A Time to Event Framework For Multi-touch Attribution

Dinah Shender, Ali Nasiri Amini, Xinlong Bao, Mert Dikmen, Amy Richardson Fricke, Jing Wang

June 8, 2022

Paper link [available on research.google.com]

Acknowledgements

This is the work of a large (15+) team at Google, as well as partner teams. I’m just the representative that’s here today :)

Agenda

- Why multi-touch attribution (MTA)?
- Describe MTA requirements
- Describe modeling framework that fulfills them
- Discuss attribution

What is Multi-touch Attribution?

Conversion

Search "Running shoes" and click on Nike ad

Watching YT and see a Nike ad

Search "Nike Free Run" and click on Nike ad

Go to Nike.com and purchase shoes

	Search "Running shoes" and click on Nike ad	Watching YT and see a Nike ad	Search "Nike Free Run" and click on Nike ad	Go to Nike.com and purchase shoes
Last Click			100%	
DDA	40%	20%	40%	

Why Do Advertisers Want MTA?

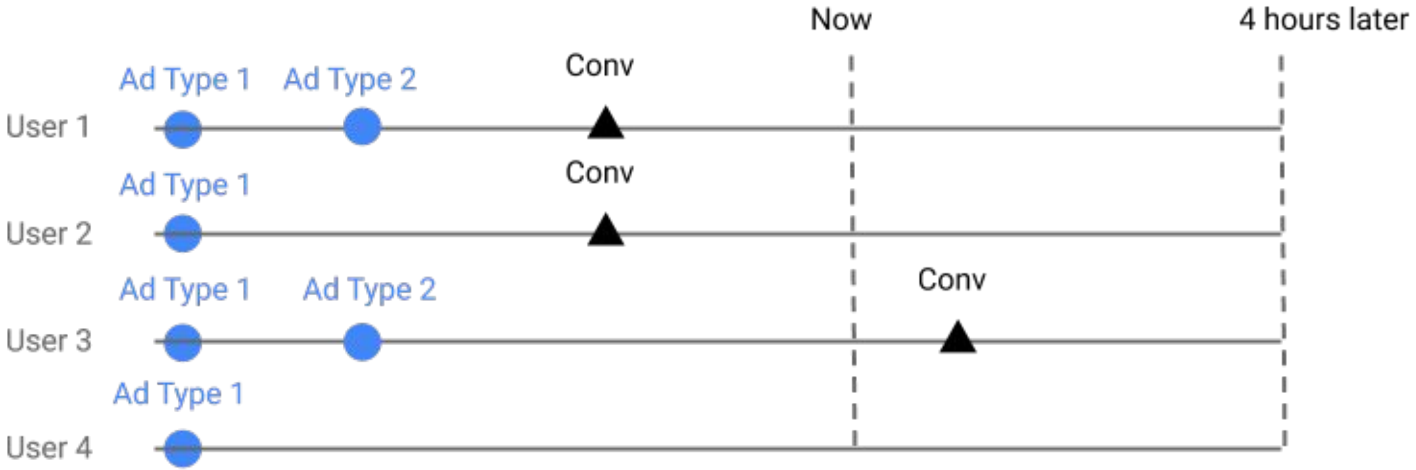
Advertiser Goal: want to ensure spend is proportional to ads' contributions to conversions (which represents ad ROI). MTA tells them where to direct spend.

Which ads are shown is determined by ad auction. Best way to align spend and attribution is to adjust auction bid according to attribution.

Often want to do so in real time, in order to capture changes in market, business fluctuations etc. So attribution system must also be real-time.

Requirement #1: Handle Incomplete Data

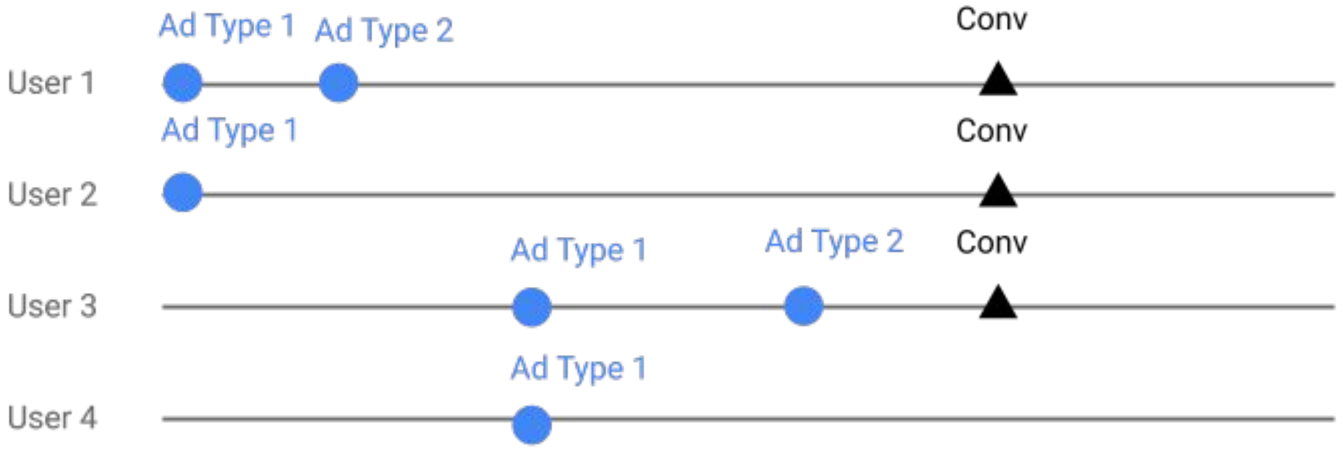
Real-time data is necessarily incomplete/censored



At t=Now, Ad Type 2 seems not to contribute to conversions. At t=Now+4h, it seems to contribute 50%

Requirement #2: Ad contribution depends on time delay

Effect of an ad changes (decreases) over time



Looking at order, Ad Type 2 increases conversions by 50%. But accounting for timestamps, more likely that it accounts for all conversions soon after, but doesn't help in long-term.

Split MTA into 2 sub-problems

- **Modeling User Conversion Behavior**
  Many interesting applications, including attribution

- **Attribution**
  How to distribute credit amongst ads, given a model of user conversion behavior?

Framework is Flexible With Respect to Type of Data

Framework for attributing credit based on estimating how ads change user conversion rates over time.

With **observational data**: correlation between ads and conversions. “What ads are most associated with conversions?”

With **experimental data**: causal relationship between ads and conversions. “What ads cause conversions?”

Overall framework for modeling user conversion behavior is the same, so won’t focus on this distinction

Proposed model

Handling censored data suggests survival-type model, but with multiple occurances ("deaths"). Use an inhomogenous Poisson process

$$Y_i(s) - Y_i(r) \sim Poisson\left(\int_r^s \lambda_i(t)dt\right)$$

$Y_i(t)$ is the number of conversions for user i at time t.

$\lambda_i(t)$ is the intensity ("conversions per unit time"). It depends on the user's path and the time from ad events to t.

Modeling Intensity Function

Model for intensity function is highly flexible. Next few slides are just examples

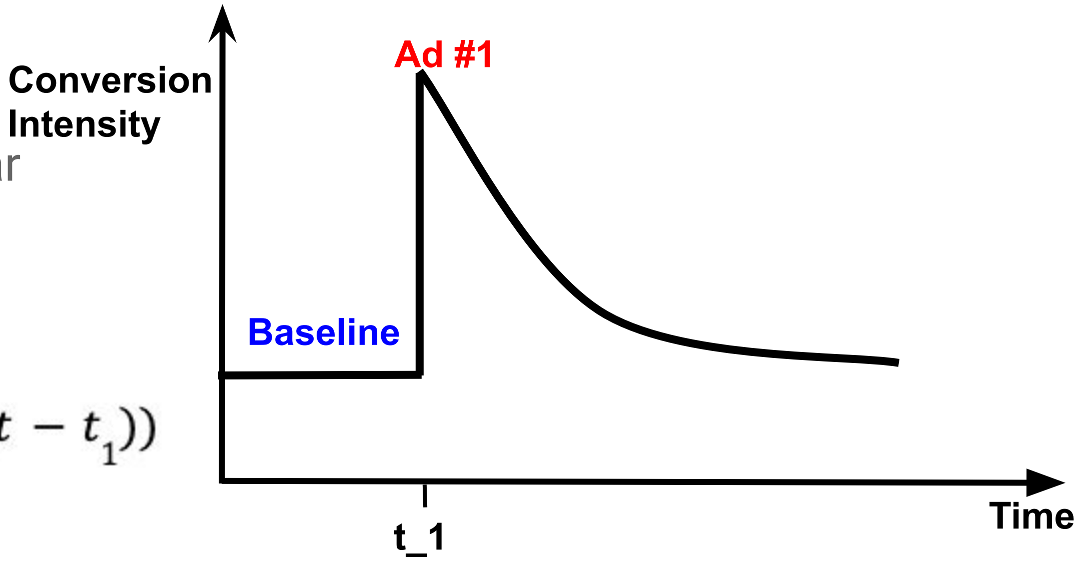
Intensity Function: Single Ad at t_1

$$log(\lambda(t)) = \alpha_0 + f(t - t_1)$$

Many ways to parameterize f(t):

- Approximate as piecewise linear
- Splines
- Exponential "basis", e.g.:

$$log(\lambda(t)) = \alpha_0 + \sum_{l=1}^{L} \beta_l \, exp(-\theta_l(t - t_1))$$



Intensity Function: Single Ad With Features

Can add ad features $\{x_{1k}\}_{k=1,\ldots,K}$:

$$log(\lambda(t)) = \alpha_0 + f(t - t_1) + \sum_{k=1}^{K} g_k(t - t_1, x_{1k})$$

Parameterize g_k similarly to f, possibly with lower dimension/fewer degrees of freedom.

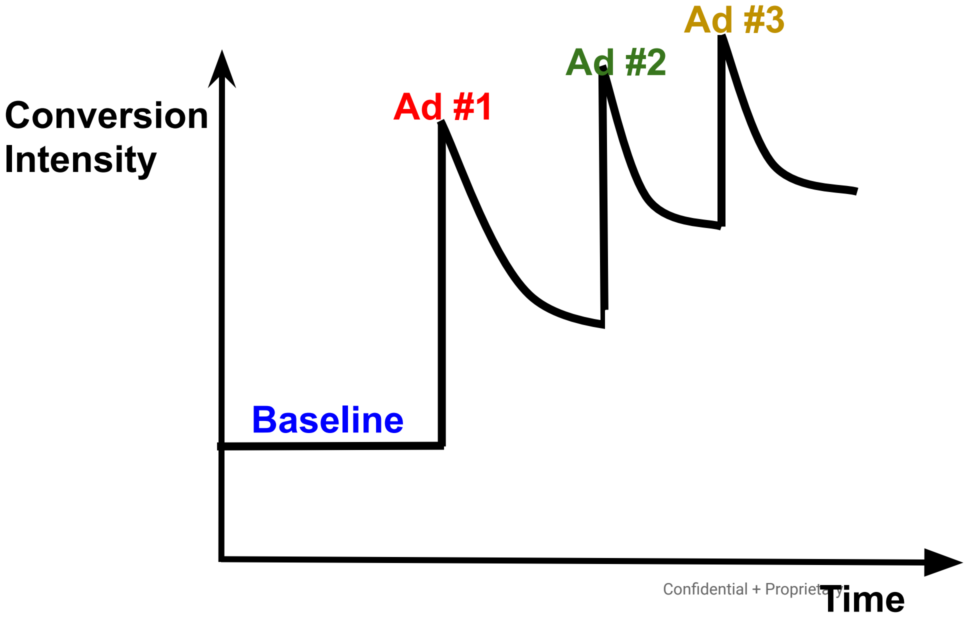
Intensity Function: Multiple Ads

Similar to previous:

$$log(\lambda(t)) = \alpha_0 + \sum_j f(t - t_j) + \sum_{j,k} g_k(t - t_j, x_{jk})$$

Can handle between-ad interactions by allowing feature to depend on current and previous ads.

See paper for more details.



Estimation Options

Many options. If $\lambda_i(t)$ piecewise constant or willing to discretize, can treat as Poisson regression. Each constant intensity interval generates a training example with

- Response = # conversions in that interval
- Offset = length of interval

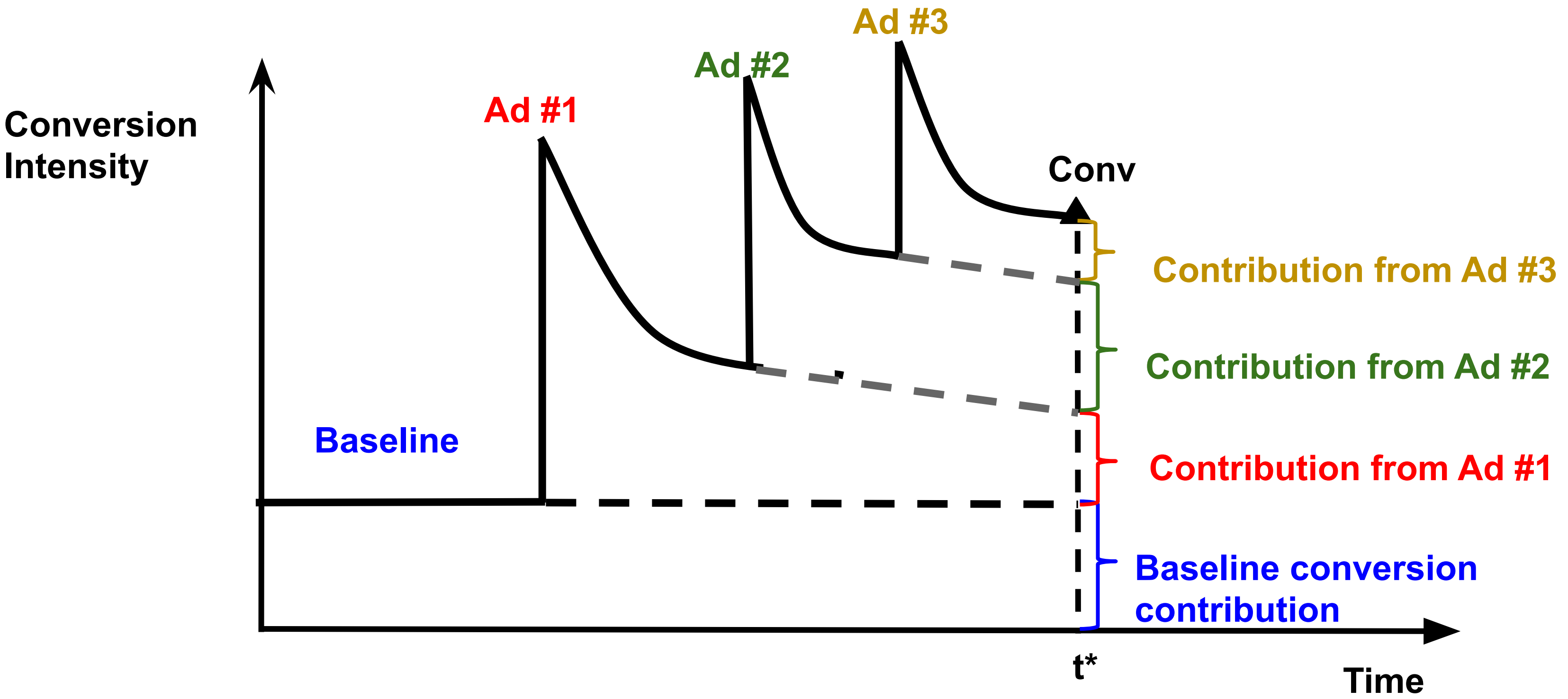
Apply your favorite type of regularization. We use empirical Bayesian approach (details).

Attribution

Many ways to do attribution!

This is just one way, but it has some attractive properties

Backwards Elimination Attribution in 1 picture



Backwards Elimination Notation

Formally,

$\mathcal{A}(j)$ = first j ads in user path

$\lambda(t^*, \mathcal{A}(j))$ = intensity at conversion time if only see first j ads

$$RawCredit(j) = \lambda(t^*, \mathcal{A}(j)) - \lambda(t^*, \mathcal{A}(j-1))$$

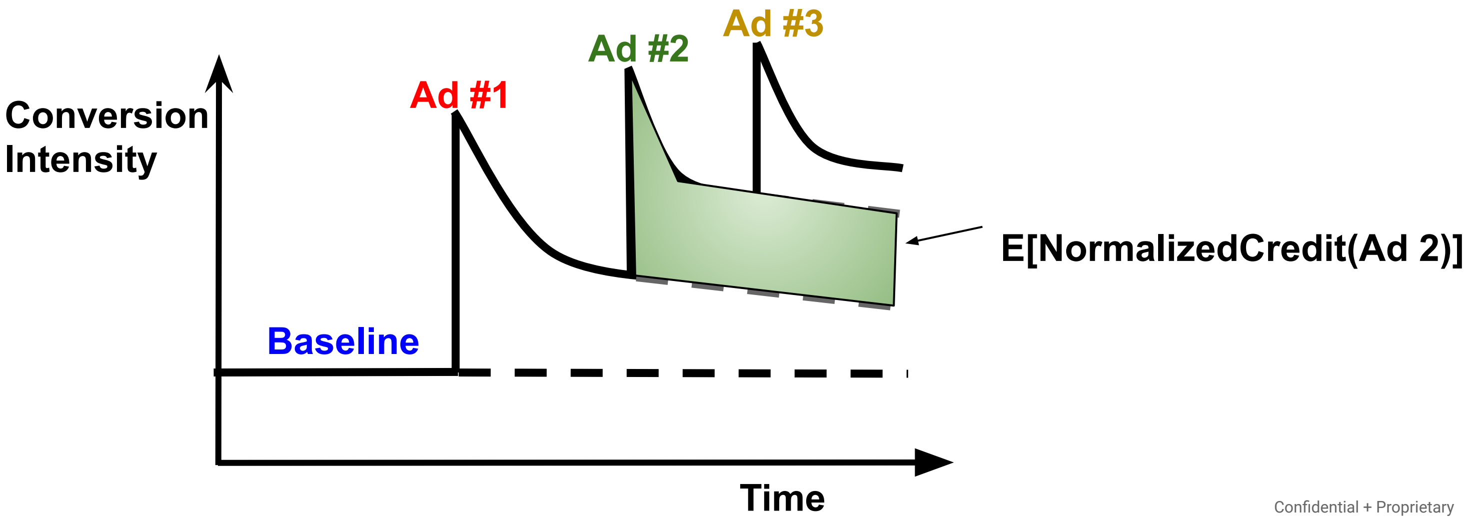
Assume n ads by time t*, then

$$NormalizedCredit(j) = \frac{\lambda(t^*, \mathcal{A}(j)) - \lambda(t^*, \mathcal{A}(j-1))}{\lambda(t^*, \mathcal{A}(n))}$$

Backwards Elimination Motivation

$$E[NormalizedCredit(j)] = \int (\lambda(t, \mathcal{A}(j)) - \lambda(t, \mathcal{A}(j-1)))\, \mathrm{d}t$$

RHS = Expected number of additional conversions from ad j



Backwards Elimination and Interactions

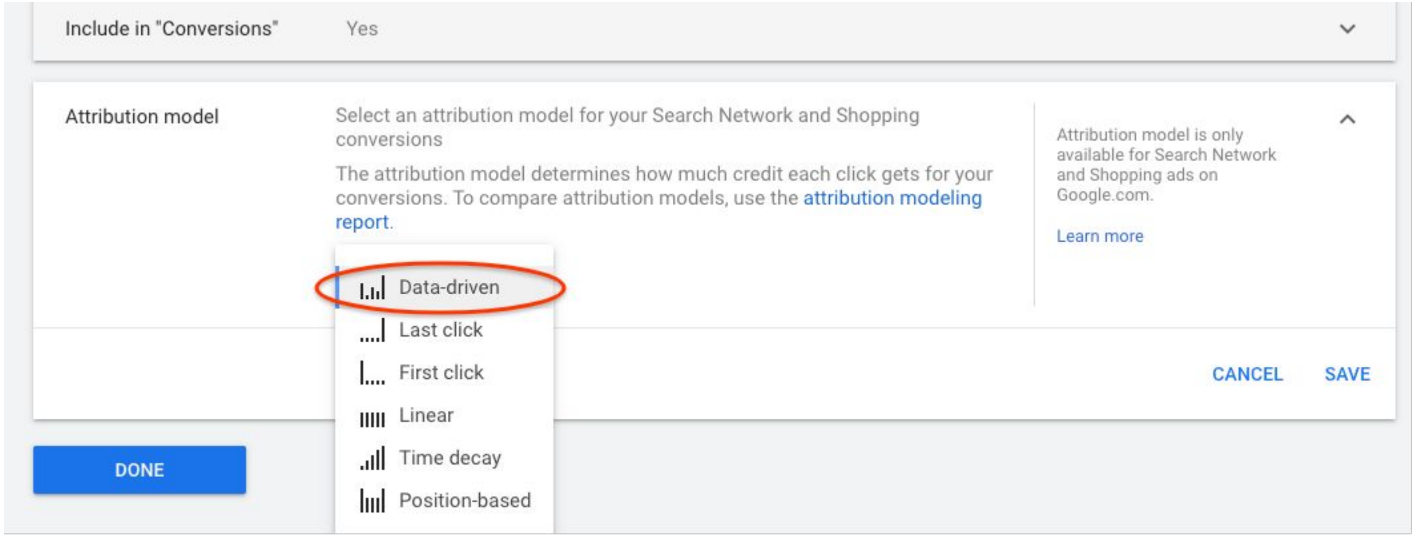
Consider the special case:

$$RawCredit(Ad2) = \lambda(t^*, \mathcal{A}(2)) - \lambda(t^*, \mathcal{A}(1))$$

If there's an interaction between ad 1 and ad 2 that changes intensity, resulting change in intensity (and ad credit) all goes to second ad.

More generally, effects of interactions affect credit for the last ad.

Methodology is Actively Used at Google!



[External DDA help page]()

Appendix

Experimental Data

Many ways to set-up experiments

Suppose that you withhold all ads for some advertiser for some users ("unexposed"), but log when they would have seen ad.

Remaining users see ads as normal ("unexposed")

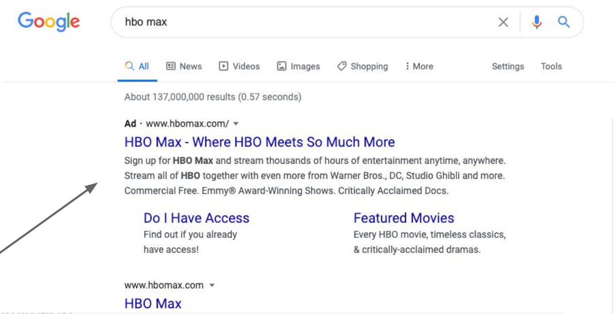
Can model this in intensity function by crossing f and g with is_exposed

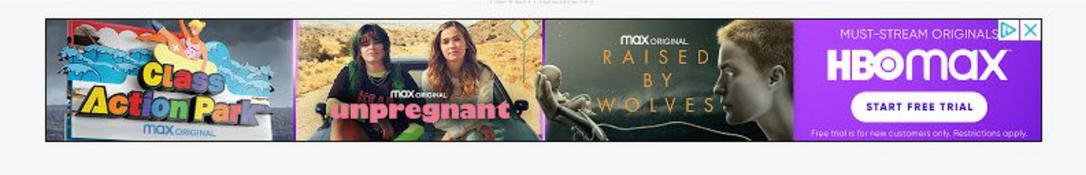
Estimation Options

Alternatively, could use SGD on the log-likelihood:

$$\sum_{i=1}^{N}\left[-\int_{0}^{\tau}\lambda_i(t)\,dt + \sum_{j=1}^{C_i} log(\lambda_i(T_{ij}))\right]$$

What is Multi-touch Attribution?





- See or click on multiple ads
- Then go buy product (“convert”)
- How to attribute credit for conversion to ads (“multi-touch path”)?

Note, more generally, conversions can include email sign-ups, purchases, or anything advertiser wants to track.