting and biased. Therefore, the predictive ability of the model must be tested in practical operations.
- **The difference between attribution modeling and setting an optimal budget for the marketing mix:** Danaher and van Heerde (2018) discuss that attribution modeling and budget optimization for the marketing mix are two different concepts. For attribution modeling, the marketing channel that occurred multiple times in the customer journey is attributed to a higher value for conversions/revenue earned. However, this does not mean that the allocation of the marketing budget based on the attribution results does not guarantee its optimal use. Attribution is dependent on exposure and not on costs. Allocating a fixed budget to maximize profits depends primarily on costs, not

[7] Influencer is the user that has an above average impact on other users in the network. For example, a social networking influencer (such as Facebook, Twitter, etc.) is a user who can influence the behavior of other users (Williams, 2016).

exposures. A higher exposure leads to a higher attribution and does not affect on the fixed budget allocation to maximize profits. Higher marketing media costs do not affect attribution but affect fixed budget allocation to maximize the profits. When using the Markov chain for attribution, the exposure of marketing media was reduced if the same marketing medium followed itself in the customer journey at least twice. This method was partly an attempt to eliminate the effect of attribution by exposure.

Although the results of the analysis are limited by the facts described above, it can be noted that only one of these limitations refers directly to the used method. Other limitations exist at the level of analytical software implementation and are associated with the data collection, not analysis. These limitations have arisen because we were not allowed to intervene right into the implementation phase (installation and configuration of the tool) as described in the (Clifton, 2015) four-phase digital analysis process. The added value of the above-mentioned analysis concerns the phase of analytical "adulthood" that is presented in the same process.

When analyzing customer journeys through attribution modeling, future research should consist of the following activities.

- Confirmation of the correct use of the Markov chain of the fourth/fifth order in a wider range of businesses in the Slovak market environment and possible consequences into the European Union markets.
- Removal of the above-mentioned survey restrictions. Priority is to extend the analysis to the customer journeys that do not end with the purchase of the product. Removing this limitation may cause a change in the results of the original analysis performed using the Markov chain.
- Verifying the results obtained during the real operations of the companies that have participated in the study conducted by us.

CONCLUSIONS

Multichannel attribution helps companies assign the value to each marketing channel to select the profitable ones. The main objective of this paper was to define the current state of multichannel attribution and, based on the literature study, analyze the data of the selected companies using the Markov chains. Attribution modeling has already been the focus of the study of several authors.

Based on the works of Anderl et al. (2014; 2016), and Bryl (2016), the Markov chain was used to evaluate the benefit, examining individual customer journeys before the purchase was made. Examined customer journeys came from four e-commerce businesses primarily focused on sales on the Slovak market.

Using the Markov chain, we worked with the assumption that the use of its first order did not correspond to real-world customer behavior. We believe that future customer interaction does not depend only on its current step, but it is influenced by interactions made in the past. Therefore, it is advisable to use a higher order of Markov chain in the analysis. This assumption was directed to the formulation of the H1 hypothesis. The previous work of Anderl et al. (2016) has already pointed out that Markov's third order provides higher accuracy of the model. Our analysis, using the GDL estimator, concluded that it is of the utmost importance for companies operating on the Slovak market to use the fourth and fifth order. This is partly related to the average and median length of customer journeys in the businesses being implemented. H1 hypothesis was therefore rejected. Using the Markov model of the fourth order helped us to uncover another phenomenon that is related to the removal effect of the marketing channel from the portfolio of available marketing channels. With the order growth, the effect of removing individual marketing channels is decreasing, but when looking at revenue, this effect decreases more slowly. This means that, despite of the growing impact of interactions between marketing channels (spillover or carryover effect), removing a given marketing channel from shopping paths will be reflected more significantly in the generated revenues.

At the end of the study, the attribution of the conversions between the newly-formed Markov model and the heuristic models was compared to see if the channel attribution of the conversions differs. When comparing the Markov chain and the heuristic model, the Linear Model (it was chosen for comparison because it considers the entire customer journey as the only one from the monitored heuristic models), it was found that the Markov model attributes a lower value to the selected marketing channels, while attributing a higher value to others.

Although several limitations were identified, we consider the proposed method to be replicable across companies. Each company using Google Analytics can obtain the input data for the proposed model and, therefore, can run the analysis mentioned in this paper. As these companies have an implementation of the software in their control, they can also overcome the limitations we were not able to eliminate.

The methodology and results might be used by companies whose customers need more than one interaction with the company to purchase a product. Moreover, the results and methods can also be used by companies that do not generate online sales to evaluate other types of conversions.

ACKNOWLEDGMENT

The research was realized within the national project "Decision Support Systems and Business Intelligence within Network Economy" (Contract No. 1/0201/19) funded by Grant Agency for Science, Ministry of Education, Science, Research and Sport of the Slovak Republic.

References

ALTOMARE, D. *Package 'ChannelAttribution'* [online]. 2016. [cit. 25.9.2019]. <https://cran.rproject.org/web/packages/ChannelAttribution/ChannelAttribution.pdf>.

ANDERL, E. *Three Essays on Analyzing and Managing Online Consumer Behavior*. Dissertation thesis, Universität Passau, Passau, 2014.

ANDERL, E. ET AL. *Mapping the customer journey: a graph-based framework for online attribution modeling* [online]. 2014. [cit. 24.8.2018]. <http://ssrn.com/abstract=2343077>.

ANDERL, E. ET AL. Mapping the customer journey: Lessons learned from graph-based online attribution modeling. *International Journal of Research in Marketing*, 2016, Vol. 33, pp. 457–474.

BARAJAS, J. ET AL. Experimental designs and estimation for online display advertising attribution in marketplaces. *Marketing Science*, 2016, Vol. 35, No. 3, pp. 465–483.

BRYL, S. *Attribution model with R (part 1: Markov chains concept)* [online]. 2016. [cit. 24.9.2018]. <http://analyzecore.com/2016/08/03/attribution-model-r-part-1/>.

BULYGO, Z. *What The Highest Converting Websites Do Differently* [online]. 2012. [cit. 24.9.2018]. <https://blog.kissmetrics.com/what-converting-websites-do/>.

CONSTANTIN, C. Principal Component Analysis-A Powerful Tool in Computing Marketing Information. *Bulletin of the Transilvania University of Brasov, Economic Sciences,* 2014, Vol. 7, No. 2, pp. 25–30.

CLIFTON, B. *Successful Analytics: Gain Business Insights by Managing Google Analytics*. West Susse, Advanced Web Metrics, 2015.

FERENCOVÁ, M. ET AL. Social Media Usage in Product Information Searching. *Applied Mechanics and Materials*, 2015, Vol. 795, pp. 69–76.

FINSTAT [online]. 2017. [cit. 27.1.2018]. <https://finstat.sk/>.

GEYIK, S. C. ET AL. Multi-touch Attribution Based Buget Allocation in Online Advertising. In: *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*, New York: ACM, 2014, pp. 1–9. ISBN 978-1-4503-2999-6

GHOSE, A. AND TODRI, V. Towards a Digital Attribution Model: Measuring the Impact of Display Advertising on Online Consumer Behavior [online]. *MIS Quarterly*, August 1, 2015. <https://ssrn.com/abstract=2638741> or <http://dx.doi.org/10.2139/ssrn.2638741>.

GOOGLE ANALYTICS [online]. 2018. [cit. 27.1.2018]. <https://analytics.google.com/>.

KAKALEJČÍK, L., BUCKO, J., DANKO, J. Impact of direct traffic effect on online sales [online]. *Journal of Research in Interactive Marketing*, 2020, Vol. 14, No. 1, pp. 17–32. <https://doi.org/10.1108/JRIM-01-2019-0012>.

KARLIN, S. AND TAYLOR, H. *A First Course in Stochastic Processes*. Academic Press, 1975.

KAUSHIK, A. *Webová analytika 2.0: Kompletní průvodce analýzami návštevnosti*. Brno: Computer Press, 2011.

LI, H. AND KANNAN, P. K. Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment. *Journal of Marketing Research*, 2014, Vol. 51, pp. 40–56.

MASSARA, F., LIU, S. S., MELARA, R. D. Adapting to a retail environment: Modeling consumer–environment interactions. *Journal of Business Research*, 2010, Vol. 63, No. 7, pp. 673–681.

MOLNAR, CH. *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable* [online]. 2020. [cit. 10.9.2020]. <https://christophm.github.io/interpretable-ml-book/>.

NISSAR, T. AND YEUNG, M. Purchase Conversions and Attribution Modeling in Online Advertising: an Empirical Investigation. *EMAC 2015,* Leuven, 2015.

R. *The R Project for Statistical Computing* [online]. 2019. [cit. 10.8.2019]. <https://www.r-project.org>.

RENTOLA, O. *Analyses of Online Advertising Performance Using Attribution Modeling*. MSc. Thesis, Aalto University, Aalto, 2014.

PETERSON, R. A. Response construction in consumer behavior research. *Journal of Business Research*, 2005, Vol. 58, No. 3, pp. 348–353.

ROBERGE, M. *The Sales Acceleration Formula: Using Data, Technology, and Inbound Selling to go from $0 to $100 Million.* Hoboken: John Wiley&Sons, 2015.

SHAO, X. AND LI, L. Data-driven Multi-touch Attribution Models. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, 2011.

SMARANDACHE, F. AND VLADUTESCU, S. Towards a Practical Communication Intervention. *Revista de Cercetare si Interventie Sociala*, 2014, Vol. 46, p. 243.

STERNE, J. *Artificial Intelligence for Marketing*. Hoboken: John Wiley&Sons, 2017, p. 368. ISBN 9781119406334

SZULC, C. Valuing Digital Marketing Channels with Attribution Models [online]. 2013. [cit. 10.10.2019]. <http://relevance.com/wp-content/uploads/2013/04/multi-touch-eBook-digitalrelevance.pdf>.

VAN DEN BERG, G. AND PIETERSMA, P. *25 Need to Know Management Models*. Harlow, Pearson, 2015.

WANG, Q. ET AL. When Structural Models Meets Big Data: Examining Multi-Device Attribution for Native Ads Using TB-Sized Data. *Proceedings of Workshop on Information Systems and Economics,* 2015. [cit. 10.10.2019]. <https://research.yahoo.com/mobstor/publication_attachments/LargeScaleStructualAnalysis.pdf>.

WILLIAMS, R. *Influencer Economy: How to Launch Your Idea, Share it with the World, and Thrive in the Digital Age*. Ryno Lab, 2016, p. 276. ISBN 978-0-996077-11-8

WOOFF, D. A. AND ANDERSON, J. M. Time-weighted multi-touch attribution and channel relevance in the customer journey to online purchase. *Journal of statistical theory and practice*, 2015, Vol. 9 No. 2, pp. 227–249.

YADAGIRI, M. M. ET AL. A non-parametric approach to the multi-channel attribution problem. *16th International Conference*, Basel: Springer International Publishing, 2015.