

ATTRIBUTION MODELLING IN MARKETING: LITERATURE REVIEW AND RESEARCH AGENDA

Jitendra Gaur, Research Scholar, Indian Institute of Management Kashipur
Dr. Kumkum Bharti, Assistant Professor, Indian Institute of Management
Kashipur

ABSTRACT

The dilemma of allocation of marketing budget to multi-channels is faced by the practitioners and has remained an interesting area of research for the academia. This paper investigates and synthesizes different models and techniques used for allocation of marketing budget for three decades i.e. 1990-2019. Authors have used the PRISMA technique to create a corpus of relevant research articles. The classification of literature is done based on three broad categories i.e. constructs studied, usage of marketing channels, and models of attribution modelling. The analysis revealed that 71% of research articles were published in A and A category of ABDC ranked journals. Around 67% of the articles on attribution modelling were published during 2014-2019 making it one of the thrust areas for research in marketing discipline. Based on the synthesis of previous literature, the prominent models used by the researchers to solve the attribution problem were Markov chain, Probit model, Linear models among others. Lastly, authors provided a conceptual framework for the ensemble model combining the properties of the Markov chain model and Shapley value to make a robust model.*

Keywords: Online Marketing, Attribution Modelling, Ensemble Model, Omnichannel Marketing, Digital Attribution.

INTRODUCTION

Marketing is a complex phenomenon that includes the study of the target market, customer needs, integrated marketing, and profitability. According to Kotler (1999), companies pursue their marketing objectives by using a combination of various marketing tools known as marketing mix i.e. product, price, place (or distribution), and promotion. In the 21st century and specifically during the last decade of 20th century marketing activities are not limited to offline media. Hamil (1997) mentioned the explosion of business activity on the internet in the 90's. The first significant commercial activity on the internet took place in 1994 which catapult the commercialization of internet commerce in the next five years (Hoffman, 2000). The rising number of internet users also accredited to the steep rise in internet commerce activities. The current internet users in the world stand at 4.48 billion encompassing 58% of the world population (Statista, 2019). This means users are engaged in commercial and non-commercial activities via the internet. Therefore, marketers have estimated USD 365 billion of digital advertising spending with a year on year growth of 9.4% by the end of 2020. In this paper, for simplicity, we are going to use digital advertising and digital marketing interchangeably from now onwards. Digital marketing is a branch of marketing that deals with online media and, digital advertising is a subset of digital marketing. Kannan & Li (2017) have defined digital marketing as an adaptive, technology-enabled process by which firms collaborate with customers and partners to jointly create, communicate, deliver, and sustain value for all stakeholders.

To sustain value for all stakeholders the marketing performance should be measured. Of various decisions to be taken by a marketer, the dilemma of allocation of marketing budget for multi-channel marketing is daunting. In this study, multi-channel marketing can be considered as two-tier circles, where the inner circle is the conversion (ex. sale, signup) and the outer circle is customers in which each channel offers an independent and separate opportunity for conversion. Therefore, to address this dilemma, marketers have used attribution modelling (from now onwards, AM). Moffett et al. (2014) define attribution modelling, where advanced analytics is used to distribute suitable credit for the sale or conversion to each showcasing touchpoint of marketing channels. AM helps a marketer to calculate the right ROI in the multi-channel budget allocation decisions. Therefore, this paper investigates and synthesizes the different models and techniques used for attribution modeling published in the selected journals in past three decades i.e. 1990-2019. In the end, researchers also propose a conceptual framework to provide an alternative viewpoint to solve this critical dilemma from the mind of a marketer.

Rationale for this Study

Traditionally marketing performance was measured using accounting tools such as the balance sheet and income statements (Lavery, 1996). The author posit that firms should focus on returns on marketing (RoM) if marketing is considered as an investment. Of late, firms have started investing heavily in online advertising over offline advertising media as it facilitates ease of traceability, trackability, and the option to gather individual-level data. The adoption and growth of online advertising is also accredited to the benefits it offers such as scalability, measurability, and precise targeting of the customers. Goodwin (1999) stated that the internet encompasses the entire customer journey, wherein a customer journey is the full set of the interplay of a consumer on the website. In addition, Goodwin (1999) mentioned click-through rates, cost per click, cost per impression and return on investment as the four existing measures of online marketing effectiveness. Of which return on investment (ROI) is the most widely used indicator by marketers to evaluate online marketing effectiveness. Though, ROI has its origin from the accounting field, but it largely used by marketers to check the effectiveness of online advertising on ROI. Flamholtz (1985) characterizes ROI as a financial ratio that expresses profit in direct relation to investment. In online commerce, the ROI and attribution is jointly studied de Haan et al. (2016). Based on google trends from 2004-2019, the search term "*return on investment*" and "*attribution*" in the marketing field has seen a downward trend. Prior to 2004, the data to track the trends on these two terms is not available. It is evident from Figure 1 that a worldwide downward trend for the popularity of the term "*return on investment*" is observed. The 0-100 scale in the below graph depicts the competitiveness of the keyword. A higher scale signifies the competitiveness of a keyword.

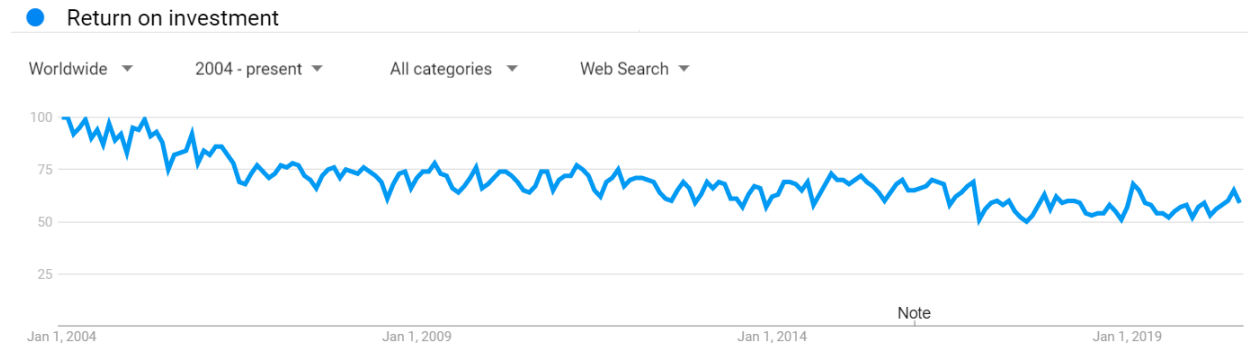


FIGURE 1
GOOGLE WORLDWIDE TREND FOR THE KEYWORD “RETURN ON INVESTMENT” SOURCE: GOOGLE TRENDS (WORLDWIDE, DEC 2019)

A probable reason for this downward trend could be the introduction and availability of more sophisticated methods that facilitate better estimation of the ROI and allocate the marketing budget. Botchkarev et al. (2011) suggested a few limitations of ROI. First, ROI focuses on the maximization of the ratio between returns and investments. But it does not lead to profit maximization. Second, ROI is a financial measure and focuses on profitability. Last, ROI does not reveal the systems’ effectiveness and efficiency.

According to Rentola (2014), ROI calculations can lead to sub-optimal decisions therefore for online advertising performance attribution modelling should be considered. Though the term ROI has seen a downward trend, in the same duration, the term “*attribution*” in the marketing field has gained popularity. Figure 2 shows the worldwide trend by Google for keyword “*attribution*” in the marketing field and shows that it has gained popularity in the field of marketing since 2011.

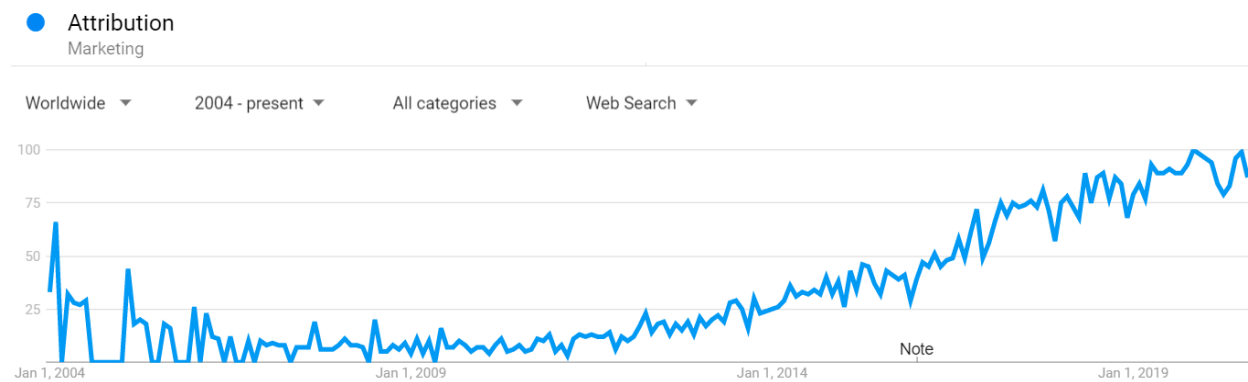


FIGURE 2
GOOGLE WORLDWIDE TREND FOR THE KEYWORD “ATTRIBUTION (MARKETING)” SOURCE: GOOGLE TRENDS (WORLDWIDE, DEC 2019)

The development of advanced statistical models for making the attribution modelling based decisions for allocation of marketing could be one of the reasons for the rising popularity of the keyword. The advanced statistical models of attribution are relevant in the digital

marketing as they provide user-level data, offer advanced tracking mechanisms for tracking online customers, increase the data storage capacity and enriched with advanced modelling techniques such as Markov chain, deep learning, vector autoregressive and others for estimation and predictions. AM is important for industry as it helps marketers to calculate the right ROI in multi-channel marketing Rentola (2014) decisions. Whereas, research on AM is relevant for academia as it is one of the top priorities for Marketing Science Institute conference MSI (2018, 2020). In this, attribution has been ranked as the number one priority by the Marketing Science Institute (MSI) for (2016-2018). Further, Gartner CMO spend survey 2018-19 mentioned that CMOs spend two third of their marketing budget on multi-digital marketing channels which surpasses spends on offline marketing. So, it becomes important to identify the right ROI in multi-digital marketing channels, for which AM is used by the marketers Rentola (2014). This is the motivation as well as the research gap to identify the various techniques used to solve the problem of AM. Therefore, the research problem is,

RP: How various techniques have been used to solve the problem of attribution modelling over the last three decades (1990-2019) by various authors?

Before 2005, researchers were using ROI as a measure for marketing performance (Rust et al., 2003; Montgomery et al., 2004). But after 2005 researchers used attribution modelling techniques such as Markov Chain (Yang & Ghose, 2010; Xu et al., 2014; Anderl et al., 2016a; Kakalejčik & Bucko, 2018) Linear models (Breuer et al., 2011; Rutz & Bucklin, 2007; Zhao et al., 2019) Probit model (Montgomery et al., 2004; Danaher & Dagger, 2013; Danaher & van Heerde, 2018) for allocation of budget to multi-channel marketing. The research objective one will help in solving the research problem.

Research Objective 1: To identify and synthesize the techniques applied by different researchers to solve the problem of attribution in the marketing literature.

Authors of the current research article observed that researchers Montgomery et al. (2004); Yang & Ghose (2010) and Danaher & Dagger (2013) have used a combination of modelling techniques also known as *ensemble modelling* techniques to solve the problem of attribution in online marketing. The authors also wanted to propose a combination of attribution models to offer a robust model to solve the attribution problem in multi-channel online marketing channels.

Research Objective 2: To propose a conceptual framework based on the ensemble technique.

Attribution modelling helps the marketer to calculate the right ROI for digital marketing. Kelly et al. (2017) define digital attribution as “*assigning credits to engagements happening before a completed conversion*”. Moffett et al. (2014) characterized attribution modelling where advanced analytics is used to distribute suitable credit for the sale/conversion to each showcasing touchpoint of marketing channels. It is critical to allocate the given budget to different marketing channels and thus the requirement to measure the marketing channel's effectiveness leads to the rise of statistical methods for attribution models. Attribution modelling upon combining offline marketing data with online marketing data could provide holistic marketing performance and can help marketing managers to solve the critical dilemma of finding out the right return on investment for the marketing channels.

However, is difficult to obtain and to combine the offline data at an individual level with individual-level data from online marketing channels. Thus, it becomes hard to gauge the customer coming from the offline marketing channel (Kannan et al., 2016). Therefore, in the current study, the authors have included only online marketing channel studies. The authors

choose to take at least two or more marketing channels under consideration for this study. Because in case a single marketing channel will be chosen by the firm then the ROI will be measured as the ratio of return and investment.

The structure of this article is as follows. In the first section, an introduction of the ROI concept and attribution modelling in marketing is given. In the second section, a thorough literature review on attribution modelling was undertaken. Here, literature was categorized based on constructs, the number of marketing channels, and usage of modelling techniques by authors. The third section of methodology explains the procedure for selection and exclusion of articles. In the fourth section, data analysis is carried out along with various studies conducted by different authors over the years. In the fifth section, the conceptual framework is explained followed by managerial implications. Lastly, thematic future agenda for the study is explained followed by a conclusion.

LITERATURE REVIEW

The review of the literature for studies on attribution modelling has been categorized based on the framework used by (Kannan et al., 2016). However, the classification of articles is done based on the authors' knowledge after thoroughly examining the corpus of research articles. Authors used three broad classification categories namely type of constructs studied in different articles, several marketing channels used i.e. whether the researchers studied multi-channel marketing or two-channel marketing. Lastly, authors studied the modelling techniques used by researchers over the time frame of 1990 to 2019 (Table 1).

Constructs Studied	Marketing Channel Used	Models Used
Effectiveness	Two Channels	Bayesian Framework
Cross Channel Advertising Effects	Multi-Channel	Regression
Carryover/Spillover Effects		Time Series
		Cooperative Game
		Deep Learning
		Ensemble Model

Effectiveness of Marketing Channels

The effectiveness of a marketing channel determines the budget allocation for each channel. This is because not all marketing channels respond similarly and have identical effects. Traditionally, it was believed that firm-initiated channels (FIC) such as emails and display advertisements underperform than the customer-initiated channels (CIC) such as search advertisements and price comparison advertisements; as it is believed that customer does not actively search for a product on the internet. Early work on media synergies measurement by Anderl et al. (2016a) has reported similar results. de Haan et al. (2016) examined the website funnel stages and advertising effectiveness and found CIC is 26.7 times more effective than FIC for revenue generation. de Haan et al. (2016) study was based on one retailer and was conducted for a short period, hence created a scope for further research in order to establish the generalizability of findings. The effectiveness of different channels within FIC and CIC was studied by (Breuer et al., 2011) and discovered that email has the strongest impact followed by display advertising and price comparison advertising. From the viewpoint of a practitioner,

display advertising is less effective, thus warrants reduced investments. However, Ghose & Todri (2015) measured viewability for display advertisements and found that mere exposure to the display advertisements without active user interaction with advertisement leads to an increase in the interest towards the brand of the advertiser by active searches. It was clear from studies on effectiveness that CIC especially search engine advertisements, email, a price comparison website, and display advertisements are more effective than FIC.

Cross Channel Advertising Effects Among Marketing Channels

Cross channel advertising effects has been measured using interaction effects, carry over effects and spill over effects by various authors. In regression, when more than one independent variable is present then it can have a simultaneous effect on the dependent variable. This effect is known as interaction effect (Dhar & Weinberg, 2016). Nottorf (2014) mentioned that an increase in intersession time leads to an increase in click probability. At the same time, with each additional exposure to display advertisements the chances of click probability decrease. In contrast to the belief that each marketing channel has a significant and positive interaction effect, Xu et al. (2014) found that display advertising has a relatively low effect on purchase conversion. However, mere exposure to display advertisements leads to an increase in search activity on search engines (Ghose & Todri, 2015; Kireyev et al., 2016). Interestingly, it was observed that search engine advertising does not lead to an increase in display advertisements. It is evident from the above studies that display advertisements have a minute and positive interaction effects on other marketing channel's performance. A marketer should not completely stop display advertisements however they should judiciously spend on display advertisement.

Carryover and Spillover Effects Among Marketing Channels

Breuer & Brettel (2012) define carryover effect "*as the percentage of advertising effect carries over from time period t to time period $t+1$.*" Spillover effects are when the promotion of one product of a firm leads to the sale of another product of the same firm (Wei et al., 2011). The measurement of marketing channel effectiveness is challenging because of the presence of spillover and carryover effects. These carryovers and spillover effects influence the short term and long-term effectiveness of marketing channels. Li & Kannan (2014) noticed significant spillover and carryover effects at the visit stage and purchase stage of sales funnel. Therefore, the authors suggest that neither the last-click attribution model nor seven-day average measures are right to estimate the real impact of advertising. Danaher & van Heerde (2018) came up with a fixed budget profit maximization (FBPM) model considering the carryover and spillover effects of marketing channels. The analysis showed that the FBPM model outperformed the last touch attribution (LTA) and the probit model.

Studies with two Marketing Channels

According to the author's knowledge and extant literature on digital marketing the adoption of more than one marketing channel started to register in research journals in 2010. Yang and Ghose (2010) studied the interdependence of sponsored search and organic search and analysed whether positive, negative, or zero relationship exists between sponsored search and organic search. Kireyev et al. (2016) studied whether display advertisements impact the search advertisements; and proposed the estimation of dynamic interactions among display

advertisements and search advertisements. Both studies revealed the presence of significant and positive interdependence among different marketing channels. Also, as expected in multi-channel marketing interaction and spillover effects were noticed and significant.

Studies with Multi-Channel Online Advertising

In multi-channel online marketing, firms often use more than two marketing channels to promote their offerings. Because, customers, usually refer more than one marketing channel to find the details of the offering with the help of online reviews, ratings, and other means before making the final purchase. Li & Kannan (2014) estimated the spillover and carryover effects of multi-channel online advertising and provided a conceptual framework for attributing and allocating the credit for conversion incorporating spillover and carryover effects for FIC and CIC using individual-level data. Anderl et al. (2016b) studied the performance of the digital marketing channels and measured the effect of the performance of one marketing channel over others. Authors found that visitors who first used FIC and later used CIC have shown an increased purchase propensity; whereas visitors who switched from branded search to generic search have decreased purchase propensity. Danaher & van Heerde (2018) also measured the carryover effects and interaction effects of online marketing channels. These authors proposed that attribution should not be based on the number of exposures to marketing channels but use channel effectiveness to determine attribution. Kakalejčík & Bucko (2018) analysed the multi-channel paths using a Markov chain model and compared it with other heuristic models. Authors found that majority of purchases came from direct traffic; also 40% of purchases were registered by customers in less than 5 steps of their customer journey. From all the above studies, it is evident that spillover, carryover, and interaction effects play an important role in multi-channel online advertising. Visitors moving from generic to branded searches are more probable to purchase similarly customers with fewer steps in customer journey i.e. fewer than 5 steps have higher purchase propensity which decreases with, an increase in the number of steps in the customer journey.

Probabilistic model with a Bayesian Framework

Probabilistic models use probability theory which factors in uncertainty rather than ignoring it. The Bayesian framework allows predictions to be made using statistical models with the use of probability incorporating uncertainty and noise within the model. Rutz & Bucklin (2007) developed a dynamic linear model based on a Bayesian framework to measure the spillover effects from generic to branded keywords in search engine advertising. Yang & Ghose (2010) used a hierarchical Bayesian modelling framework and assessed the model utilizing the Markov Chain Monte Carlo technique to investigate the relationship of organic and paid search listing. Nottorf (2014) used the Bayesian framework to develop a binary logit choice model which was used to measure the interaction effects among various channels. Xu et al. (2014) also used a Bayesian framework in conjunction with Markov Chain Monte Carlo technique similar to what Yang and Ghose did in their study. Overall, it was evident from the studies that a positive relationship exists in organic and paid searches, generic and branded keywords, display, and search advertisements. Spillover from generic keywords leads to branded keyword searches also, these effects reduce from step 3 and onwards of the customer journey. Extra exposure to display advertisement results decrease in click probability. Search advertisements have a positive effect on display advertisements and vice versa but with a small positive effect.

Probabilistic Model with Regression (Probit, Logit, Tobit)

In the studies of attribution modelling, seldom, a probabilistic model is combined with regression model. A regression model is an equation that describes the average relationship between the independent and dependent variables. Thus, to solve the attribution modelling problems, different probabilistic models, based on regression including probit, logit, and tobit were used. Rust et al. (2003) used a multinomial logit model to estimate the return on marketing and to develop a new customer lifetime value (CLV), model. The new CLV model developed by the authors permits the modelling of brand switching patterns and competitive effects. Montgomery et al. (2004) studied latent utility by using the dynamic multinomial probit model with the introduction of a vector autoregressive model to capture dynamics in choice. Breuer et al. (2011) estimated carryover effects and interaction effects using a direct aggregation model and GLS regression. Contrary to other authors they did not find any interaction effects among online marketing channels. Danaher & Dagger (2013) studied purchase incidence, purchase outcome, dollar sales, and profit by using the Type II tobit model. The estimation of purchase incidence was done using probit model, whereas, the estimation of the purchase outcome was done by a tobit model. Nottorf (2014) studied the interaction effects among various channels by using a binary logit choice model (BLCM). Danaher & van Heerde (2018) studied interaction effects, spillover effects, and carryover effects by using a probit model to derive the fixed budget profit maximization (FBPM) allocation rule. The probabilistic model developed by different researchers using probit, logit, or tobit models performed better predictions compared to baseline predictions. Strong carryover effects were observed in email, followed by display advertising and lastly by PCA.

Probabilistic model with Time Series

Probabilistic time series models are used to find the predictive distribution for a value of time series at future points from the model. Researchers have included time series in the probabilistic model to solve the attribution modelling problem and captured the dynamics in choice in the model. de Haan et al. (2016) examined the effectiveness of various online marketing channels, and wherein sales funnel the effects of an online marketing channel is strongest. The authors used the vector auto-regressive (VAR) model for aggregate level time series and found that customer-initiated contact (CIC) was 26.7 times more effective than the firm initiated the contact. When CIC was additionally split then it was discovered that content separated activities were 9-10 times increasingly powerful and content integrated activities were 44.3 occasions more successful than FIC for revenue generation.

Model with Cooperative Game Theory

Cooperative game theory alternatively known as coalitional game theory, models the interaction of decision-makers that focuses on the group behaviour of players. Berman (2018) conducted a study to develop payment and measurement schemes that result in yield improvement to the advertisers. The author used the cooperative game-theoretical model based on the shapely value which allocates value among cooperative game players. Efficiency, symmetry, pay to play and marginality are the four properties of shapely value. The findings revealed that cost per thousand impressions (CPM) campaigns outperform the cost per thousand acquisition (CPA).

Probabilistic Model with Deep Learning

Deep learning models are based on neural networks and suitable for solving the problems even with humongous data that is generated online daily. Arava et al. (2018) proposed a data-driven attribution and conversion prediction model which they named as Deep Neural net with Attention for Multi-Touch Attribution (often called DNAMTA). The authors stated that they were the first ones to use a deep learning algorithm to solve the multi-touch attribution problem in marketing. Arava et al. (2018) compared DNAMTA with models such as logistic regression, LSTM, and last touch attribution and found that DNAMTA outperformed all other models. However, more channels and data from different industries would help in proving generalizability using DNAMTA.

Ensemble Techniques

The ensemble modelling technique is a combination of two or more performing models or classifiers that jointly increase prediction accuracy. Chatterjee et al. (2015) in their study found that the chosen ensemble method outperforms with an accuracy of 97% over all other models. Dietterich (2000) gave three reasons that make an ensemble model outperform any single classifier. First, the ensemble model is statistically robust. Since the ensemble method takes an average vote from each classifier, therefore, there are fewer chances of picking up a wrong classifier. Second, ensemble models are better off because of computational reasons. Lastly, ensemble models perform better because of representational reasons. With the weighted sums of the hypothesis, it becomes possible in the ensemble model to expand the representable functions space. In addition to these three reasons, the ensemble model also enjoys merit over standalone models because it carries the potential to reduce the generalization error (it measures the accuracy of the algorithm prediction for unseen data). One of the reasons for the reduction of the generalization error in the ensemble model is due to a reduction in prediction errors as long as base models are diverse and independent (Kotu & Deshpande, 2015).

Paradigm Shift in Attribution Modelling

Customer lifetime value (CLV) and return on investment

This study spans 29 years starting from 1990-2019. During the first fifteen years of the study i.e. during 1990-2005, academic research was concentrated on customer lifetime value (CLV) and return on investments (ROI). Rust et al. (2003) in their seminal paper discussed ROI and CLV calculations for the marketing field using customer equity. However, ROI calculations are not straight forward as it requires longitudinal and historical data, so ROI and CLV measures were rare until the 1990s. Rust et al. (2003) provided the first broad framework to measure CLV by incorporating the competitive impacts and brand switching patterns. These authors suggested that future studies should investigate customer equity and researchers should develop the CLV dynamic models to understand how customer value changes over time.

Heuristic models

During the period, 2006-2015 google introduced heuristic models to measure the right ROI for advertising source/medium. Seven attribution models were developed by Google Inc. namely linear, last interaction, first interaction, time decay, position-based, last non-direct

interaction, and last AdWords interaction. These seven attributions models are popular among practitioners and academia involved in online marketing.

Attribution models

Between 2010-2019, researchers have developed several attribution models using a statistical approach. With the advent of more consumers moving towards online platforms for search and purchase, this trend of developing new models of AM is continuing. Marketing Science Institute (2002-2004) has kept “*assessing marketing productivity*” (return on marketing) and “*marketing metrics*” as its top priority in Table 2.

Dimensions	Authors (Year)	Definitions
Carryover effect	Li & Kannan (2014)	The visit involvement can impact ensuing visits to the site through the same channel as well as conceivable transformations through that channel.
Customer Equity	Rust et al. (2003)	The cumulative sum of organizations’ present and prospective customers lifetime values.
Customer Lifetime Value	Berger & Nasr (1998)	The cumulative discounted order values of past, present, and future order values.
Efficiency	Berman (2018)	If two distributors play the game, the method will trait all transformations to the two distributors.
Exciting effects	Xu et al. (2014)	The event of a prior advertisement click influences the likelihood of an event of consequent promotion clicks.
Interaction effect	Rosnow & Rosenthal (1989)	The simultaneous impact of two or more independent variables is significantly larger and lesser than the sum of the individual effect on a dependent variable.
Marginality	Berman (2018)	Publishers who contribute more to the transformation will get higher attribution.
Pay to play	Berman (2018)	Publisher will receive zero attribution if no advertisement shown by the publisher.
Spillover effect	Li & Kannan (2014)	A visit may lead to visits and conversions via different channels
Symmetry	Berman (2018)	In an event, if both publishers show off equal effort ($q_1 = q_2$) then they will receive equal attribution.
Viewability	Ghose & Todri (2015)	If the impression was visible for more than one second on a consumer’s screen area, then it is rendered viewable.

METHODOLOGY

For this research article, full length, peer-reviewed research papers published in the English language in scholarly journals between 1990 and 2019 were selected. The choice for catchphrases for looking through articles came after a careful assessment of the relevant definition of attribution modelling in marketing and their potential equivalent words in the academic literature. “*Attribution Model*”, “*Online Marketing*” and “*Omnichannel Marketing*” were chosen as the search phrases in titles, abstracts, and keywords of the research articles. The databases chosen were Web of Science, EBSCO, Scopus, and Google Scholar. These databases were chosen as it covers the greater part of the scholarly journals pertinent to the theme under consideration and aligned with the work of (Guo et al., 2019). The article selection was done

using the PRISMA technique as mentioned in Figure 3. Two additional criteria were added for shortlisting the relevant research articles. First, an article must be accessible on the web and preferably published in the Australian Business Deans Council (ABDC) ranked journals. Second, a research article should have at least one citation if it is not published in the ABDC ranked journal. The initial screening gave 335 articles and ended with 21 research papers for the final review. The exclusion criteria for articles are mentioned in the Figure. 3 Consequently, the 21 research papers were read for descriptive analysis and arranged with an informational index of 20 fields, including the name of the journal, year of publication, type of study, and key constructs, among others.

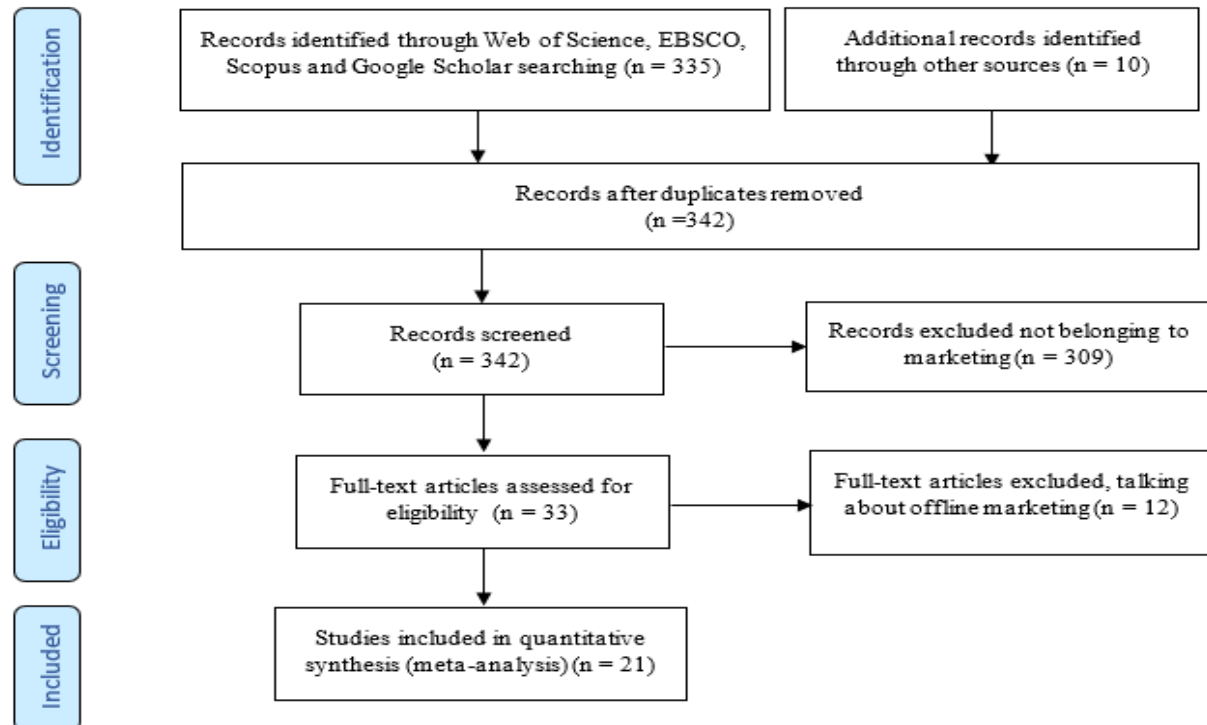


FIGURE 3
PRISMA TECHNIQUE FOR ARTICLE SELECTION

DATA ANALYSIS

Synthesis of Papers

It has been identified that 81% of the articles (17/21) were listed in ABDC (2019 Journal Quality List), and most prominently 71% of research articles were published in either A* or A category journals (specifically, 14 in A* and 1 in A category). This can be, thus, inferred that the research topic is of interest to the high rated journals.

Journal Wise Breakup

Our analysis revealed that about 19% of the articles on attribution modelling were published in the Journal of Marketing Research followed by 14% each in Marketing Science and

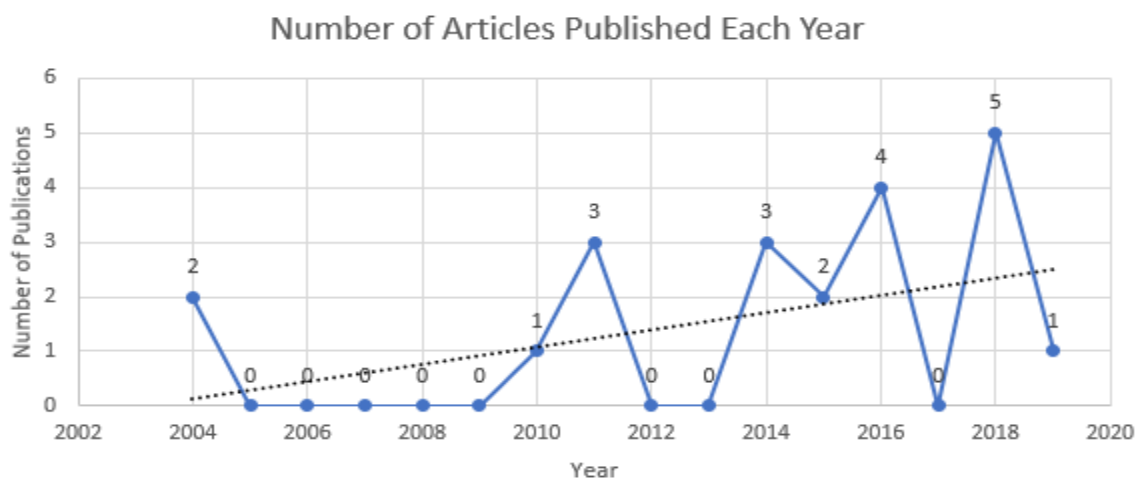
International Journal of Research in Marketing. Though attribution modelling articles were mostly published in the marketing journals but retailing, information science, and finance were few more domains that published articles on AM in Table 3.

Journal	2004	2005-2009	2010	2011	2013	2014	2015	2016	2017	2018	2019	Total
International Journal of Market Research											1	1
ai.google										1		1
arXiv preprint arXiv:1809.02230 (Cornell University)										1		1
Electronic Commerce Research and Applications						1						1
International Journal of Research in Marketing								3				3
Journal of Applied Management and Investments										1		1
Journal of Marketing	1											1
Journal of Marketing Research				1	1	1				1		4
Journal of Retailing								1				1
Management Science						1						1
Marketing Letters				1								1
Marketing Science	1		1				1					3
MIS Quarterly							1					1
Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining				1								1
Total	2	0	1	3	1	3	2	4	0	4	1	21

The maximum number of articles published on attribution problems is four by the Journal of Marketing followed by three publications each in the Journal of Marketing Research, International Journal of Research in Marketing and Marketing Science. One article each was published in the International Journal of Market Research, Journal of Marketing, Journal of Retailing, Marketing Science, Marketing Letters, MIS Quarterly, Journal of Applied Management and Investments, ai.google., and arXiv (Cornell University). A single article is published in the proceedings of 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining was also made part of this review process as it attempts to solve the attribution problem with the latest modelling technique of deep learning which is based on neural networks.

Year-Wise Articles Publication

Year-wise analysis suggests that attribution modelling is a new topic and roughly 90% of the articles were published in the last 10 years. The findings of this study reveal that before this emergent trend of research on attribution modelling, researchers were conducting studies on CLV/ROI, especially during 1990-2005 to study the impact of advertisement spending on performance. In this paper, two seminal studies on CLV and ROI were used to showcase the paradigm shift from 1990-2005. But, post-2005 with the development of heuristic models for attribution modelling by Google, a new avenue for calculation of ROI for online marketing channels were opened. The use and application of heuristics models not only attracted the attention of practitioners but also of academia. Resultantly, in the last five years, a 66.6% of the articles on attribution modelling were published across journals. The black dotted trendline (Figure 4) shows a progressive upward moving trend for publication of articles in recent years.



**FIGURE 4
YEAR-WISE ARTICLES PUBLISHED**

Models Used by Overall and Year-Wise

Of the total research papers, 90% of the research papers were quantitative and 10% were conceptual papers. The most frequently used models for attribution modelling in these research papers were Markov Chain (18%), probit model (14%), linear regression (14%), logit model (9%), and vector auto regressive model (9%). Further, 4% of the articles used game-theoretical model, ARW algorithm, Type II tobit model, three-level measurement model, incremental lift approach, persistence modelling, proportional hazard model, and DNAMTA.

In the next section, the frequency of the various attribution models used in the articles is provided accompanied by the author and year information. The analysis revealed that the Markov chain model has been used by 18% of the articles and still popular for predicting the attribution values. The below table suggests the number of articles published during the particular year, constructs studied, and the objective for which the model was used in Table 4.

Authors (Year)	Models	No. of Articles	Objective of the study
2004		2	
Rust et al. (2003)	Logit Model	1	To propose a model of CLV by consolidating the effect of brand exchange and offerings by competitors.
Montgomery et al. (2004)	Probit model	1	To demonstrate empirically that succession of web viewings is educational in foreseeing a client's path.
2010		1	
Yang & Ghose (2010)	Markov chains	1	To discover the effect of search engine advertising on consumers' reactions in the presence of organic listings of the same firm.
2011		3	
Papadimitriou et al. (2011)	ARW Algorithm	1	To study display advertisement, social influence and search lift using a controlled experiment.
Rutz & Bucklin (2011); Breuer et al. (2011)	Linear Model	2	To study whether spillover occurs from generic search to branded search.
2013		1	
Danaher & Dagger (2013)	Type II Tobit model	1	To investigate the relative effectiveness of marketing channels.
2014		3	
Nottorf (2014)	Logit Model	1	To explore the interaction effects among different channels and clarify shopper online behaviour conduct.
Xu et al. (2014)	Markov chains	1	To capture the exciting effects among advertisement clicks and to measure the effectiveness of online marketing.
Li & Kannan (2014)	Three-level measurement model	1	To build a coordinated model to assess the carryover and spillover effects of earlier contacts of various channels, and to help select the optimizing marketing budgets.
2015		1	
Ghose & Todri (2015)	Incremental lift approach	1	To examine more than one interaction between display ads and online search behaviour.
2016		4	
Anderl et al. (2016a)	Markov chains	1	To quantify the worth and overall performance of digital channels and measure how one digital channel affects the performance of another channel and generalizes it.
Kireyev et al. (2016)	Persistence modelling extends multivariate time series methods	1	To estimate the dynamic associations from search and display spends to sales, and also among display and search channels.
Anderl et al. (2016b)	Proportional hazard model	1	To propose a scientific classification-based methodology for marketing

			channels that are based on the dimensions of contact origin and brand usage.
de Haan et al. (2016)	Vector autoregressive (VAR)	1	To study the relative efficacy of various online marketing channels; how long the impacts last, and where the impacts are more rooted in the funnel?
2018		3	
Arava et al. (2018)	Deep Neural Net with Attention multi-touch attribution model (DNAMTA) model	1	To propose information-driven multi-touch attribution and conversion prediction model (DNAMTA) that outperforms different approaches.
Kakalejčik & Bucko (2018)	Markov chains	1	To characterize the present condition of multichannel attribution and, given the literature, to inspect the information gathered from a chosen company by utilizing the Markov chains approach.
Danaher & van Heerde (2018)	Probit model	1	To propose an attribution definition dependent on the relative incremental contribution made by each medium to purchase, considering interaction and carryover effects.
Berman (2018)	Game theoretical model (Shapley value)	1	To establish measurement and payment schemes that reduce the impact of moral hazards and asymmetric knowledge and result in improved results for the advertiser.
Kelly et al. (2018)	Conceptual Study	1	To evaluate additional conversions generated by a single ad channel.
2019		1	
Zhao et al. (2019)	Linear Model	1	To propose several attribution modelling methods that measure how revenue ought to be attributed to online marketing channels.

Collaboration of Authors by Countries

We have analyzed the authors from various countries, who are collaborating on the AM topic. It is evident from below Figure 5 that the USA is the epicentre of research in attribution modelling and spearheading by having collaboration with countries like Germany, Turkey, Switzerland, Netherlands, and Australia. Size of the country name in Figure 5 shows the number of articles published by the countries. In Figure 5, USA has the biggest size as most researches on AM topic prior to 2015 was done in USA and the new authors from other countries were either doing joint researches or few independent researches. However, with spread of the knowledge about attribution modeling, after 2015 authors from Germany and Turkey started collaborating with USA. Germany is also collaborating with other countries such as Switzerland and Netherlands. Whereas, authors from Netherlands also collaborated with the authors from Australia and Turkey. In days to come with percolation of attribution modeling in various school of thoughts we might see more countries collaborating on this research topic.

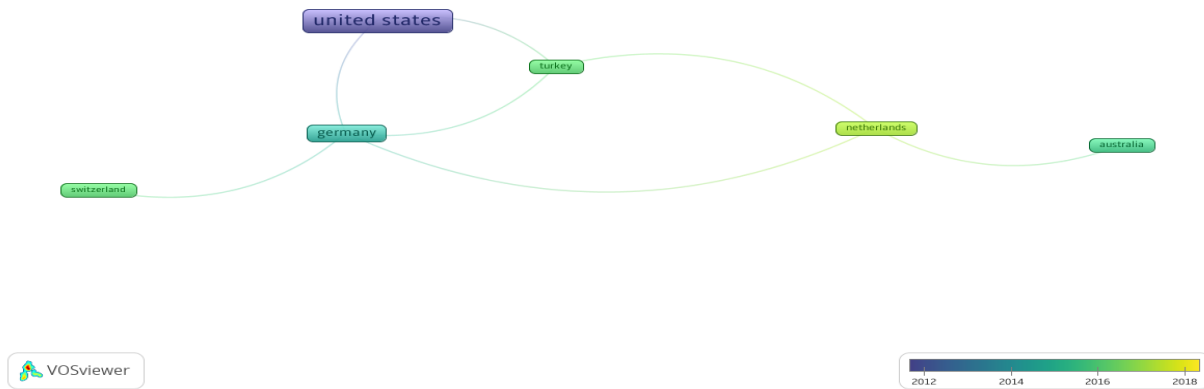


FIGURE 5
AUTHORS COLLABORATION BY COUNTRIES ON ATTRIBUTION MODELING

Keyword Analysis of Research Articles on Attribution Modelling

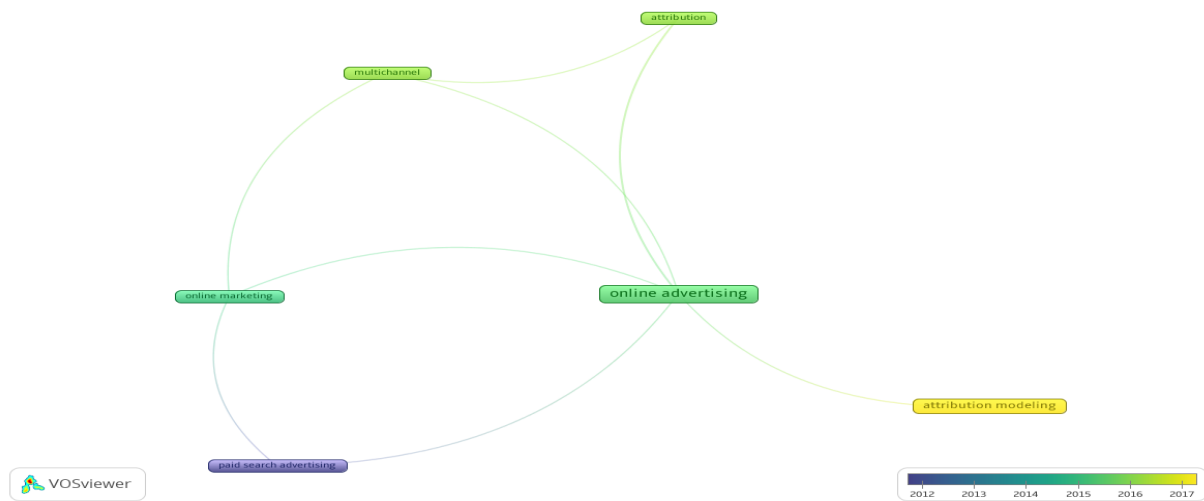


FIGURE 6
KEYWORDS USED IN ARTICLES ON ATTRIBUTION MODELING

Authors have conducted keyword analysis of the research articles. The analysis was conducted using the VOS viewer. Figure 6 suggests that before 2015 authors were concentrating on topics such as paid search advertising, online marketing, and online advertising. After 2015 the focus of the authors changed to multichannel, attribution and attribution modelling. In the days to come more research will be focused on omni-channel marketing, multi-channel marketing, and attribution modelling as we can see the shift from one or two channels to multi-channel marketing.

Conceptual Framework for Ensemble Modelling

An ensemble model is a machine learning model created by combining multiple algorithms to provide better prediction performance Qiu et al. (2014). The authors proposed an ensemble model by combining the results from the Markov chain model and Shapley value. Markov chain is a probabilistic model used for attribution modelling (Yang & Ghose, 2010; Xu et al., 2014; Anderl et al., 2016a; Kakalejčik & Bucko, 2018). A high adoption rate of the Markov model in AM studies is due to its flexibility and parsimonious nature. The transition probabilities calculated with the Markov chain model can correctly attribute every touchpoint used for marketing. Further, Markov chain model with removal effects is useful to measure the individual contribution of marketing channels. In removal effects, contribution of each channel is calculated by removing it from the customer journey and observe how many conversions happen without that channel. Markov models visualize customer journeys as chains, for example, the first-order Markov model depends only on the previously visited state and so on for higher-order Markov models. Shapley (1953) defined the Shapley value as a cooperative game theory that is used to distribute the value of the payoff among the players. Berman (2018) also used Shapley value in his research paper to study interaction effects among various channels. The Shapley value has the four desirable properties namely, efficiency, symmetry, null player, and marginality.

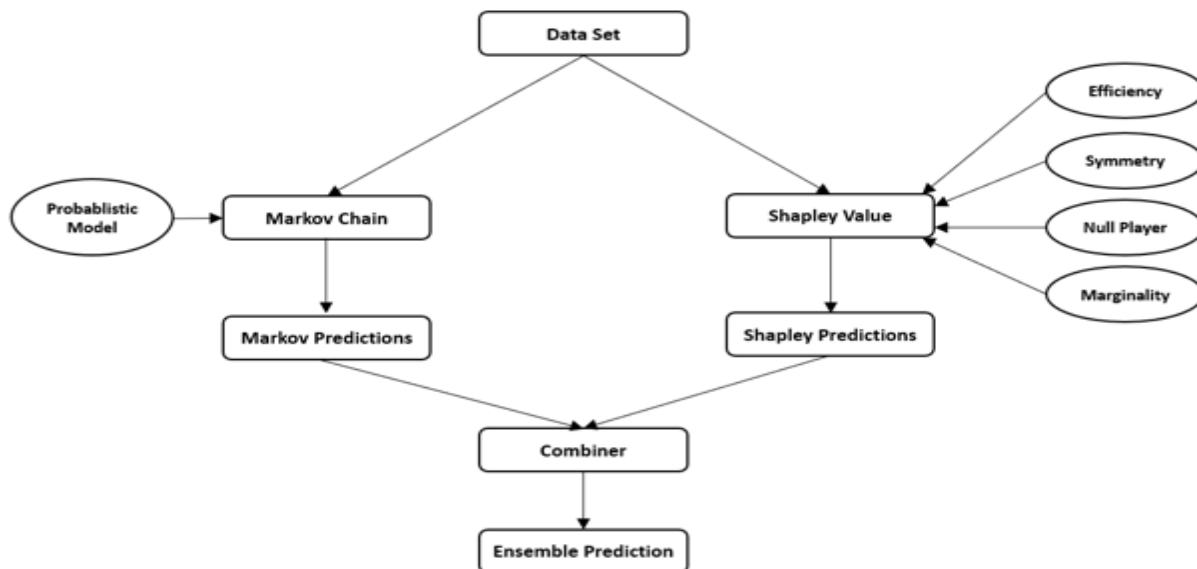


FIGURE 7
CONCEPTUAL FRAMEWORK FOR ENSEMBLE MODELING

As explained earlier, ensemble model is a combination of multiple algorithms. In current case, authors combined Markov chain model which is the most used modelling technique based on probabilistic model and Shapley value based on the model properties could provide a robust ensemble model. The ensemble model outperforms a single model as the ensemble model reduces the generalization errors. In Figure 7 authors has used unidirectional arrows in the

conceptual model. On right hand side of the conceptual model efficiency, symmetry, null player and marginality are the important properties of Shapley value which provide its uniqueness whereas Markov chain is a probabilistic model. Thus, combining two different modelling techniques shall cover the deficiency of other modelling technique making it more robust.

Research Implications

The research implications have been subdivided into two categories.

Implications for Managers

Our ensemble model suggests that combining the predictions from the Markov chain probabilistic model and Shapley value would be a robust model than a standalone model. It is of utmost necessity in the multi-channel environment to set up the right marketing budget for each marketing channel. Managers will be able to better allocate the budget among better performing online channels based on ROI generated from the concerned marketing channel with the proposed model. It will help in reducing the budget allocation to the non-performing channels. The proposed attribution model will help in facilitating the performance-based advertising for the various marketing channel.

Implications for A Researcher

In academia, most of the studies have used standalone modelling techniques to solve the attribution modelling problem. Ensemble models have gained popularity where researchers combine two or more models and overcome the disadvantages of the standalone model. The ensemble model has opened the new avenue for research altogether. The first click/ last click heuristic models were not the right measure for budget allocations. In such a scenario the ensemble model with combined statistical model techniques should prove a more robust model for the researchers. According to the author's knowledge, the current research paper is one of the most comprehensive studies synthesizing relevant articles on attribution modelling in online marketing. The research paper could be used by researchers aiming to solve the attribution problem in online marketing as a comprehensive study to know the work done until the year 2019.

Future Research Agenda

In the attribution modelling body of literature, thematically, a large portion of the existing research is aimed at solving the online attribution problem with the help of heuristic models.

Thematic agenda point 1) The future online attribution research should focus on granular user-level data. In the current study, the authors observed that all empirical studies do not have granular user-level data. The future scope should include consumer-level data from multiple advertisers rather than just one advertiser (Nottorf, 2014; Yang & Ghose, 2010). With the advent of technology, more granular-level customer data over a period is available for analysis.

Thematic agenda point 2) The future online attribution research should focus on generalizability. Notably, online attribution in the marketing field is concentrated mainly in the domains of the retail and advertising industry. Authors such as (Danaher & Dagger, 2013; de Haan et al., 2016; Montgomery et al., 2004; Ghose & Todri, 2015) mentioned that future

researches should focus on generalizability. Montgomery et al. (2004) mentioned that their study included data from one online retailer for one-month time period, so expansion is warranted to compare the results. de Haan et al. (2016) mentioned that generalizability is one of the primary concerns for their study and advocated that time-varying parameters should also be checked in future studies.

Methodological agenda point 1) The future research in online attribution should incorporate realistic cost estimates data. Future researches should include realistic cost estimates as highlighted by many authors such as (Li & Kannan, 2014; Xu et al., 2014; Anderl et al., 2016a; Danaher & van Heerde, 2018). Anderl et al. (2016a) mentioned that future studies could include revenue and profit data to measure CLV. Danaher & van Heerde (2018) suggested that future investigations ought to include realistic cost estimates for online marketing channels.

Methodological agenda point 2) The future research in online attribution should measure interaction effects, carry-over effects, and long-term impacts. The authors also suggested that in future studies long terms sales effects should be included in empirical models. Anderl et al. (2016a) mentioned that future studies should test the framework using offline data also the potential of interaction effects for longer customer journeys to be measured.

Methodological agenda point 3) The future research in attribution modelling should focus on ensemble modelling. According to the authors, understanding, and analysis combining two or more modelling techniques can be useful to overcome the drawbacks of a model and prove to be a more robust model.

CONCLUSION

In this research, the authors have presented the paradigm shift (from ROI/CLV calculation to heuristic models to attribution models) in the last three decades to measure the effectiveness of online advertising spending. It started with the measurement of CLV and ROI measures. Later, with the introduction of a multi-source/medium environment, the attribution model came to the rescue of the marketer to measure marketing effectiveness as the measurement is not left that straightforward and simple. Google came up with some heuristic models such as first click, last click, linear model and others to help practitioners. Academia has also shown interest in the field of attribution starting from the year 2010. Researchers have used various models and techniques to solve the attribution problem in marketing and the author's analysis revealed that the Markov chain model is most prominently used by academia followed by the linear model, probit model, logit model, vector autoregression model, and others respectively. Authors presented the attribution studies done in last three decades, the constructs studied and with the analysis it was found that previous studies were focused on CLV and ROI measurement but during last 10 years and more specifically in last 5-6 years the academia focused on solving attribution problem in marketing with the use of statistical advanced techniques. Authors apart from investigating of past research articles also proposed a conceptual framework based on the ensemble model combining the power of Markov chain and Shapley value.

Limitations

The authors built the corpus after searching many databases and included the relevant studies. The studies on offline media were excluded from the current study scope as it is difficult to track the customers researching online and buying offline or vice versa. Future research can

include studies for offline media also. Also, authors considered ABDC journals list as inclusion criteria for the corpus so, further research articles from other journal, proceedings not listed in ABDC journals covering the topic could be incorporated in future study. The conceptual framework proposed by the authors is robust as per their understanding. However other possible alternative models could also be analysed, and results compared with the current proposed model and with other heuristic models. Future research can also test the results across various industries and can check if the results are stable and generalizable.

REFERENCES

- Anderl, E., Becker, I., Von Wangenheim, F., & Schumann, J.H. (2016a). Mapping the customer journey: Lessons learned from graph-based online attribution modeling. *International Journal of Research in Marketing*, 33(3), 457-474.
- Anderl, E., Schumann, J.H., & Kunz, W. (2016b). Helping firms reduce complexity in multichannel online data: A new taxonomy-based approach for customer journeys. *Journal of Retailing*, 92(2), 185-203.
- Arava, S.K., Dong, C., Yan, Z., & Pani, A. (2018). Deep neural net with attention for multi-channel multi-touch attribution. *arXiv preprint arXiv:1809.02230*.
- Berger, P.D., & Nasr, N.I. (1998). Customer lifetime value: Marketing models and applications. *Journal of Interactive Marketing*, 12(1), 17-30.
- Berman, R. (2018). Beyond the last touch: Attribution in online advertising. *Marketing Science*, 37(5), 771-792.
- Botchkarev, A., Andru, P., & Chiong, R. (2011). A Return on Investment as a Metric for Evaluating Information Systems: Taxonomy and Application. *Interdisciplinary Journal of Information Knowledge and Management*, 6.
- 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.
- Chatterjee, S., Dash, A., & Bandopadhyay, S. (2015). Ensemble support vector machine algorithm for reliability estimation of a mining machine. *Quality and Reliability Engineering International*, 31(8), 1503-1516.
- 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.
- Danaher, P.J., & Van Heerde, H.J. (2018). Delusion in attribution: Caveats in using attribution for multimedia budget allocation. *Journal of Marketing Research*, 55(5), 667-685.
- De Haan, E., Wiesel, T., & Pauwels, K. (2016). The effectiveness of different forms of online advertising for purchase conversion in a multiple-channel attribution framework. *International Journal of Research in Marketing*, 33(3), 491-507.
- Dhar, T., & Weinberg, C.B. (2016). Measurement of interactions in non-linear marketing models: The effect of critics' ratings and consumer sentiment on movie demand. *International Journal of research in Marketing*, 33(2), 392-408.
- Dietterich, T.G. (2000). Ensemble methods in machine learning. In *International workshop on multiple classifier systems* (1-15). Springer, Berlin, Heidelberg.
- Flamholtz, E. (1985). *Human Resource Accounting: Advances in Concepts, Methods and Applications*. San Francisco: Jossey-Bass.
- Ghose, A., & Todri, V. (2015). Towards a digital attribution model: Measuring the impact of display advertising on online consumer behavior. *Available at SSRN 2672090*.
- Goodwin, T. (1999). Measuring the effectiveness of online marketing. *Market Research Society. Journal.*, 41(4), 1-6.
- Guo, F., Ye, G., Hudders, L., Lv, W., Li, M., & Duffy, V.G. (2019). Product placement in mass media: a review and bibliometric analysis. *Journal of Advertising*, 48(2), 215-231.
- Hamil, J. (1997). The Internet and international marketing. *International Marketing Review*, 14(5), 300-323.
- Hoffman, D.L. (2000). The revolution will not be televised: Introduction to the special issue on marketing science and the Internet. *Marketing Science*, 19(1), 1-3.
- Kakalejčik, L., Bucko, J., Resende, P.A., & Ferencova, M. (2018). Multichannel Marketing Attribution Using Markov Chains. *Journal of Applied Management and Investments*, 7(1), 49-60.

- Kannan, P.K. (2017). Digital marketing: A framework, review and research agenda. *International Journal of Research in Marketing*, 34(1), 22-45.
- Kannan, P.K., Reinartz, W., & Verhoef, P.C. (2016). The path to purchase and attribution modeling: Introduction to special section.
- Kelly, J., Vaver, J., & Koehler, J. (2018). A causal framework for digital attribution.
- Kireyev, P., Pauwels, K., & Gupta, S. (2016). Do display ads influence search? Attribution and dynamics in online advertising. *International Journal of Research in Marketing*, 33(3), 475-490.
- Kotler, P. (1999). *Marketing management: The millennium edition* (Vol. 199). Upper Saddle River, NJ: Prentice Hall.
- Kotu, V., & Deshpande, B. (2014). *Predictive analytics and data mining: concepts and practice with rapidminer*. Morgan Kaufmann.
- Laverty, K.J. (1996). Economic "short-termism": The debate, the unresolved issues, and the implications for management practice and research. *Academy of Management Review*, 21(3), 825-860.
- Li, H., & 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.
- Moffett, T., Pilecki, M., & McAdams, R. (2014). The forrester wave: Cross-channel attribution providers, Q4 2014. November. Available at <https://www.forrester.com/report/The+Forrester+Wave+CrossChannel+Attribution+Providers+Q4+2014/-/E-RES115221>.
- 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.
- Nottorf, F. (2014). Modeling the clickstream across multiple online advertising channels using a binary logit with Bayesian mixture of normals. *Electronic Commerce Research and Applications*, 13(1), 45-55.
- Papadimitriou, P., Garcia-Molina, H., Krishnamurthy, P., Lewis, R.A., & Reiley, D.H. (2011, August). Display advertising impact: Search lift and social influence. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (1019-1027).
- Qiu, X., Zhang, L., Ren, Y., Suganthan, P.N., & Amaratunga, G. (2014). Ensemble deep learning for regression and time series forecasting. In *2014 IEEE symposium on computational intelligence in ensemble learning (CIEL)* (1-6). IEEE.
- Rentola, O. (2014). Analyses of online advertising performance using attribution modeling. 7.
- Rosnow, R. L., & Rosenthal, R. (1989). Definition and interpretation of interaction effects. *Psychological Bulletin*, 105(1), 143-146.
- Rust, R.T., Lemon, K.N., & Zeithaml, V.A. (2004). Return on marketing: Using customer equity to focus marketing strategy. *Journal of Marketing*, 68(1), 109-127.
- 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.
- Shapley, L.S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2(28), 307-317.
- Wei, A.P., Chen, M.L., & Peng, C.L.E. (2011). The advertising spillover effect: Implications for mutual fund families. *Journal of Management*, 28(4), 361-377.
- Xu, L., Duan, J.A., & Whinston, A. (2014). Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science*, 60(6), 1392-1412.
- 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.
- Zhao, K., Mahboobi, S.H., & Bagheri, S.R. (2019). Revenue-based attribution modeling for online advertising. *International Journal of Market Research*, 61(2), 195-209.