Mapping the customer journey:

A graph-based framework for online attribution modeling

Eva Anderl*, Ingo Becker, Florian v. Wangenheim, Jan H. Schumann

October 18, 2014

Eva Anderl is a Ph.D. Candidate at Passau University, Chair of Marketing and Innovation, Innstr. 27, D-94032 Passau, Germany, eva.anderl@uni-passau.de.

Ingo Becker is a Ph.D. Candidate at ETH Zurich, Chair of Technology Marketing, Weinbergstr. 56/58, CH-8092 Zurich, Switzerland, ibecker@student.ethz.ch.

Florian v. Wangenheim is Professor of Technology Marketing, ETH Zurich, Chair of Technology Marketing, Weinbergstr. 56/58, CH-8092 Zurich, Switzerland, fwangenheim@ethz.ch.

Jan H. Schumann is Professor of Marketing and Innovation, Passau University, Chair of Marketing and Innovation, Innstr. 27, D-94032 Passau, Germany, jan.schumann@uni-passau.de.

The authors gratefully acknowledge support from the German Federal Ministry of Education and Research (Project “Fre(E)S – Productivity of Free E-Services”, FKZ 01FL10038) and the Wharton Customer Analytics Initiative.

Mapping the customer journey:

A graph-based framework for online attribution modeling

Advertisers employ various channels to reach consumers over the Internet but often do not know to what degree each channel actually contributes to their marketing success. This attribution challenge is of great managerial interest, yet so far academic approaches have not found wide application in practice. The authors introduce a graph-based framework to analyze multichannel online customer path data as first- and higher-order Markov walks. According to a comprehensive set of criteria for attribution models, embracing both scientific rigor and practical applicability, four model variations are evaluated on four, large, real-world data sets from different industries. Results indicate substantial differences to existing heuristics such as "last click wins" as well as alternative attribution approaches. Applying the proposed framework to four different data sets enables generalizations and helps identify avenues for future research. The framework offers support to practitioners by facilitating objective budget allocation and allows for future applications such as real-time bidding.

1 Introduction

Online advertising is essential to many industries' promotional mix (Raman, Mantrala, Sridhar, & Tang, 2012). While advertisers use a variety of online marketing channels,[1] including paid search and display marketing, as well as e-mail, mobile, and social media advertising to reach consumers, they lack transparency about the degree to which each channel or campaign contributes to their companies' success. Marketing executives thus call for performance measures of the contributions of each online marketing channel

[1] In this study, we use the term "online marketing channels" to cover different online marketing instruments, including search engine advertising, display, or social media advertising.

(Econsultancy, 2012a). The challenge of attributing credit to different channels (Neslin & Shankar, 2009) involves finding ways to measure “the partial value of each interactive marketing contact that contributed to a desired outcome” (Osur, 2012, p. 3). To date, many advertisers apply simple heuristics, for example “last click wins”, such that the value gets attributed solely to the marketing channel that directly preceded the conversion (Econsultancy, 2012a; The CMO Club & Visual IQ, Inc., 2014), and any prior customer interactions are disregarded.

Modern technological advancements enable recording of customer journey data though, enabling new ways to address the attribution problem. We define an online customer journey of an individual customer as including all touch points over all online marketing channels preceding a potential purchase decision that lead to a visit of an advertiser's website.[2] Although some software tool providers now offer multitouch attribution solutions,[3] last and first click wins heuristics remain among the most widely used attribution methods in practice (Econsultancy, 2012a). Furthermore, even the multitouch attribution tools used in practice largely rely on simple heuristics, such as linear attribution, which splits the contribution evenly across all channels (Econsultancy, 2012a, 2012b; Osur, 2012). Only a small number of vendors offer statistical or algorithmic attribution methodologies (Osur, 2012), but their mechanisms remain publicly unavailable and irreproducible (Dalessandro, Perlich, Stitelman, & Provost, 2012).

In turn, despite its practical relevance, the attribution problem only recently has become a focus for marketing researchers (Abhishek, Fader, & Hosanagar, 2012; Berman, 2013; Haan, Wiesel, & Pauwels, 2013; Kireyev, Pauwels, & Gupta, 2013; Li & Kannan,

---

[2] Similar concepts have been referred to as the path to purchase (Xu, Duan, & Whinston, 2014) or the online purchase journey (Li & Kannan, 2014). .

[3] Examples of companies used in this context include, but are not limited to, Adclean, Adometry, Atlas, C3 Metrics, ClearSaleing, Coremetrics (IBM), Google, Theorem, Trueeffect, Visual IQ, Icrossing, and [x+1].

2014; Xu, Duan, & Whinston, 2014). However, to the best of our knowledge, sophisticated attribution approaches have not found wide application in practice. Acceptance and adaption of marketing models demands more than analytical rigor (Lehmann, McAlister, & Staelin, 2011; Little, 1970; Wübben & Wangenheim, 2008); managers hesitate to base their decisions on mechanisms whose results are not available when they need them (Lodish, 2001) or if they do not understand how the insights are generated (Lilien, Roberts, & Shankar, 2013; Little, 1970, 2004). Thus, a practice-oriented attribution approach needs to fulfill important criteria for managerial acceptance, such as ease of interpretation, versatility, or algorithmic efficiency.

In response, we introduce a novel, practice-oriented attribution framework based on Markovian graph-based data mining techniques, extending an approach originally developed in the context of search engine marketing (Archak, Mirrokni, & Muthukrishnan, 2010). We model and analyze individual-level multichannel customer journeys as first- and higher-order Markov graphs, using a property we call removal effect to determine channel contributions. Based on four large, real-world data sets, we evaluate our models according to a set of criteria for attribution models, building on prior research on managerial decision models (Lehmann et al., 2011; Lilien, 2011; Little, 1970, 2004; Lodish, 2001). We compare our suggested framework against existing attribution approaches and apply it to a real-life system implemented at a German multichannel tracking provider. Thus, we contribute to marketing theory and practice in multiple ways.

First, we extend the emerging attribution literature (Abhishek et al., 2012; Berman, 2013; Haan et al., 2013; Kireyev et al., 2013; Li & Kannan, 2014; Xu et al., 2014) by presenting a novel framework that meets both academic standards and criteria relevant for implementation in practice. Second, we reduce the thresholds for applying attribution techniques (Econsultancy, 2012a) and foster cross-industry acceptance (Dalessandro et al., 2012) by developing a comprehensive set of evaluation criteria for attribution models. Third, we contribute new insights into online marketing effectiveness in a multichannel setting. By using four, real-world data sets we are able to extend prior studies (Abhishek et al., 2012; Li & Kannan, 2014; Xu et al., 2014) and move towards empirical generalizations. Fourth, our practice-oriented approach helps to bridge the widening gap between marketing theory and practice (Jaworski, 2011; Lilien, 2011; Reibstein, Day, & Wind, 2009) by providing solutions to several explicit problems that online marketers confront. Our framework facilitates an objective logic for deriving budget optimization processes and updates the mental models of decision makers. Finally, the versatility of our framework allows for future applications such as real-time bidding in ad exchanges.

2 Research background

2.1 Research on attribution modeling

Academic research on attribution is still scarce (Raman et al., 2012; Tucker, 2012) but can build on prior studies pertaining to online advertising effectiveness. Most existing research focuses on single channels, such as search (Ghose & Yang, 2009; Rutz & Bucklin, 2011; Yang & Ghose, 2010) or display (Braun & Moe, 2013; Goldfarb & Tucker, 2011). Studies comparing the short- and long-term effectiveness of different online advertising channels based on aggregate data relate to the attribution problem (Breuer & Brettel, 2012; Breuer, Brettel, & Engelen, 2011), yet they do not attempt to award credit for conversions. Jordan, Mahdian, Vassilvitskii, and Vee (2011) examine allocation decisions for publishers, using multiple attribution approaches, and derive optimal allocation and pricing rules for publishers selling advertising slots. In a study of the economic welfare consequences of the use of attribution technologies, Tucker (2012) finds evidence for more conversions at lower costs, due to the ability to systematically substitute towards selected campaigns across advertising platforms. Practice-oriented literature on attribution mainly highlights the relevance of the topic or summarizes ongoing industry activities, without providing methodological details

(Chandler-Pepelnjak, 2008, Econsultancy, 2012a, 2012b; Lovett, 2009; Osur, 2012; Riley, 2009).

Furthermore, we know of few academic studies that address the online attribution problem: Shao and Li (2011) introduce two attribution approaches, a bagged logistic regression model and a simple probabilistic model. Building on their work, Dalessandro et al. (2012) propose a more complex, causally motivated attribution methodology based on cooperative game theory. Based on a set of simulated campaign data they find that advertisers tend to assign credit to conversions that are driven by the users' volition to convert rather than on the factual influence of the advertisement. Focusing on the interplay between advertisers and publishers, Berman (2013) evaluates the impact of different incentive schemes and attribution methods on publishers' propensity to show ads and the resulting profits of advertisers. He introduces an analytical model based on Shapley value, similar to the model proposed by Dalessandro et al. (2012), and compares it to the last click wins heuristic. Abhishek et al. (2012) suggest a dynamic hidden Markov model (HMM), based on individual consumer behavior, that captures a consumer’s deliberation process along typical stages of the conversion funnel: disengaged, active, engaged, and conversion. They find that different channels affect the consumers in different states of their decision process. For example, display ads usually impact consumers early in the decision process, moving them from a disengaged state to activity or engagement. Li and Kannan (2014) propose a Bayesian model to measure online channel consideration, visits, and purchases using individual conversion path data and validate it in a field experiment. They use the estimated carryover and spillover effects to attribute conversion credit to different channels and find that these channels' relative contributions are significantly different from last click wins. By means of a mutually exciting point process model, Xu et al. (2014) calculate average conversion probabilities for different online advertising channels, showing that the conversion rate measure underestimates the effect of display ads compared to search ads. A multivariate time-series model based on aggregate data by Kireyev et al. (2013) analyzes attribution dynamics for display and search advertising. They derive spillover effects from display towards search conversion; however, display ads also increase search clicks, thereby increasing costs for search engine advertising. Finally, Haan et al. (2013) propose a structural vector autoregression (SVAR) model, also based on aggregate data, to determine the effectiveness of various offline and online advertising channels.

2.2 Criteria for attribution modeling

Putting academic marketing models to work in practice is challenging, because the most complex model is not necessarily the one that will affect an organization's productivity (Little, 1970; Lodish, 2001). In the following, we therefore conceptualize marketing attribution modeling with a catalogue of six criteria that reflect scientific rigor as well as aspects relevant to the implementation in practice. We build on prior research into the acceptance of marketing decision models (Leeflang & Wittink, 2000; Lilien, 2011; Little, 1970, 1979, 2004; Lodish, 2001; Reibstein et al., 2009) and connect it with criteria previously discussed in the context of attribution modeling (Dalessandro et al., 2012; Shao & Li, 2011). Table 1 provides an overview of the six criteria we propose and their relation to prior literature.

Marketing decision models should enable the computation of the relative impacts of different decision variables and enable *objectivity* in budget decisions (Lilien, 2011). In the case of attribution, models need to be able to assign credit to individual channels or campaigns in accordance with their factual ability to generate value, such as by contributing to conversions or increasing revenues (Dalessandro et al., 2012). Although objectivity seems to be an obvious criterion, common models applied in practice break this rule. For example, models that condense user journeys to one click (e.g., first- or last-click heuristics) omit any additional marketing contacts, and models based on predefined weights fail to attribute credit fairly across channels.

Although attribution primarily takes a retrospective view, attribution models should be able to correctly predict conversion events (Shao & Li, 2011). In addition to ensuring scientific rigor, this classification helps to persuade managers of the model's credibility (Lodish, 2001). We therefore introduce *predictive accuracy* as a second criterion. With the exception of approaches based on Shapley value (Berman, 2013; Dalessandro et al., 2012), all of the cited studies evaluate *predictive accuracy* using measures such as log-likelihood (Abhishek et al., 2012; Li & Kannan, 2014), mean absolute percentage error (Haan et al., 2013; Li & Kannan, 2014), or the sum of squared errors (Xu et al., 2014).

Robustness is another important metric to evaluate model fitness (Little, 1970, 2004; Shao & Li, 2011). Robustness conveys the ability of a model to deliver stable and reproducible results if the model runs multiple times (Little, 1970) and is indispensable for sustainable budget decisions. Only Shao and Li (2011) explicitly evaluate *robustness*, which they call variability. Li and Kannan (2014) provide additional validation using a field experiment and Xu et al. (2014) use out-of-sample validation.

To ensure managerial acceptance, models need to be simple and easy to communicate (Little, 1970, 2004), which we summarize as *interpretability*. Simplicity comprises the intuitive understandability of a model "by all parties with material interest" (Dalessandro et al., 2012, p. 2). The interpretability of the results is of utmost importance for practical acceptance, because managers often refuse to use black box approaches that conceal how they work or how they generate results (Lilien, 2011; Little, 1970; Lodish, 2001).

Table 1

Evaluation criteria for attribution models

Criterion	Definition	Relation to prior research: Studies	Relation to prior research: Description
Objectivity	Models must be able to assign credit to individual channels or campaigns in accordance with their factual ability to generate value, such as contributing to conversions or increasing revenues.	Lilien (2011)	Models should allow for computing the relative impact of decision variables and enable objectivity in evaluating decisions options.
		Dalessandro et al. (2012)	Attribution systems should reward an individual channel in accordance with its ability to affect the likelihood of conversion (fairness).
Predictive accuracy	Models should be able to predict conversion events correctly.	Lodish (2001)	Predictive validity is important to persuade managers of a model's credibility.
		Shao and Li (2011)	Attribution models should have high accuracy in predicting active or inactive users (accuracy).
Robustness	Models should deliver stable and reproducible results if they run numerous times.	Little (1970, 2004)	Models should be robust to avoid bad, unstable results.
		Shao and Li (2011)	Attribution models should deliver stable estimates (variability).
Interpretability	Model structure should be transparent to all stakeholders with reasonable effort, and the results should be interpretable with relative ease.	Little (1970)	Model users should be able to transfer model results directly into managerial decisions.
		Little (1970, 2004)	Models should be simple and easy to communicate.
		Little (1970); Lodish (2001); Lilien (2011)	Models should be easy to interpret, because managers refuse to apply black box approaches.
		Dalessandro et al. (2012)	An attribution system needs to be accepted "by all parties with material interest" based on its "statistical merit" and the "intuitive understanding" of the system's components.
Versatility	Versatility combines adaptability and ease of control. Adaptability is the capability to incorporate new information that becomes available over time. Ease of control enables users to adjust inputs to fit company-specific requirements and derive appropriate outputs.	Little (1970)	Models should be "adaptive" and "easy to control." "Adaptive" describes the capability to update the model as soon as new information become available; "easy to control" enables the user to adjust inputs to modify outputs.
		Lodish (2001)	Models should deliver an adequate level of aggregation to achieve acceptance by managers.
Algorithmic efficiency	The speed of computing model outputs when they are requested.	Little (1970, 2004)	Model structures should be complete in relevant issues and able to handle many phenomena without being bogged down.
		Lodish (2001)	Models need to provide results as soon as managers require them to be applicable in practice.
		Archak, Mirrokni, and Muthukrishnan (2010)	As a basic precondition for practical purposes, a methodology must be able to handle large data volumes fast and efficiently.

Little (1970) posits that models should be adaptive and easy to control, which we combine to *versatility*. Adaptability encompasses the capability of incorporating new information that becomes available over time, which is particularly critical in rapidly changing environments (Leeflang & Wittink, 2000). In the online environment, the set of available channels is constantly evolving (Evans, 2009). An attribution framework therefore should be able to include varying channels and should easily be extended toward innovative forms of advertising. Furthermore, a model should allow for different aggregation levels (Lodish, 2001). Though some existing models are highly flexible (Berman, 2013; Dalessandro et al., 2012; Shao & Li, 2011; Xu et al., 2014), *versatility* can be limited by explicit assumptions about the customer decision process as well as channel characteristics. For example, Haan et al. (2013) exclude channels with performance-based payment models, such as affiliates, to avoid endogeneity.

Finally, we introduce *algorithmic efficiency*, the speed with which the model computes outputs when requested, as a sixth criterion. With recent advances in online tracking technologies, clickstream data sets can be of tremendous size, comprising millions of clicks or even billions of impressions (Bucklin & Sismeiro, 2009). An attribution methodology must be able to handle these volumes efficiently, because practitioners will not apply results that are not available when required (Lodish, 2001). Next, we seek to develop a model that meets all of the suggested criteria and evaluate it using real-life data sets.

3 Data

Our research is based on four clickstream data sets provided by online advertisers, in collaboration with a multichannel tracking provider. Clickstream data record each user's Internet activity and thus trace the navigation path he or she takes (Bucklin & Sismeiro, 2009). For each visit to the advertiser's website during the observation period, the data include detailed information about the source of the click and an exact timestamp. Clicks either represent a direct behavioral response to an advertising exposure or result from the user entering the advertiser's uniform resource locator (URL) directly into the browser, so these sources comprise all online marketing channels, as well as direct type-ins. We also know for each visit whether it was followed by a conversion, in this case a purchase transaction. We use this information to construct customer journeys that describe the click pattern of individual consumers across all online marketing channels and their purchase behavior. Thus, we not only track successful journeys ending with a conversion but also journeys that never lead to a conversion, within a timeframe of 30 days of the last exposure.

The data collection occurs at the cookie-level, such that we identify individual consumers—or more accurately, individual devices. The use of cookie data suffers several limitations, such as an inability to track multidevice usage or bias due to cookie deletion (Flosi, Fulgoni, & Vollman, 2013; Rutz, Trusov, & Bucklin, 2011), yet cookies remain the industry standard for multichannel tracking (Tucker, 2012). We do not include information on offline marketing channels, because measuring individual-level exposure to multiple offline media proves highly difficult in practice (Danaher & Dagger, 2013).

The advertisers that provide the data sets for this study operate in different industries: Data set 1 was provided by an online travel agency. Data sets 2, 3, and 4 originate from specialized online retailers selling apparel and luggage, respectively. The three retailers address a broad audience and sell a wide range of brands. All the advertisers in our sample are pure online players, so we can exclude online/offline cross-channel effects (Wiesel, Pauwels, & Arts, 2011). Each data set includes a minimum of 405,000 journeys per advertiser. Their average length is 1.3–2.5 contacts, and between 0.9% and 2.0% of all journeys lead to a successful conversion. All advertisers included in the evaluation distinguish seven or eight different online channels, though the channels used differ partly across firms. Search engine advertising (SEA), search engine optimization (SEO), affiliate, and newsletter appear in all four data sets. Other channels used by the advertisers include display, price comparison, social media advertising, and retargeting banners. In Table 2, we present detailed descriptions of our data sets. Table 3 provides information on the distribution of clicks per channel, illustrating the variation in our data sets.

Table 2

Descriptions

Description	Data set 1	Data set 2	Data set 3	Data set 4
Industry	Travel	Fashion retail	Fashion retail	Luggage retail
Number of different channels	8	8	7	7
Number of clicks	1,478,359	1,639,467	1,125,979	615,111
Number of journeys	600,978	1,184,583	862,112	405,339
Thereof with length ≥ 2	206,519	170,914	142,039	105,031
Thereof with length ≥ 5	48,344	30,095	12,416	11,475
Journey length	2.46 (8.860)	1.38 (1.916)	1.31 (1.238)	1.52 (4.587)
Number of conversions	9,860	10,153	16,200	8,115
Journey conversion rate	1.64%	0.86%	1.88%	2.00%

Note. Standard deviations are in parentheses.

Table 3

Number of clicks per channel per journey

Description	Data set 1		Data set 2		Data set 3		Data set 4	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Type In	n/a	n/a	.29	1.049	.25	.810	.13	4.416
SEA	.58	.898	.15	.718	.18	.548	1.13	1.122
SEO	.17	.613	.17	.795	.43	.280	.16	.538
Affiliate	.06	.339	.64	1.015	n/a	n/a	.00	.164
Price Comparison	.09	1.829	.00	.053	n/a	n/a	.04	.271
Newsletter	.03	.235	.07	.724	.02	.156	n/a	n/a
Referrer	n/a	n/a	.06	.260	.14	.379	.04	.209
Retargeting	.01	.164	n/a	n/a	.00	.007	.02	.203
Display	1.46	8.746	.02	.295	n/a	n/a	n/a	n/a
Social Media	n/a	n/a	n/a	n/a	.28	.815	n/a	n/a
Other	.06	.357	n/a	n/a	n/a	n/a	n/a	n/a

4 Model development

We propose a graph-based Markovian framework to analyze customer journeys and derive an attribution model, adapting an approach proposed by Archak et al. (2010) in the context of search engine advertising. Markov chains are probabilistic models that can represent dependencies between sequences of observations of a random variable. They have a long history in marketing (Styan & Smith, 1964) and have been used frequently to model customer relationships (Homburg, Steiner, & Totzek, 2009; Pfeifer & Carraway, 2000). Other applications include advertising frequency decisions (Bronnenberg, 1998) and brand loyalty (Che & Seetharaman, 2009).

In our model, we represent customer journeys as chains in directed Markov graphs.[4] A Markov graph $M = \langle S, W \rangle$ is defined by a set of states

$$S = \{s_1, \ldots, s_n\} \tag{1}$$

and a transition matrix W with edge weights

$$w_{ij} = P\left(X_t = s_j \middle| X_{t-1} = s_i\right), 0 \le w_{ij} \le 1, \sum_{j=1}^{N} w_{ij} = 1 \; \forall \, i. \tag{2}$$

Using this graph-based approach allows us to represent and analyze customer journeys in an efficient way as the size of the final graph does not depend on the number of journeys in the data set but only on the number of states.

4.1 Base model

Customer journeys contain one or more contacts across a variety of channels. In the base model, each state s_i corresponds to one channel. If an advertiser employs three different channels C1, C2, and C3 in his online marketing mix, the model would include three states

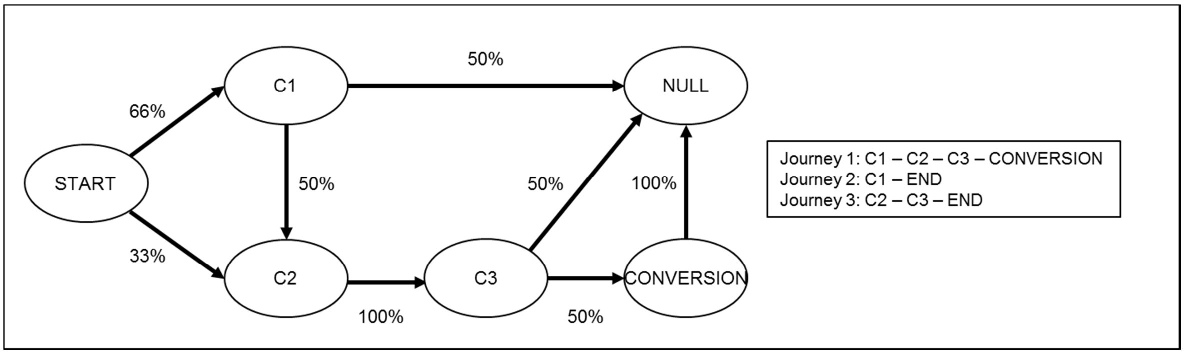
[4] Called adgraphs by Archak, Mirrokni, and Muthukrishnan (2010).

C*1*, *C2*, and *C3*.[5] Additionally, all graphs contain three special states: a *START* state that represents the starting point of a customer journey; a *CONVERSION* state representing a successful conversion; and an absorbing *NULL* state for customer journeys that have not ended in a conversion during the observation period. The full set of states S in our example would hence look as follows: S = {*START*, *CONVERSION*, *NULL*, *C1*, *C2*, *C3*}.

The transition probability w_{ij} in the base model corresponds to the probability that a contact in channel i is followed by a contact in channel j. For the first channel in each journey, we add an incoming connection from the *START* state. If a customer journey ends in a conversion, we connect the state representing the last channel in the journey to the *CONVERSION* state, otherwise it leads to the *NULL* state. For modeling reasons, we always add a connection from the *CONVERSION* state to the *NULL* state. Cycles in the graph are possible, such as when a sequence of two identical channels appears in a customer journey. Figure 1 shows an exemplary Markov graph based on three customer journeys. Figure 2 provides a graphical structure of the simple model for data set 1.

Figure 1

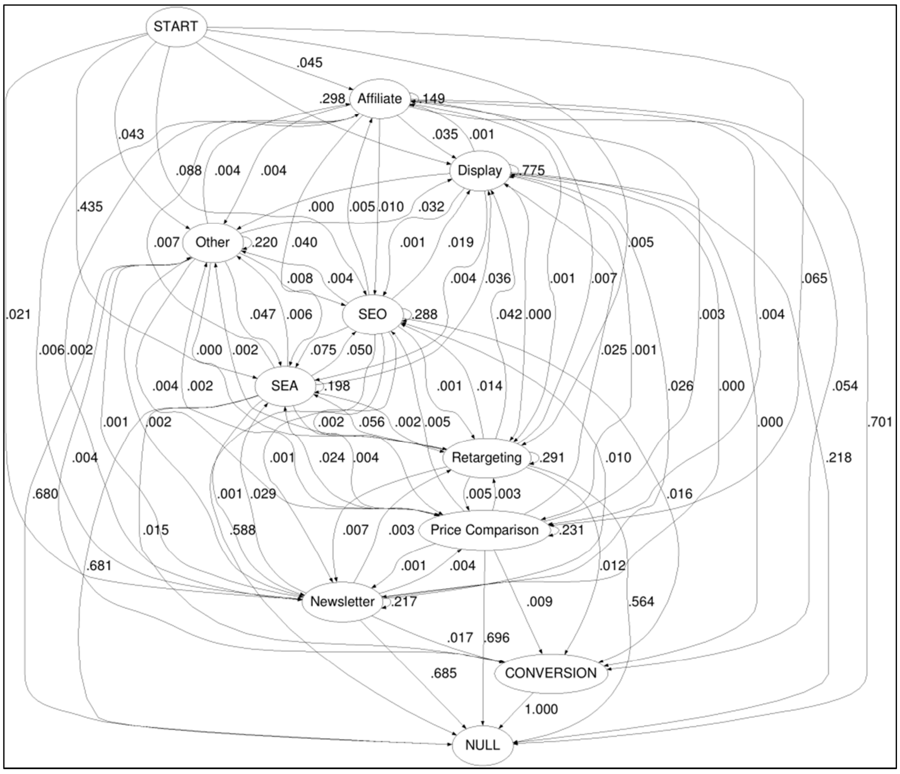
Exemplary Markov graph



[5] As we do not make any assumptions on the channels used, we employ dummy channels in our examples. In practice, the set of channels—and thus the set of states—depends on the actual channels used by the advertiser.

Figure 2

Markov graph for data set 1 (base model)



4.2 Higher-order models

Markovian models suggest that the present only depends on the first lag and do not incorporate previous observations. Because prior research suggests that clickstreams should not be regarded as strictly Markovian though (Chierichetti, Kumar, Raghavan, & Sarlós, 2012; Montgomery, Li, Srinivasan, & Liechty, 2004), we extend the approach proposed by Archak et al. (2010) by introducing alternative higher-order models, in which the present depends on the last k observations. Transition probabilities thus can be defined as follows:

$$P(X_t = s_t | X_{t-1} = s_{t-1}, X_{t-2} = s_{t-2}, \dots, X_1 = s_1) \qquad (3)$$
$$= P(X_t = s_t | X_{t-1} = s_{t-1}, X_{t-2} = s_{t-2}, \dots, X_{t-k} = s_{t-k}).$$

For our implementation, we exploit the knowledge that a Markov chain of order k, over some alphabet A, is equivalent to a first-order Markov chain over the alphabet A_k of k-tuples. States in higher-order models therefore include k-tuples of states in the first-order models. Unfortunately, the number of independent parameters increases exponentially with the order of the Markov chain and quickly becomes too large to be estimated efficiently with real-world data sets (Berchtold & Raftery, 2002). Considering these implementation issues in relation to algorithmic efficiency, we limit our analyses to Markov chains of a maximum order of four.

4.3 Removal effect

The representation as Markov graphs allows identifying structural correlations in the customer journey data that can be used to develop an attribution model. Archak et al. (2010) propose a set of *ad factors* to capture the role of each state, such as *Eventual Conversion*(s_i), i.e. the probability of reaching conversion from a given state s_i. *Visit*(s_i) is the probability of passing s_i on a random walk beginning in the *START* state. For attribution modeling, we propose using the ad factor *Removal Effect*(s_i), defined as the change in probability of reaching the *CONVERSION* state from the *START* state when we remove s_i from the graph. As *Removal Effect*(s_i) reflects the change in conversion rate if the state s_i was not present, it is well suited to measure the contribution of each channel (or channel sequence). Using the assumption that all incoming edges of the state s_i that we remove are redirected to the absorbing *NULL* state, *Removal Effect*(s_i) is equivalent to the multiplication of *Visit*(s_i) and *Eventual Conversion*(s_i).[6] The removal effect can thus be efficiently calculated using matrix multiplication or applying local algorithms provided by Archak et al. (2010).

Removal Effect(s_i) can take values between 0 and the total conversion rate. However, as most existing attribution heuristics use percentage values, we report removal effects per state as percentages of the sum of all removal effects (excluding the special states START, CONVERSION, and NULL), when comparing our results to other models. Higher-order

[6] For a proof, see Archak et al. (2010).

models allow us to calculate removal effects for states representing channel sequences; in addition, we also aggregate the mean values for each channel to get information on a channel level.

5 Results

We evaluate our models according to the previously established criteria—objectivity, predictive accuracy, robustness, interpretability, versatility, and algorithmic efficiency—and compare our results against existing attribution approaches.

5.1 Application of evaluation criteria

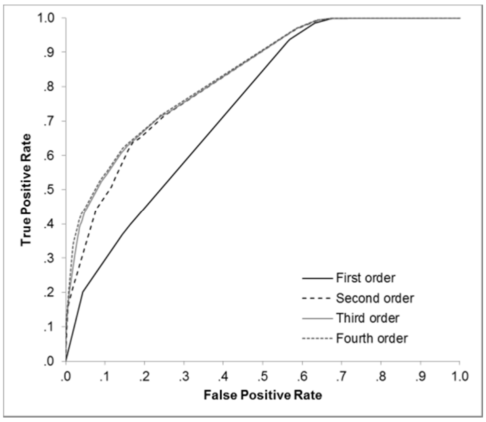
The graph-based framework we propose satisfies the *objectivity* criterion, as it includes all contacts in the analysis and makes no previous assumptions about the importance of individual channels or channel order. In contrast with existing practical applications, the analyses are completely data driven, and the mechanics of model building and ad factor calculation are fully disclosed and reproducible.

Predictive accuracy measures how many conversion events get classified correctly. We use the 10-fold cross-validation, which is superior to leave-one-out validation or bootstrapping for measuring predictive performance both within and out of sample, since all the data serve as the holdout once (Kohavi, 1995; Sood, James, & Tellis, 2009). However, standard metrics for classification accuracy, such as percentage correctly classified or log-likelihood, are poor metrics for measuring classification performance in the case of unequal misclassification costs or when class distribution is skewed (He & Garcia, 2009; Provost, Fawcett, & Kohavi, 1998). As the discriminative power of these measures is limited in our context, where journey conversion rates do not exceed 2%, we turn to alternative measures to evaluate predictive accuracy.

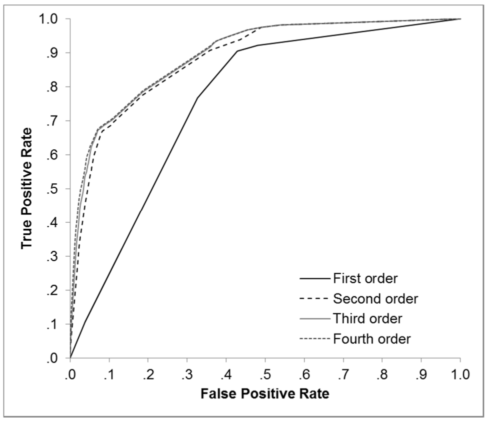
First, we choose the receiver operating characteristic (ROC) curve that decouples classification performance from class distributions and misclassification costs. A ROC curve is a two-dimensional graph; the true positive rate α is plotted on the x-axis, while the false positive rate 1 - β appears on the y-axis (Bradley, 1997; Fawcett, 2006). To compare our models, we reduce ROC performance to a single scalar value, the area under the ROC curve (AUC), which we calculate using linear interpolation (Bradley, 1997). Figure 3 contains the ROC curves for all models based on a within-sample evaluation of all journeys. As a second measure, we calculate the top-decile lift for each model. Top-decile lift is defined as the proportion of the 10% of journeys predicted to be most likely to convert who actually end in a conversion relative to the baseline conversion rate (Neslin, Gupta, Kamakura, Lu, & Mason, 2006). In Table 4, we report both measures and compare our approach to the last click wins and first click wins heuristics as well as a logit model.

Figure 3

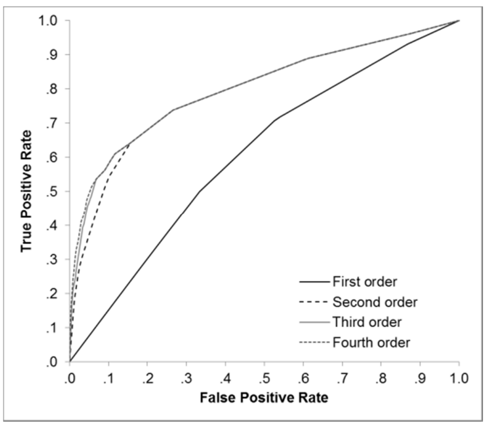
ROC curves (within sample)



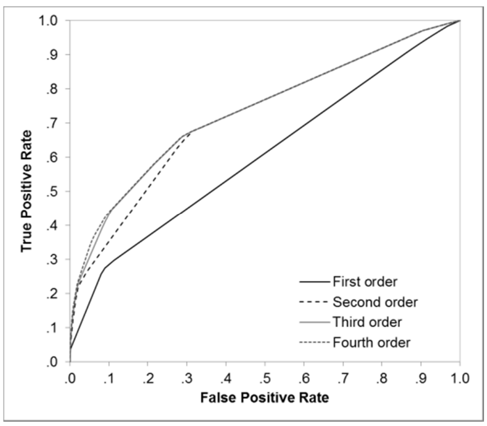
Data set 1



Data set 2



Data set 3



Data set 4

Table 4

Predictive accuracy

Measure	Sample	Data set	Base model	Second order	Third order	Fourth order	Last click wins	First click wins	Logit
AUC	Within sample	DS 1	.7398 (.0006)	.8214 (.0007)	.8323 (.0008)	.8357 (.0010)	.7398 (.0006)	.7133 (.0005)	.8185 (.0022)
		DS 2	.7604 (.0006)	.8834 (.0006)	.8932 (.0005)	.8967 (.0005)	.7609 (.0007)	.7726 (.0008)	.8654 (.0011)
		DS 3	.6079 (.0009)	.7933 (.0015)	.7996 (.0014)	.8040 (.0018)	.6079 (.0009)	.5980 (.0006)	.8037 (.0011)
		DS 4	.6029 (.0010)	.7191 (.0029)	.7309 (.0009)	.7333 (.0009)	.6029 (.0010)	.5528 (.0007)	.7281 (.0013)
	Out-of-sample	DS 1	.7396 (.0061)	.8207 (.0022)	.8305 (.0062)	.8294 (.0063)	.7406 (.0056)	.7136 (.0050)	.8185 (.0074)
		DS 2	.7607 (.0056)	.8832 (.0054)	.8924 (.0044)	.8933 (.0050)	.7619 (.0056)	.7728 (.0071)	.8653 (.0086)
		DS 3	.6079 (.0079)	.7930 (.0086)	.7988 (.0088)	.8006 (.0089)	.6093 (.0082)	.5990 (.0064)	.8036 (.0088)
		DS 4	.6030 (.0093)	.7180 (.0083)	.7280 (.0080)	.7230 (.0072)	.6050 (.0080)	.5550 (.0063)	.7276 (.0084)
Top-decile lift	Within sample	DS 1	2.9056 (.0211)	4.7181 (.0249)	5.2798 (.0198)	5.3587 (.0203)	2.9056 (.0211)	2.4333 (.0203)	4.9226 (.0261)
		DS 2	2.4880 (.0407)	6.8070 (.0156)	6.9306 (.0146)	6.9546 (.0158)	2.4955 (.0312)	2.8692 (.0311)	6.6011 (.0211)
		DS 3	1.5065 (.0254)	5.1744 (.0207)	5.6426 (.0221)	5.6450 (.0232)	1.5065 (.0254)	1.4967 (.0188)	5.7731 (.0197)
		DS 4	2.8089 (.0137)	3.4708 (.1148)	4.1838 (.0134)	4.2897 (.0161)	2.8089 (.0137)	1.6820 (.0144)	4.0112 (.1042)
	Out-of-sample	DS 1	2.9075 (0.1385)	4.6937 (.1593)	5.2393 (.1563)	5.2913 (.1704)	2.9175 (.1508)	2.4296 (.0934)	4.9260 (.1487)
		DS 2	2.4748 (.0742)	6.7889 (.1442)	6.9277 (.1178)	6.9113 (.1202)	2.5103 (.1078)	2.8709 (.1532)	6.6109 (.1684)
		DS 3	1.4896 (.0910)	5.1619 (.1433)	5.6311 (.1993)	5.6107 (.2055)	1.5418 (.1002)	1.5143 (.1368)	5.7714 (.1704)
		DS 4	2.8112 (.1074)	3.4556 (.1737)	4.1198 (.1273)	4.1553 (.1464)	2.8256 (.1114)	1.6897 (.1337)	4.0124 (.1332)

Note. Standard deviations are in parentheses.

Although the overall predictive accuracy varies substantially between data sets, the relative predictive performance of the different model types is comparable, leading to similar rankings of the model types. Within and out-of-sample performance for all models is similar, indicating a low risk of overfitting. With the exception of data set 2, the base model outperforms the first click wins heuristics and leads to similar results as the last click wins approach, yet does not predict conversions as well as a logit model. Increasing the memory capacity substantially improves the predictive performance of our graph-based models. With the exception of data set 3, third-order and fourth-order models outperform the logit model as well as the two heuristic approaches in both AUC and top-decile lift. For data set 3, which has the shortest journeys in our sample, the AUC performance of the higher-order models is similar to the logit models; however, they do not attain the top-decile lift of the logit models.

The third evaluation criterion, *robustness*, applies to two measures. First, predictive accuracy should be robust across all cross-validation repetitions. Table 4 lists the standard deviations of the predictive performance measures for each model as well as for the three models we use as a comparison. The results imply low overall variation without systematic differences between models. Second—and even more important, the variable used for attribution modeling should provide stable attribution results that offer a reliable basis for managerial decisions, such as budget shifts. Therefore, we specifically test the robustness of the *Removal Effect*(s_i) ad factor. For each model state s_i, we compute the average standard deviation of *Removal Effect*(s_i) across ten cross-validation repetitions. We report the stability of the removal effects as percentages of the average removal effect across all states, as the number of states per model and correspondingly the mean *Removal Effect*(s_i) varies. We summarize these validation results in Table 5. For all data sets in our sample, the average standard deviation as a percentage of the average removal effect increases with model order, leading to a necessary trade-off between predictive accuracy and robustness.

Table 5

Removal effect: Average standard deviation as % of average removal effect (10-fold cross-validation)

	Model			
Data set	Base model	Second order	Third order	Fourth order
Data set 1	1.14%	1.92%	3.25%	5.43%
Data set 2	1.51%	2.10%	3.81%	7.57%
Data set 3	1.31%	1.72%	2.78%	5.15%
Data set 4	1.34%	1.80%	2.97%	4.67%

Although objectivity, predictive accuracy, and robustness represent necessary conditions for attribution models, additional criteria such as *interpretability* must be fulfilled to foster acceptance and application in practice. The graphical representation (see Figure 1 and Figure 2) of our framework can help marketing executives understand the basic concept. In discussions with online marketing managers, we discovered that despite their initial skepticism toward algorithmic attribution approaches in general, the proposed framework was regarded as easy to interpret and well accepted. The output metrics can be provided in the same format as existing heuristics and are intuitively interpretable and easy to communicate to other stakeholders.

Because it requires no preliminary assumptions about channels or decision processes, our framework is highly *versatile*. The only prerequisite for building the graphical models we propose is the availability of historical, individual-level tracking data. Our framework can evaluate various conversion types, including sales, sign-ups, or leads, and easily integrate new online marketing channels. Analyses might run on different aggregation levels, such that users can analyze not only channels but also advertising campaigns or even different creatives.

Considering the large data volumes in online marketing (Bucklin & Sismeiro, 2009) and practitioners' requests for regular updates (Econsultancy, 2012a), *algorithmic efficiency* has become a decisive criterion for attribution models. Removal effects can be calculated efficiently in $O(|S|^2)$ time (Cormen, Leiserson, Rivest, & Stein, 2009) and hence allow for frequent model updates. However, as the number of states increases exponentially with the order of the Markov chain, we limit our analyses to lower-order models in order to allow for updates in near real-time. Combining our evaluation results with practical considerations, we recommend third-order models for standard attribution analyses.[7] Using higher-order models also yields additional insights into channel interactions, further increasing managers' understanding of the interplay across channels.

5.2 Attribution results

We compare the attribution results of our proposed framework with the attribution heuristics most widely used in industry practice, namely last click wins, first click wins, and linear attribution (Econsultancy, 2012a), as well as the Shapley value approach (Berman, 2013; Dalessandro et al., 2012). Given our evaluation results, we use third-order Markov models in our comparison. Our analyses are based on the full data sets and show the contribution of each channel towards final conversions. We present the results in Table 6.

We observe significant differences for the results of the Markov model and those of the other models. The channels SEO, display, newsletter, and retargeting are consistently undervalued by the heuristic attribution approaches. In contrast, the contribution of SEA is overestimated by all heuristic approaches. The remaining channels leave a more ambiguous picture: Direct type-ins, when users directly access the company website, seem to be overvalued by the last click wins approach, yet the results compared to the other attribution

[7] However, depending on the use case, an advertiser may choose a different trade-off between predictive accuracy, robustness, and algorithmic efficiency.

approaches are less consistent across data sets. Affiliate is underestimated by the first click wins and the linear heuristic; for last click wins the results differ between data sets. The results obtained by the Shapley value approach deviate substantially from all other approaches and—with the exception of SEA—show high variance across data sets. Overall, while each advertiser needs to derive and verify detailed implications individually, our findings indicate potential for empirical generalizations of multichannel online advertising effectiveness across industries.

In addition to channel-level attribution, higher-order models offer a more detailed view on the interplay of channels, which we illustrate using the second-order model for data set 1 in Table 7. The second-order model attribution results for the other data sets can be found in Appendix A. Across all data sets, the increase in overall purchase probability for most channels is highest right after the START state, near the beginning of the journey[8]—which corresponds with the high share of one-click journeys in our data. Sequences of identical channels show high removal effects in all data sets. For example, in data set 1, affiliate preceded by affiliate has a percentage removal effect of 5.31%, whereas the average removal effect for affiliate preceded by another channel is only 1.22%. The consistent importance of same channel sequences across data sets may indicate idiosyncratic channel preferences for some users comparable to those found in multichannel relational communication (Godfrey, Seiders, & Voss, 2011). Future research should investigate the existence of such preferences, their antecedents, and their implications for multichannel online advertising.

---

[8] A notable exception is retargeting, which explicitly targets customers who have previously visited the advertiser's website (Lambrecht & Tucker, 2013).

Table 6

Attribution results in comparison to other models (in %)

	Data set 1					Data set 2					Data set 3					Data set 4				
	Markov graph (Third order)	Last click wins	First click wins	Linear attri-bution	Shapley value	Markov graph (Third order)	Last click wins	First click wins	Linear attri-bution	Shapley value	Markov graph (Third order)	Last click wins	First click wins	Linear attri-bution	Shapley value	Markov graph (Third order)	Last click wins	First click wins	Linear attri-bution	Shapley value
Type In	n/a	n/a	n/a	n/a	n/a	35.23%	43.91%	40.28%	40.39%	24.48%	27.55%	29.77%	25.51%	26.95%	21.84%	18.89%	22.02%	13.71%	17.62%	21.04%
SEA	46.29%	53.19%	56.36%	54.21%	14.00%	19.81%	22.27%	23.60%	22.18%	15.93%	18.53%	20.16%	20.70%	20.11%	14.53%	59.91%	60.98%	76.26%	67.72%	13.35%
SEO	19.54%	16.76%	16.67%	17.11%	22.24%	15.79%	13.66%	13.24%	14.14%	16.54%	22.03%	21.33%	21.12%	21.56%	12.27%	9.24%	7.79%	5.30%	7.26%	11.21%
Affiliate	19.71%	20.17%	13.66%	16.98%	26.06%	8.72%	7.83%	6.87%	8.05%	7.87%	n/a	n/a	n/a	n/a	n/a	4.41%	3.67%	0.42%	2.33%	51.78%
Price Comparison	5.29%	4.78%	6.05%	5.37%	7.24%	0.21%	0.11%	0.12%	0.14%	1.01%	n/a	n/a	n/a	n/a	n/a	3.56%	2.17%	2.21%	2.37%	-3.77%
Newsletter	4.47%	2.93%	4.28%	3.55%	28.89%	14.54%	8.76%	11.94%	10.96%	16.95%	1.21%	1.15%	1.32%	1.24%	18.03%	n/a	n/a	n/a	n/a	n/a
Referrer	n/a	n/a	n/a	n/a	n/a	1.99%	1.67%	2.58%	1.90%	3.29%	5.72%	6.85%	7.52%	6.91%	14.12%	1.88%	1.65%	1.96%	1.63%	5.83%
Retargeting	1.25%	0.67%	0.78%	0.82%	14.22%	n/a	n/a	n/a	n/a	n/a	0.07%	0.01%	0.00%	0.00%	1.81%	2.09%	1.72%	0.16%	1.07%	0.56%
Display	0.93%	0.14%	0.21%	0.17%	-13.67%	3.70%	1.79%	1.37%	2.25%	13.93%	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Social Media	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	24.87%	20.73%	23.83%	23.22%	17.40%	n/a	n/a	n/a	n/a	n/a
Other	2.52%	1.36%	1.97%	1.80%	1.02%	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$\chi 2$		12,911	18,019	10,992	n/a[a]		40,931	30,076	17,480	111,146		5,948	4,822	2,719	216,499		3,357	36,979	8,056	n/a[a]
df		7	7	7	n/a[a]		7	7	7	7		6	6	6	6		6	6	6	n/a[a]
p		<.001	<.001	<.001	n/a[a]		<.001	<.001	<.001	<.001		<.001	<.001	<.001	<.001		<.001	<.001	<.001	n/a[a]

Note. $\chi 2$ values in comparison to Markov graph attribution results.

[a] $\chi 2$ test not possible because of negative values.

Regarding search channels, prior research has argued that click propensities for paid and organic search results are moderated by involvement (Jerath, Ma, & Park, 2014). However, our findings indicate that personal channel preferences also play a role, as spillover effects from SEA to SEO are much stronger than from SEO to SEA. Whereas consumers who have clicked on sponsored search ads react to both paid and organic results, consumers clicking on organic search results seem to generally prefer organic results. Future research should investigate the role of channel preferences and their interaction with involvement. Besides increasing predictive performance, the application of models with higher orders thus enables a more detailed understanding of the interplay across channels and allows to identify several promising avenues for future research.

Table 7

Attribution results for second order model (data set 1)

	Preceding channel								
Current channel	START	Affiliate	Display	News-letter	Price Com-parison	Retar-geting	SEA	SEO	Other
Affiliate	8.10%	5.31%	0.15%	0.20%	0.23%	0.08%	2.61%	1.00%	0.15%
Display	0.62%	0.03%	0.42%	0.02%	0.02%	0.01%	0.09%	0.02%	0.01%
Newsletter	1.91%	0.07%	0.04%	0.99%	0.02%	0.02%	0.27%	0.13%	0.01%
Price Comparison	3.07%	0.04%	0.03%	0.04%	1.36%	0.01%	0.22%	0.05%	0.04%
Retargeting	0.34%	0.04%	0.02%	0.02%	0.03%	0.24%	0.20%	0.06%	0.01%
SEA	32.52%	0.84%	0.35%	0.48%	0.34%	0.15%	14.27%	2.43%	0.47%
SEO	7.78%	0.42%	0.07%	0.10%	0.11%	0.06%	4.40%	4.75%	0.08%
Other	1.12%	0.03%	0.02%	0.02%	0.01%	0.01%	0.27%	0.06%	0.50%

6 Discussion

Our research contributes to marketing theory and practice in at least five ways. First, we extend the growing literature on online attribution (Abhishek et al., 2012; Berman, 2013; Dalessandro et al., 2012; Haan et al., 2013; Kireyev et al., 2013; Li & Kannan, 2014; Shao & Li, 2011; Xu et al., 2014) by introducing a novel attribution framework that meets both academic standards and criteria relevant for implementation in practice. Our framework represents multichannel online customer journeys as Markov walks in directed graphs. In addition to our base model, we further introduce higher-order Markov graphs, in which the present depends on the last k observations. We use the removal effect, defined as the change in probability of reaching the *CONVERSION* state from the *START* state when s_i is removed from the graph, to derive state and channel contributions, respectively. Using four, large, real-life data sets, we evaluate four model alternatives according to our criteria and compare them to other approaches. Higher-order models outperform existing heuristics and a logit model regarding predictive accuracy. A comparison of the attribution results shows substantial differences between the results of the Markov graphs and existing attribution approaches, including heuristical attribution as well as attribution based on Shapley value (Berman, 2013; Dalessandro et al., 2012). Thus we provide a valid alternative to widely used, often misleading attribution heuristics applied in practice.

Second, we develop a comprehensive set of six criteria required for successful attribution models. Building on existing literature on the acceptance of marketing decision models in practice (Leeflang & Wittink, 2000; Lehmann et al., 2011; Lilien, 2011; Little, 1970, 1979, 2004; Lodish, 2001), we ensure scientific rigor by assessing objectivity, predictive accuracy, and robustness; we also include criteria to encourage application in practice, namely, interpretability, versatility, and algorithmic efficiency. Whereas previous studies have discussed selected properties for attribution methods (Dalessandro et al., 2012; Shao & Li, 2011), we present the first exhaustive set of criteria that acknowledges practitioners' requirements. Clear criteria reduce the barriers to applying attribution techniques in managerial practice (Econsultancy, 2012a) and foster standardization and cross-industry acceptance (Dalessandro et al., 2012). The incentives of advertisers and other market actors, such as publishers or agencies, are seldom congruent (Abou Nabout, Skiera, Stepanchuk, & Gerstmeier, 2012; Berman, 2013), which creates a demand for independent, objective criteria to assess attribution models.

Third, we offer new insights into online marketing effectiveness in a multichannel setting. By using four, real-world data sets, each including at least seven different online channels, we are able to extend prior studies (Abhishek et al., 2012; Li & Kannan, 2014; Xu et al., 2014) and move towards empirical generalizations. We find that SEO, display, newsletter, and retargeting are consistently undervalued by heuristic attribution approaches. In contrast, the contribution of SEA is overestimated by all heuristic approaches. Furthermore, higher-order models significantly outperform both first-order models and one-click heuristics regarding predictive accuracy, which indicates that channels in customer journeys should not be analyzed in isolation. Hence our results add to the evidence that one-click heuristics cannot capture the full contribution of online channels (Abhishek et al., 2012; Li & Kannan, 2014; Xu et al., 2014). In addition, the more detailed view on the interplay of channels provided by higher-order models allows identifying several promising avenues for future research. For example, we find evidence for idiosyncratic channel preferences. Future research should investigate the existence of such preferences, their antecedents, and their implications for multichannel online advertising.

Fourth, our practice-oriented approach helps to bridge the often-lamented gap between marketing theory and practice (Jaworski, 2011; Lilien, 2011; Reibstein et al., 2009). Existing marketing research does not fully capture the increasing richness and complexity of firms' online marketing activities (Yadav & Pavlou, 2014). Besides considering practical considerations in deriving our criteria, we implemented our attribution framework in a real industry environment. We developed a prototype of our framework, including all four model types and implemented it as a real-life system at a German multichannel tracking provider. Thus far, several clients, operating in the fashion, sports equipment, and telecommunications industry, have applied our attribution model in practice, confirming its high usability and positive impact on marketing effectiveness; however, we cannot disclose explicit test results for confidentiality reasons.

Scientifically validated attribution models help resolve several managerial problems. The proliferation of online channels makes budget allocation decisions increasingly complex (Raman et al., 2012). Our framework can facilitate independent budget decisions by providing easy-to-interpret, objective information that factors out subjective influences. Budgets should be allocated across channels according to their factual value contribution; if the budget share of a channel is higher than its actual contribution as measured by our attribution framework, advertisers should readjust their budget allocations. By continuously recalibrating their budget split based on attribution results, managers can move towards an optimal budget allocation. Such usage also should enhance decision makers' expertise and update their mental models, which are prone to systematic errors and biases (Tversky & Kahneman, 1974). The introduction of data-driven attribution also suggests effects on hierarchical structures within organizations and group decision making. In meeting the objectivity criterion, our approach is devoid of personal assumptions, preferences, and other biases that could adversely affect the decision process (Bruggen, Smidts, & Wierenga, 1998; Leeflang & Wittink, 2000)—in marked contrast with existing, widely used attribution methodologies that rely on the (pre)definition of channel or position weights by advertisers (Econsultancy, 2012a; The CMO Club & Visual IQ, Inc., 2014).

Fifth and finally, compared with other attribution models that are purely retrospective, such as the linear-weighted heuristic or the Shapley value approach, our proposed graph-based framework can also be used prospectively. Thus a possible application is real-time bidding in ad exchanges, where advertisers can bid on advertising slots for specific users using information such as the user's location or previous surfing behavior (Muthukrishnan, 2009). Using our framework, advertisers can more accurately calculate the conversion probability *Eventual Conversion*(s_i) of a customer, given his or her previous customer journey. The predicted change in *Eventual Conversion*(s_i) when the advertiser wins the auction and the advertisement is shown to the user also can be used to calculate the value of this slot on an individual user level and thus determine a maximum cost-effective bid. It would be interesting to apply our framework in a real-time bidding setting and compare it to alternative approaches (e.g., Perlich, Dalessandro, Raeder, Stitelman, & Provost, 2013).

7 Outlook

Our research has several limitations that may stimulate research on attribution and online marketing effectiveness. We used four data sets from different industries, but some findings may still be company specific. The customer journeys in these data sets were short on average, including a high number of one-click journeys. However, sophisticated attribution is not required for journeys consisting of just a single click: In that case, both the “last click wins” and the “first click wins” heuristics deliver objective results that would satisfy our criteria, whereas longer journeys increase advertisers’ need to understand channel contributions. Furthermore, although our framework should be well suited to handle such information, our data sets do not include views and exposures to offline channels. We therefore recommend applying this framework to further industries and to integrate not just clicks but views as well as offline channels.

To further evaluate the effectiveness of online marketing channels, companies should consider revenues and profits from conversions, and—potentially in a second step—the CLV of customers acquired. As Chan, Wu, and Xie (2011) show, customers acquired through different online marketing channels differ in CLV. Our graph-based approach is well suited for such extensions with additional data, thus further research should include this information to advance attribution. The timing between contacts along the customer journey would be another interesting extension.

The attribution problem is by definition endogenic; it measures the relative effectiveness of channels in a given setting (Li & Kannan, 2014), so the results are conditional on a number of management decisions, such as channels used, budget limits per channel, or ad creatives employed. Therefore, developing an optimal budget allocation remains an iterative process. Nevertheless, objective attribution is a necessary prerequisite for managers to optimize their budget decisions. Subject to the availability of longitudinal data, attribution results calculated using our framework could also serve as a basis for developing optimization algorithms.

Finally, a strict causal interpretation of customer journeys is difficult, because alternative explanations may exist for the correlations between conversions and advertising exposures. Some channels, such as retargeting, explicitly try to target customers who have a higher propensity to purchase (Lambrecht & Tucker, 2013). Even without special targeting, observed correlations might be due to selection effects, such as an activity bias (Lewis, Rao, & Reiley, 2011). To establish a causal relationship, large-scale field experiments with randomized exposure are required. Such experiments are hard to implement in practice, especially in multichannel settings, but comparing our attribution modeling framework against experimental results would be a valuable follow-up.

References

Abhishek, V., Fader, P. S., & Hosanagar, K. (2012). *Media exposure through the funnel: A model of multi-stage attribution*. Retrieved from http://ssrn.com/abstract=2158421

Abou Nabout, N., Skiera, B., Stepanchuk, T., & Gerstmeier, E. (2012). An analysis of the profitability of fee-based compensation plans for search engine marketing. *International Journal of Research in Marketing*, *29*(1), 68–80.

Archak, N., Mirrokni, V. S., & Muthukrishnan, S. (2010). Mining advertiser-specific user behavior using adfactors. In M. Rappa & P. Jones (Eds.), *Proceedings of the 19th International Conference on World Wide Web* (pp. 31–40). ACM.

Berchtold, A., & Raftery, A. E. (2002). The mixture transition distribution model for high-order Markov chains and non-Gaussian time series. *Statistical Science*, *17*(3), 328–356.

Berman, R. (2013). *Beyond the last touch: Attribution in online advertising*. Retrieved from http://ron-berman.com/papers/attribution.pdf

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, *30*(7), 1145–1159.

Braun, M., & Moe, W. W. (2013). Online display advertising: Modeling the effects of multiple creatives and individual impression histories. *Marketing Science*, *32*(5), 753–767.

Breuer, R., & Brettel, M. (2012). Short- and long-term effects of online advertising: Differences between new and existing customers. *Journal of Interactive Marketing*, *26*(3), 155–166.

Breuer, R., Brettel, M., & Engelen, A. (2011). Incorporating long-term effects in determining the effectiveness of different types of online advertising. *Marketing Letters*, *22*(4), 327–340.

Bronnenberg, B. J. (1998). Advertising frequency decisions in a discrete Markov process under a budget constraint. *Journal of Marketing Research*, *35*(3), 399–406.

Bruggen, G. H. van, Smidts, A., & Wierenga, B. (1998). Improving decision making by means of a marketing decision support system. *Management Science*, *44*(5), 645–658.

Bucklin, R. E., & Sismeiro, C. (2009). Click here for Internet insight: Advances in clickstream data analysis in marketing. *Journal of Interactive Marketing*, *23*(1), 35–48.

Chan, T. Y., Wu, C., & Xie, Y. (2011). Measuring the lifetime value of customers acquired from Google search advertising. *Marketing Science*, *30*(5), 837–850.

Chandler-Pepelnjak, J. (2008). *Measuring ROI beyond the last ad*. Retrieved from http://advertising.microsoft.com/espana/WWDocs/User/es-es/Advertise/engagement_mapping.pdf

Che, H., & Seetharaman, P. B. (2009). “Speed of replacement”: Modeling brand loyalty using last-move data. *Journal of Marketing Research*, *46*(4), 494–505.

Chierichetti, F., Kumar, R., Raghavan, P., & Sarlós, T. (2012). Are Web users really Markovian? In A. Mille, F. L. Gandon, J. Misselis, M. Rabinovich, & S. Staab (Eds.), *Proceedings of the 21st International Conference on World Wide Web* (pp. 609–618). ACM.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). *Introduction to algorithms* (3rd ed.). Cambridge, MA: Massachusetts Institute of Technology.

Dalessandro, B., Perlich, C., Stitelman, O., & Provost, F. J. (2012). *Causally motivated attribution for online advertising*. New York, NY. Retrieved from http://m6d.com/blog/wp-content/uploads/2012/02/m6d_causal_attribution_2012.pdf

Danaher, P. J., & Dagger, T. S. (2013). Comparing the relative effectiveness of advertising channels: A case study of a multimedia blitz campaign. *Journal of Marketing Research*, *50*(4), 517–534.

Econsultancy. (2012a). *Marketing attribution: Valuing the customer journey*. London, UK. Retrieved from http://econsultancy.com/de/reports/marketing-attribution-valuing-the-customer-journey

Econsultancy. (2012b). *Quarterly digital intelligence briefing: Making sense of marketing attribution*. London, UK. Retrieved from https://econsultancy.com/reports/quarterly-digital-intelligence-briefing-making-sense-of-marketing-attribution

Edelman, D. C. (2010). Branding in the digital age. *Harvard Business Review*, *88*(12), 62–69.

Evans, D. S. (2009). The online advertising industry: Economics, evolution, and privacy. *Journal of Economic Perspectives*, *23*(3), 37–60.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874.

Flosi, S., Fulgoni, G. M., & Vollman, A. (2013). If an advertisement runs online and no one sees it, is it still an ad? Empirical generalizations in digital advertising. *Journal of Advertising Research*, *53*(2), 192–199.

Ghose, A., & Yang, S. (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science*, *55*(10), 1605–1622.

Godfrey, A., Seiders, K., & Voss, G. B. (2011). Enough is enough! The fine line in executing multichannel relational communication. *Journal of Marketing*, *75*(July), 94–109.

Goldfarb, A., & Tucker, C. (2011). Online display advertising: Targeting and obtrusiveness. *Marketing Science*, *30*(3), 389–404.

Haan, E. de, Wiesel, T., & Pauwels, K. (2013). *Which advertising forms make a difference in online path to purchase?* (Marketing Science Institute Working Paper Series No. 13). Boston, MA. Retrieved from http://www.msi.org/reports/which-advertising-forms-make-a-difference-in-online-path-to-purchase/

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284.

Homburg, C., Steiner, V. V., & Totzek, D. (2009). Managing dynamics in a customer portfolio. *Journal of Marketing*, *73*(5), 70–89.

Jansen, B. J., & Schuster, S. (2011). Bidding on the buying funnel for sponsored search and keyword advertising. *Journal of Electronic Commerce Research*, *12*(1), 1–18.

Jaworski, B. J. (2011). On managerial relevance. *Journal of Marketing*, *75*(4), 211–224.

Jerath, K., Ma, L., & Park, Y.-H. (2014). Consumer click behavior at a search engine: The role of keyword popularity. *Journal of Marketing Research*, *51*(4), 480–486.

Jordan, P., Mahdian, M., Vassilvitskii, S., & Vee, E. (2011). The multiple attribution problem in pay-per-conversion advertising. In W.-Y. Ma, J.-Y. Nie, R. A. Baeza-Yates, T.-S. Chua, & W. B. Croft (Eds.), *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM.

Kireyev, P., Pauwels, K., & Gupta, S. (2013). *Do display ads influence search? Attribution and dynamics in online advertising* (HBS Working Paper No. 13-070). Boston, MA. Retrieved from http://www.hbs.edu/faculty/Publication%20Files/13-070.pdf

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (Vol. 2, pp. 1137–1145). San Francisco, CA: Morgan Kaufmann.

Lambrecht, A., & Tucker, C. (2013). When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research*, *50*(5), 561–576.

Leeflang, P. S. H., & Wittink, D. R. (2000). Building models for marketing decisions: Past, present and future. *International Journal of Research in Marketing*, *17*(2-3), 105–126.

Lehmann, D. R., McAlister, L., & Staelin, R. (2011). Sophistication in research in marketing. *Journal of Marketing*, *75*(4), 155–165.

Lewis, R. A., Rao, J. M., & Reiley, D. H. (2011). Here, there, and everywhere: Correlated online behaviors can lead to overestimates of the effects of advertising. In W.-Y. Ma, J.-Y. Nie, R. A. Baeza-Yates, T.-S. Chua, & W. B. Croft (Eds.), *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM.

Li, H. A., & Kannan, P. K. (2014). Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research*, *51*(1), 40–56.

Lilien, G. L. (2011). Bridging the academic – practitioner divide in marketing decision models. *Journal of Marketing*, *75*(4), 196–210.

Lilien, G. L., Roberts, J. H., & Shankar, V. (2013). Effective marketing science applications: Insights from ISMS-MSI Practice Prize finalist papers and projects. *Marketing Science*, *32*(2), 229–245.

Little, J. D. C. (1970). Models and managers: The concept of a decision calculus. *Management Science*, *16*(8), B-466–B-485.

Little, J. D. C. (1979). Decision support systems for marketing managers. *Journal of Marketing*, *43*(3), 9–26.

Little, J. D. C. (2004). Comments on "Models and managers: The concept of a decision calculus". *Management Science*, *50*(12 Supplement), 1854–1860.

Lodish, L. M. (2001). Building marketing models that make money. *Interfaces*, *31*(3), S45.

Lovett, J. (2009). *A framework for multicampaign attribution measurement*. Cambridge, MA. Retrieved from http://www.forrester.com/A+Framework+For+Multicampaign+Attribution+Measurement/fulltext/-/E-RES48148

Montgomery, A. L., Li, S., Srinivasan, K., & Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science*, *23*(4), 579–595.

Muthukrishnan, S. (2009). Ad exchanges: Research issues. In S. Leonardi (Ed.): *Vol. 5929. Lecture Notes in Computer Science, Internet and Network Economics* (pp. 1–12). Berlin, Germany: Springer.

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., & Mason, C. H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. *Journal of Marketing Research*, *43*(2), 204–211.

Neslin, S. A., & Shankar, V. (2009). Key issues in multichannel customer management: Current knowledge and future directions. *Journal of Interactive Marketing*, *23*(1), 70–81.

Osur, A. (2012). *The Forrester Wave: Interactive attribution vendors, Q2 2012*. Cambridge, MA. Retrieved from http://www.forrester.com/The+Forrester+Wave+Interactive+Attribution+Vendors+Q2+2012/fulltext/-/E-RES72681

Perlich, C., Dalessandro, B., Raeder, T., Stitelman, O., & Provost, F. J. (2013). Machine learning for targeted display advertising: transfer learning in action. *Machine Learning*, *95*(1), 103–127.

Pfeifer, P. E., & Carraway, R. L. (2000). Modeling customer relationships as Markov chains. *Journal of Interactive Marketing*, *14*(2), 43–55.

Provost, F. J., Fawcett, T., & Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. In J. W. Shavlik (Ed.), *Proceedings of the 15th International Conference on Machine Learning* (pp. 445–453). Morgan Kaufmann.

Raman, K., Mantrala, M. K., Sridhar, S., & Tang, Y. (2012). Optimal resource allocation with time-varying marketing effectiveness, margins and costs. *Journal of Interactive Marketing*, *26*(1), 43–52.

Reibstein, D. J., Day, G. S., & Wind, J. (2009). Guest editorial: Is marketing academia losing its way? *Journal of Marketing*, *73*(4), 1–3.

Riley, E. (2009). *The Forrester Wave: Interactive attribution, Q4 2009*. Cambridge, MA. Retrieved from http://www.forrester.com/The+Forrester+Wave+Interactive+Attribution+Q4+2009/fulltext/-/E-RES55372

Rutz, O. J., & Bucklin, R. E. (2011). From generic to branded: A model of spillover in paid search advertising. *Journal of Marketing Research*, *48*(1), 87–102.

Rutz, O. J., Trusov, M., & Bucklin, R. E. (2011). Modeling indirect effects of paid search advertising: Which keywords lead to more future visits? *Marketing Science*, *30*(4), 646–665.

Shao, X., & Li, L. (2011). Data-driven multi-touch attribution models. In C. Apté, J. Ghosh, & P. Smyth (Eds.), *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 258–264). ACM.

Sood, A., James, G. M., & Tellis, G. J. (2009). Functional regression: A new model for predicting market penetration of new products. *Marketing Science, 28*(1), 36–51.

Styan, G. P. H., & Smith, H. (1964). Markov chains applied to marketing. *Journal of Marketing Research*, *1*(1), 50–55.

The CMO Club & Visual IQ, Inc. (2014). *Big data, attribution & omni-channel*. Retrieved from http://www.visualiq.com/sites/default/files/PDF-download-graphics/study_building_bridges_to_the_promised_land.pdf

Tucker, C. (2012). The implications of improved attribution and measurability for online advertising markets. In *Competition in the Online Environment.* Retrieved from https://www.ipria.org/events/conf/Competition_Conference/Measureabiltiy_and_online_ads.pdf

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

Wiesel, T., Pauwels, K., & Arts, J. (2011). Marketing's profit impact: Quantifying online and off-line funnel progression. *Marketing Science, 30*(4), 604–611.

Wübben, M., & Wangenheim, F. von. (2008). Instant customer base analysis: Managerial heuristics often “get it right”. *Journal of Marketing*, *72*(3), 82–93.

Xu, L., Duan, J. A., & Whinston, A. B. (2014). Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science*, *60*(6), 1392–1412.

Yadav, M. S., & Pavlou, P. A. (2014). Marketing in computer-mediated environments: Research synthesis and new directions. *Journal of Marketing*, *78*(1), 20–40.

Yang, S., & Ghose, A. (2010). Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence? *Marketing Science, 29*(4), 602–623.

Appendix A

Table 8

Attribution results for second order model (data set 2)

	Preceding channel								
Current channel	START	Affiliate	Display	News-letter	Price Com-parison	Referrer	SEA	SEO	Type In
Affiliate	4.24%	2.54%	0.09%	0.12%	0.02%	0.14%	0.52%	0.40%	0.74%
Display	0.62%	0.19%	0.82%	0.23%	0.01%	0.06%	0.33%	0.29%	0.72%
Newsletter	4.61%	0.18%	0.22%	3.86%	0.00%	0.07%	0.79%	0.50%	1.84%
Price Comparison	0.07%	0.01%	0.00%	0.01%	0.03%	0.00%	0.03%	0.02%	0.03%
Referrer	1.26%	0.15%	0.04%	0.07%	0.01%	0.19%	0.16%	0.12%	0.18%
SEA	10.25%	0.64%	0.35%	0.89%	0.03%	0.24%	5.25%	1.97%	1.52%
SEO	6.25%	0.49%	0.16%	0.43%	0.02%	0.39%	1.93%	4.70%	1.13%
Type In	16.85%	1.08%	0.77%	2.49%	0.04%	0.35%	2.21%	1.69%	11.33%

Table 9

Attribution results for second order model (data set 3)

	Preceding channel							
Current channel	START	News-letter	Referrer	Re-targeting	SEA	SEO	Social Media	Type In
Newsletter	0.71%	0.29%	0.01%	0.00%	0.04%	0.04%	0.03%	0.04%
Referrer	4.07%	0.01%	0.77%	0.00%	0.28%	0.36%	0.35%	0.00%
Retargeting	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
SEA	11.07%	0.09%	0.49%	0.00%	3.74%	1.68%	1.00%	0.82%
SEO	12.23%	0.05%	0.35%	0.00%	1.61%	6.53%	0.62%	1.10%
Social Media	12.74%	0.03%	0.29%	0.00%	0.63%	0.58%	8.27%	1.16%
Type In	13.26%	0.06%	0.52%	0.00%	1.49%	1.72%	2.36%	8.14%

Table 10

Attribution results for second order model (data set 4)

	Preceding channel							
Current channel	START	Affiliate	Price Com-parison	Referrer	Re-targeting	SEA	SEO	Type In
Affiliate	0.28%	0.92%	0.18%	0.11%	0.03%	1.21%	0.24%	0.51%
Price Comparison	1.20%	0.02%	0.72%	0.03%	0.01%	0.57%	0.18%	0.06%
Referrer	1.03%	0.03%	0.02%	0.19%	0.00%	0.26%	0.03%	0.08%
Retargeting	0.10%	0.00%	0.03%	0.02%	0.41%	1.01%	0.06%	0.21%
SEA	39.91%	0.22%	0.41%	0.19%	0.28%	18.19%	1.72%	1.37%
SEO	3.47%	0.10%	0.13%	0.06%	0.03%	3.09%	2.26%	0.38%
Type In	7.43%	0.29%	0.16%	0.17%	0.20%	5.04%	0.83%	4.30%