Multi-Touch Attribution

A Guide to Methods, Math and Meaning
Marketers today use multiple marketing channels that generate impression-level data that can be linked to unique individual, or household, customers. While most of these channels are digital, some are not, such as direct mail, call-center records or catalog mailings. A more appropriate and widely-accepted term is “addressable”.

Done right, multi-touch attribution is more than just a scheme to give credit to the addressable channels that precede a conversion (such as a purchase, request for information, or sign-up). It’s a way to assess performance, measure return on investment and, ultimately, guide marketing budget allocations to the most effective channels.

The ultimate goal of marketing is to change consumer behavior. Thus, the true measure of marketing effectiveness is not how likely a customer is to buy, but whether and by how much marketing increases a customer’s likelihood to buy. For that reason, any method used for multi-touch attribution must be based on estimates of incremental effects. In addition, the incremental effects must express causality, not just correlation.

In this world of targeting, customers are typically exposed to marketing because they are already considered more likely to make a purchase. But that doesn’t necessarily mean the marketing is precipitating a change in behavior. Similarly, executing a search or clicking on a display ad is a strong indicator of purchase intent, but that doesn’t mean the customer is influenced by the display ad or the landing page.

The distinction between a propensity to buy between different customers and the incremental persuasion power of marketing to an individual is at the crux of a viable multi-touch attribution model, a fact that is often missed in the midst of fancy buzzwords. (Particular “buzzwords” to be skeptical of include Shapley value, game theory, and post-hoc control group.)
What Makes Multi-Touch Attribution Viable?

Many attribution methods are based on pre-determined weights that are used to proportionately assign attribution to the marketing treatments preceding an outcome. Clearly, simple weight-based allocations like first- or last-click, equal attribution or time-dependent weights do not get to true incrementality.

Any viable multi-touch attribution methodology (one that is not inherently biased leading to wrong conclusions) must account for the following four concepts:

1. **INCREMENTALITY**: You should first understand that marketing is not responsible for the entire purchase. Each customer has an innate propensity to purchase without any exposure to marketing. It is the change in propensity that matters and must be measured.

2. **HETEROGENEITY BETWEEN CUSTOMERS**: Each customer has a different base propensity to purchase. Customers are exposed (or not exposed) to different marketing tactics based on their unique propensity to purchase, which accounts for differences in conversion rates. These differences must be separated from the actual persuasive effect of marketing.

3. **EXTERNAL EFFECTS**: Base conversion probabilities change over time, often rapidly, if the customer environment changes. These non-addressable factors such as the economy, price, competitive behavior, weather and, of course, non-addressable advertising are often correlated with online marketing and, if not taken into account, will bias results.

4. **DATA VISIBILITY BIAS**: Often, a customer action or impression is required for the customer to enter the universe of people that leave a trace in the data files. The makeup of the observed population changes, depending on what kind of marketing is executed. While a well targeted display campaign might add many high-propensity customers to the dataset, an indiscriminate campaign or cheap affiliate traffic could yield a high number of non-converting leads. As a result, the average conversion probability of the targeted population will vary.

Each customer has an innate propensity to purchase without any exposure to marketing. It is the change in propensity that matters and must be measured.
Neustar’s Approach to Multi-Touch Attribution

To account for all of the above factors, Neustar uses an automated model-based approach to multi-touch attribution. This relies on a flexible, proprietary model creation platform that generates faster, more relevant and more accurate insights and recommendations.

Neustar’s model-based attribution process includes the following:

- Historical transaction data, customer attribute data, and offline, market-level data is joined into customer histories.
- Model features that describe the customer, the environment and the received marketing exposures are extracted from the data.
- The model is calibrated by statistical estimation of the response parameters.
- For each customer, the model is used to score the customer’s history, to assess the incremental effect of each marketing exposure on conversion probability.
- Conversion metrics (e.g. revenue) are attributed to marketing exposures based on their incremental impact on conversion probability.

(Neustar only uses securely de-identified customer ID; no PII is used within Neustar analytics.) Neustar uses logit models to express the relationship of influencing variables to conversion probability:

\[
\ln \left( \frac{p}{1-p} \right) = \sum_i a_i C_i + \sum_j \beta_j M_j (X_1, ..., X_n) + \sum_k Y_k A_k
\]

The model has three distinct sets of terms capturing different aspects of the customer decision:

- The first sum captures differences in individual customers due to their distinct attributes. These attributes can be demographic or summaries of past interaction behavior.
- The second sum captures the influences of variables that are collected at the market level. These variables are aggregated into market-level conduit variables \(M_i\) that depend on market-level drivers \(X_m\). These conduit variables are derived from market-level models.
- The third sum captures multi-touch attribution variables derived from the customer’s interaction history. These variables capture recency (or the effect of time since last marketing impression) and frequency (the effects of multiple impressions of the same type on an individual customer) for different types of marketing treatments and their interactions.

By estimating and scoring these models, we obtain a complete database of impression-level contributions that can be summarized and compared along any dimension of the dataset. These contributions are a true reflection of the lift caused by marketing and when combined with impression-level cost, can provide marketing’s real ROI.
Why Use Customer Response Models?

Assessing advertising efficiency, at its core, is an exercise in marketing science and quantitative customer psychology more than an exercise in computer science. Central to Neustar’s approach is a customer response model that predicts the probability of each individual customer to convert as a function of the customer’s demographic, psychographic and behavioral attributes, as well as the customer’s history of received (or self-initiated) marketing exposure.

Such customer response models have a long tradition in marketing analytics and have distinct advantages over other approaches that include:

- They are based on well-understood theories of customer behavior. These theories—including concepts like recency, frequency or saturation of advertising, and interaction effects—can be expressed in intuitive ways and have been used in marketing science for decades.

- They account for non-addressable influencers of customer behavior by incorporating population-level data, along with using customer-level data.

- They provide insights at the customer level that can be used for multi-touch attribution and other marketing applications. Examples of such models include Targeting applications, Lifetime Value models and Next-Best-Action applications.

- Calibration of customer response models from data is a well-understood statistical discipline. The models themselves can be assessed for statistical as well as business validity.
Why Capturing Non-Addressable Influencers is Critical

Variables that are often outside the addressable dataset affect customer propensity to convert. Ignoring such non-addressable variables results in a less accurate multi-touch attribution model, often inflating the effect that digital marketing has on sales.

Thus, it’s essential that your multi-touch attribution solution includes these non-addressable variables, as well as a logical way to incorporate this information into the model.

Depending on data availability and level of detail required, we build anywhere from simple to highly sophisticated aggregate market response models that feed into our multi-touch attribution models. These models contain all relevant drivers of conversion, including seasonality, price, competitive activity, economy, weather, off-line advertising, and on-line advertising as appropriate. We incorporate this data in aggregate time series form and use it to assess the incremental effects of offline variables on conversions over time.

With aggregate-level models, the problem becomes how to use them in attribution. If all the variables affect individual customer propensity to convert, they ought to be part of the individual customer response model.

A naïve application of an incrementality percentage derived from market-level models indiscriminately to all customer histories—what we call the “haircut method” — will bias attribution substantially. In these methodologies, highly effective digital marketing treatments will be penalized while ineffective ones will be favored. As a result, differentiation will be dampened and reallocation opportunities might be squandered.

Our approach uses aggregate-level models to establish the relative impact of different variables in each time period and to summarize these impacts in a manageable number of “conduit variables” that channel the offline effects into the customer-level dataset for estimation. As a result, online effects are estimated in the presence of offline effects but the relative effects of offline variables are controlled. This produces a consistent, unbiased and rich customer-level model that has all offline variables present for analysis.

Our approach uses aggregate-level models to establish the relative impact of different variables in each time period and to summarize these impacts in a manageable number of “conduit variables” that channel the offline effects into the customer-level dataset for estimation.
A Complete Multi-Touch Attribution Solution

In short, here are five things that make our multi-touch attribution solution superior to others:

1. A marketing-science rooted, model-based approach that ensures accurate and unbiased results.

2. Systematic and objective methods that incorporate offline effects and non-addressable marketing treatments into attribution models.

3. An automated yet flexible model creation platform for quick turn-around that ensures model accuracy, relevancy and recency.

4. True event-level attribution—the ability to analyze results across any attribute or dimension.

5. True and comprehensive customer response models that can be used beyond multi-touch attribution.
Every day, the world generates roughly 2.5 quadrillion bits of data. Neustar (NYSE: NSR) isolates certain elements and analyzes, simplifies and edits them to make precise and valuable decisions that drive results. As one of the few companies capable of knowing with certainty who is on the other end of every interaction, we’re trusted by the world’s great brands to make critical decisions some 20 billion times a day. We help marketers send timely and relevant messages to the right people. Because we can authoritatively tell a client exactly who is calling or connecting with them, we make critical real-time responses possible. And the same comprehensive information that enables our clients to direct and manage orders also stops attackers. We know when someone isn’t who they claim to be, which helps stop fraud and denial of service before they’re a problem. Because we’re also an experienced manager of some of the world’s most complex databases, we help clients control their online identity, registering and protecting their domain name, and routing traffic to the correct network address. By linking the most essential information with the people who depend on it, we provide more than 11,000 clients worldwide with decisions — not just data.