April 21, 2015

ABSOLUTE ATTRIBUTION USING FREQUENCY EXTRAPOLATION

Robert Marsa

Follow this and additional works at: http://www.tdcommons.org/dpubs_series

Recommended Citation
http://www.tdcommons.org/dpubs_series/64

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.
ABSOLUTE ATTRIBUTION USING FREQUENCY EXTRAPOLATION

Previous solutions to the attribution problem only assign relative credit among known factors. High level mixed media models will sometimes use R² to determine the approximate credit for unknowns. A way of quantifying the effects of unknown factors in influencing behavior in order to assign accurate credit to the known factors is introduced here.

A method of absolute attribution using frequency extrapolation may include examining a lift of being exposed to additional advertising to quantify the effect of that additional advertising. For example, if people who see N+1 events are 10% more likely to convert than people who see N events, then the N+1st event deserves credit proportional to this lift. The method may further include directly measuring this lift for the cases of N=1 and higher. Additionally, the method may include using extrapolation to determine a baseline for zero events. In this way, the incremental effect of seeing one event only can be determined.

One advantage of absolute attribution over relative attribution is that it allows advertisers to know the absolute credit deserved by each advertising event instead of just the relative split of credit between events. The approach presented here thus provides a more accurate and granular measurement of the effects of unknowns than using R² at the model level.

FIG. 1 depicts a diagrammatic representation of an example network environment for fractional cross-channel attribution. In the example of FIG. 1, a user 102 may “convert,” or perform a desired action, after clicking a link 104 (e.g., a banner ad on a publisher web site 114, a search engine ad 110, or an ad on another channel 112), via a user device 106 at a particular Internet Protocol (IP) address and being directed via network 122 to the advertiser’s web page 116. Conversion 118 can be a purchase transaction, but could also include such actions as registering with a web site, signing up for product information, and the like. An attribution platform 120 allows the advertiser 116 to make informed decisions about payment for advertisements and future ad campaigns.

Data from the click 101 and ultimate conversion 118 may be collected in a variety of ways. One or more computers in the network 122 may collect click data. A click data collecting computer may be a server machine residing in a publisher 114’s or other party’s computing environment or network. The click data collecting computer may collect click streams associated
with visitors to one or more web sites. The collected information may be stored in one or more log files. The information associated with clicks may include visitor Internet Protocol (IP) address information, date and time information, publisher information, referrer information, user-agent information, searched keywords, cookies, and so on.

The attribution platform 120 employs “ad tags” for monitoring impression data and “page tags” for monitoring click data. Ad tags can be 1 x 1 pixels embedded in page code at the publisher site and can be used to determine where the ad is on a page (above or below a “fold,” i.e., visible with or without scrolling) and whether and how long a user sees it. Page tags can be embedded in a similar manner on the landing page, and can identify whether a user has arrived and where the user comes from. Such ad tags or page tags can be transmitted to the attribution platform 120 responsive to a user viewing or clicking on an ad and viewing or clicking on an associated web page.

FIG. 2 depicts a diagrammatic representation of example system architecture 200 including one or more clients 202 and attribution platform 220. A user may browse a publisher site 204 which maintains one or more ad tags 205. Ad tag data can be sent to a tag server 210, responsive to a user viewing or clicking an ad, which stores in a database 216, impression data sorted by customer. Such data may include, e.g., where, when, and how long a user viewed the ad. An ad server 212 may be used to maintain the ad on the publisher’s web site 204. The user 202 may click an ad to arrive at a landing page 208. Embedded on the landing page 208 includes a page tag 207, which identifies user accesses to the landing page 208 and may be sent to a database 214 accessible by the attribution platform 220. An advertiser 206 records a conversion 218, if any, and likewise provides the information to the attribution platform 220.

Attribution platform 220 may reside in a computing environment including one or more server machines. Each server machine may include a central processing unit (CPU), read-only memory (ROM), random access memory (RAM), hard drive (HD) or non-volatile memory, and input/output (I/O) device(s). An I/O device may be a keyboard, monitor, printer, electronic pointing device (e.g., mouse, trackball, etc.), or the like. The hardware configuration of this server machine can be representative of other devices and computers alike at a server site (represented by platform 220) as well as at a client site.
Platform 220 may include a system and a computer program product implementing a method for absolute attribution using frequency extrapolation in a network environment. Platform 220 may be owned and operated independent of the clients that it services. For example, company A operating platform 220 may provide attribution services to company B operating a client (not shown). Companies A and B may communicate over a network or over a secure channel in a public network such as the Internet. Example clients may include advertisers, publishers, and ad networks.

The system may run on a web server. The computer program product may include one or more computer readable storage media storing computer instructions translatable by multiple processors to process attribution data. The input data may be from a log file, a memory, a streaming source, or ad and page tags. Within this document, the term “attribution data” refers to any and all data associated with online advertising events such as clicking on an ad, viewing an ad (an impression), entering a search query, conversion, and so on, and may include click history data, click intelligence data, post-click data, visitor profile data, impression data, etc.

Software running on a server computer in platform 220 may receive a client file containing attribution data from an attribution data collecting computer associated with a client. For example, a client may represent an online retailer and may collect click stream data from visitors to a Web site own and/or operated by the online retailer. The attribution data thus collected can provide a detailed look at how each visitor got to the Web site, what pages were viewed by the visitor, what products and/or services the visitor clicked on, the date and time of each visit and click, and so on. As discussed above, the specific attribution data that can be collected from each click stream may include a variety of entities such as the Internet Protocol (IP) address associated with a visitor (which can be a human or a bot), timestamps indicating the date and time at which each request is made or click is generated, target URL or page and network address of a server associated therewith, user-agent (which shows what browser the visitor was using), query strings (which may include keywords searched by the visitor), and cookie data. For example, if the visitor found the Web site through a search engine, the corresponding click stream may contain the referrer page of the search engine and the search words entered by the visitor. Attribution data can be created using a corporate information infrastructure that supports a Web-based enterprise computing environment. A skilled artisan can
appreciate what typical attribution click streams may contain and how they are generated and stored.

Attribution data may include an impression/click record for every ad impression/click received from a given client of the system. An example impression/click record may include one or more of the following attributes:

- Impression/click timestamp;
- visitor cookie (if available, may be set up as a domain cookie for persistent visitor identification);
- visitor IP address;
- visitor browser user-agent;
- impression/click source (may be a publisher ID or a referrer domain);
- click destination (landing page Web address or bid keywords for advertisers); and
- conversion data (whether the visitor executed a desired conversion).

To determine the effect of unknown factors that may not be in the attribution data, a baseline may be used for people not exposed to the known factors. This baseline can be determined by extrapolation based on ad exposure frequency.

FIG. 3 shows a plot of conversion rate versus exposure frequency. As illustrated in FIG. 3, more exposure leads to higher conversion rates, although the marginal effect decreases. So, the second event seen may provide a larger effect than the third event, etc. Accordingly, a curve is fitted to this function and extrapolated to 0. This point provides a baseline conversion rate for users not exposed to the advertising events in question.

The curve fitting can be done in several ways, including:

- Fit a line to the linear region early in the curve
- Fit a non-linear function to the entire curve
Fit separate lines to separate regions of the curve to establish multiple baselines

One issue is that groups of users that see very different numbers of ads may be different in behavior. Generally speaking, for instance, users that are more active on the internet will see more online ads. Users that watch more TV will see more TV ads. These groups of active users may be different behaviorally from users who spend less time online or watch less TV. Therefore, these groups of users may have different baselines, meaning they respond differently to ads. One way of addressing this is to use approach 3 to determine separate baselines for groups of users with different ad exposure rates.

This frequency extrapolation may be performed for each independent advertising channel for which data is available. The extrapolation is also performed for each channel at multiple levels of granularity. For instance, if a system has paid search and display channels, a baseline is established for each channel. As an example, consider a user who sees two events, \( E_a \) \( E_b \). A system as disclosed herein may first determine the conversion rate for all users who have seen these two events. The system may then compute the lift over the baseline. The credit due to advertising events for this path can be determined as follows, in some implementations:

\[
\text{Lift for this path is: } \frac{\text{CR}(a,b)}{\text{Baseline}} - 1
\]

\[
\text{Absolute credit for this path is } 1 - \frac{\text{Baseline}}{\text{CR}(a,b)}
\]

So, if the baseline conversion rate is 0.2 and the conversion rate for seeing both event a as well as event b is 0.25, then the absolute credit due to the path is \( 1 - 0.2/0.25 = 0.2 \). The remaining credit (0.8) for the baseline is due to unknown factors, that is, things other than events a and b. In some implementations, the relative credit for events a and b can be assigned using a relative attribution algorithm.

At a different level of the hierarchy, say site, a system as disclosed herein may perform the following.

A user sees site A 2 times and site B 3 times. The system may compute the conversion rate (“CR”) for all users who have seen this path. The system may then determine a baseline for site A by extrapolating the conversion rate vs. frequency curve for users who have seen site A
Likewise, the system may determine a baseline for site B by extrapolating the conversion rate vs. frequency curve for users who have seen site B (“B(B)”).

Using these two baselines, the system may compute two different estimates of the absolution credit due to the path as follows:

\[ C(A) = 1 - \frac{B(A)}{CR} \]

\[ C(B) = 1 - \frac{B(B)}{CR} \]

The system may combine these two estimates using a weighted average where the weights are given as:

\[ W(A) = \frac{\text{number of people who saw A 2 times}}{\text{total}} \]

\[ W(B) = \frac{\text{number of people who saw B 3 times}}{\text{total}} \]

\[ \text{Total} = \text{number of people who saw A 2 times} + \text{number of people who saw B 3 times} \]

\[ W(A) \times C(A) + W(B) \times C(B) \] gives the absolute credit for the path.

As above, the relative weights for events within the path can be assigned using a relative attribution algorithm, in some implementations.

FIG. 4 depicts a diagrammatic representation of a data processing system. As shown in FIG. 4, data processing system 400 may include one or more central processing units (CPU) or processors 401 coupled to one or more user input/output (I/O) devices 402 and memory devices 403. Examples of I/O devices 402 may include, but are not limited to, keyboards, displays, monitors, touch screens, printers, electronic pointing devices such as mice, trackballs, styluses, touch pads, or the like. Examples of memory devices 403 may include, but are not limited to, hard drives (HDs), magnetic disk drives, optical disk drives, magnetic cassettes, tape drives, flash memory cards, random access memories (RAMs), read-only memories (ROMs), smart cards, etc. Data processing system 400 can be coupled to display 406, information device 407 and various peripheral devices (not shown), such as printers, plotters, speakers, etc. through I/O devices 402. Data processing system 400 may also be coupled to external computers or other...
devices through network interface 404, wireless transceiver 405, or other means that is coupled to a network such as a local area network (LAN), wide area network (WAN), or the Internet.

Abstract

Systems and methods pertaining to attribution are provided. In some implementations, a method enables assigning accurate credit to known factors by quantifying the effects of unknown factors in influencing behavior. The method may include examining a lift of being exposed to additional advertising to quantify the effect of the additional advertising. The method may further include measuring the lift directly for the non-zero events and using extrapolation to determine a baseline for the zero event.
**FIG. 3**

![Conversion Rate by Impression Frequency](image)

**FIG. 4**

![Diagram](image)