Making Sense of Attribution in Online Advertising

By Antoine Desir , INSEAD

While online advertising has grown rapidly, methods to justify marketing spend on digital platforms have yet to catch up.

Since 1994, when the first clickable banner advertisement (ad) by AT&T was placed on HotWired.com, online advertising has spread like wildfire. Every individual who owns a smart phone or a computer is exposed to a daily deluge of advertising messages. With about **5 billion internet users** and 4.65 billion active social media users as of April 2022, online advertising is big business. In 2021, digital marketing **grew by 35 percent to US$189 billion**.

While online advertising has developed rapidly, one key challenge remains: the problem of attribution. How can an advertiser attribute value to marketing activities across different channels and media? Attribution involves assessing the contribution of individual advertiser actions (ad actions) – such as display ads, paid search and direct electronic mailers – to eventual purchase.

The attribution question drives many marketing decisions. And yet, in spite of huge sums invested by advertisers in online ads, methods to justify these “investments” do not have any theoretical justifications. In our study on **attribution in online advertising**, my co-authors* and I propose a novel attribution metric with desirable properties, and which overcomes the shortcomings of existing heuristics.

What is the question to your answer?

In the online advertising space, marketing executives and advertising agencies make complex decisions on advertising spend, such as allocating budgets across advertising channels and media, as well as optimising tactical decisions for each channel. The contribution, or value of each channel, is an important input to media mix optimisation. It helps build an understanding of the customer journey and helps a company justify its marketing spend.

Commonly used rule-based methods attribute the value generated to the different ad actions on the user’s path to a purchase based on predetermined rules. For instance, last-touch rule attributes all the value generated to the last ad action whereas uniform rule attributes the same value to each ad action. Custom weights attribution is sometimes applied to tailor weights to a series of ad actions.

While these heuristics are easy to implement and understand, they are hardly fair or systematic. Last-touch attribution is akin to the common winner-takes-it-all narrative in soccer, where the player who scores the goal gets the glory. But what about the player who made the final pass, the team effort and interplay of actions that made that goal possible? Can we give credit where it is due, instead of relying on simplistic metrics that do not fully capture the value of each link in the chain?

For instance, it is possible that a user who is **inclined to buy a product** performs a Google search and clicks on a paid search, leading to a purchase. However, even if the paid search click did not occur, the user, who is in a state of “desire”, might have bought the product by finding the product’s link through organic search, in what is known as the counterfactual scenario. Currently, existing methods do not account for causality or counterfactual effects. Given that a purchase may be driven by external factors or a customer's pre-condition, attribution methods should take into account the baseline, or the expected outcome if the user had not been exposed to the ad.

Fundamentally, with current attribution methods, the question that is core to justifying advertising spend remains unanswered: To what extent does a specific ad influence the customer's purchase decision?

A novel attribution metric

Attribution is not unlike assigning credit to individual players in a cooperative game, considering that the value generated by online advertising is the outcome of the cooperative effect of actions taken across channels and media platforms.

Borrowing concepts from cooperative game theory, we propose a novel attribution metric, which we call a counterfactual adjusted Shapley value (CASV). Our method inherits the desirable properties of the traditional Shapley value while overcoming its shortcomings in the online advertising context.

In particular, our method captures causality and accounts for the difference in value when the customer has seen the ad and a hypothetical baseline where the customer has not been exposed to the ad. Unlike rule-based method, value will not be assigned to an ad that has no effect on the customer's purchase decision. Our method is also fair: two ad actions that have the same effect receive the same value.

With the availability of user data in the online advertising environment, our new attribution metric is computationally viable even for large-scale data set. The massive volume of economic activities taking place in the digital realm has enabled data collection at an unprecedented scale, allowing companies to tap into user data to better understand customer behaviour and improve service quality.

In our study, we showcased the applicability of our method and validated the robustness of the model using real-world data with several million user paths and a few hundred thousand purchases (conversions).

Starting with the customer journey

Where it comes to converting customers, ad actions can have disparate impacts on conversion, depending on where the customer is in the conversion funnel commonly known in marketing literature. Through the lens of the customer journey and with insights from user data, we can better understand customer behaviour based on their “state” – from being unaware of the product to being aware, interested and converted.

Accurately defining the state of the customer is a crucial part of this method. Borrowing concepts from the conversion funnel, we estimate the value generated by each ad action based on the assumption that the customer’s browsing behaviour is Markovian in nature, i.e., we assume that the customer’s history can be succinctly summarised in the current state.

To the advertiser, the state observed serves as the basis for appropriate ad actions and can improve the chances of reaching the desired customer demographics at the right time.

From the right question to the appropriate metric

Having ascertained the “right” question advertisers should be asking, advertisers now need to better understand the metrics we choose to drive the business.

In our approach to attribution, we propose to begin by defining the required properties of the attribution method. Based on this approach, our proposed metric is the only one that satisfies the four desirable properties: efficiency, linearity in value attributed, symmetry across channels with the same behaviour, and attributing zero value to channels that do not contribute. Even if a different set of properties were more appropriate in a different context, the techniques in our study could be applied to derive an appropriate method.

Since metrics define the “rules” of the game in the complex online advertising landscape, defining metrics is an important decision. However, it is not sufficiently questioned. What principles are they based on and what are the implications on others? At the end of the day, developments in attribution are a work in progress and asking the right questions is a good way to start.

Raghav Singal, Tuck School of Business; Omar Besbes, Columbia Business School; Vineet Goyal and Garud Iyengar, Columbia University.

Find article at

https://knowledge.insead.edu/operations/making-sense-attribution-online-advertising

About the author(s)

Antoine Desir is an Assistant Professor of Technology and Operations Management at INSEAD. His research focuses on developing efficient data-driven algorithms for revenue management and improving the design of online marketplaces.

About the research

Shapley Meets Uniform: An Axiomatic Framework for Attribution in Online Advertising is published in *Management Science*.