Attribution Modeling Increases Efficiency of Bidding in Display Advertising

Julien Meynet, Eustache Diemert, Pierre Galland, Damien Lefortier*

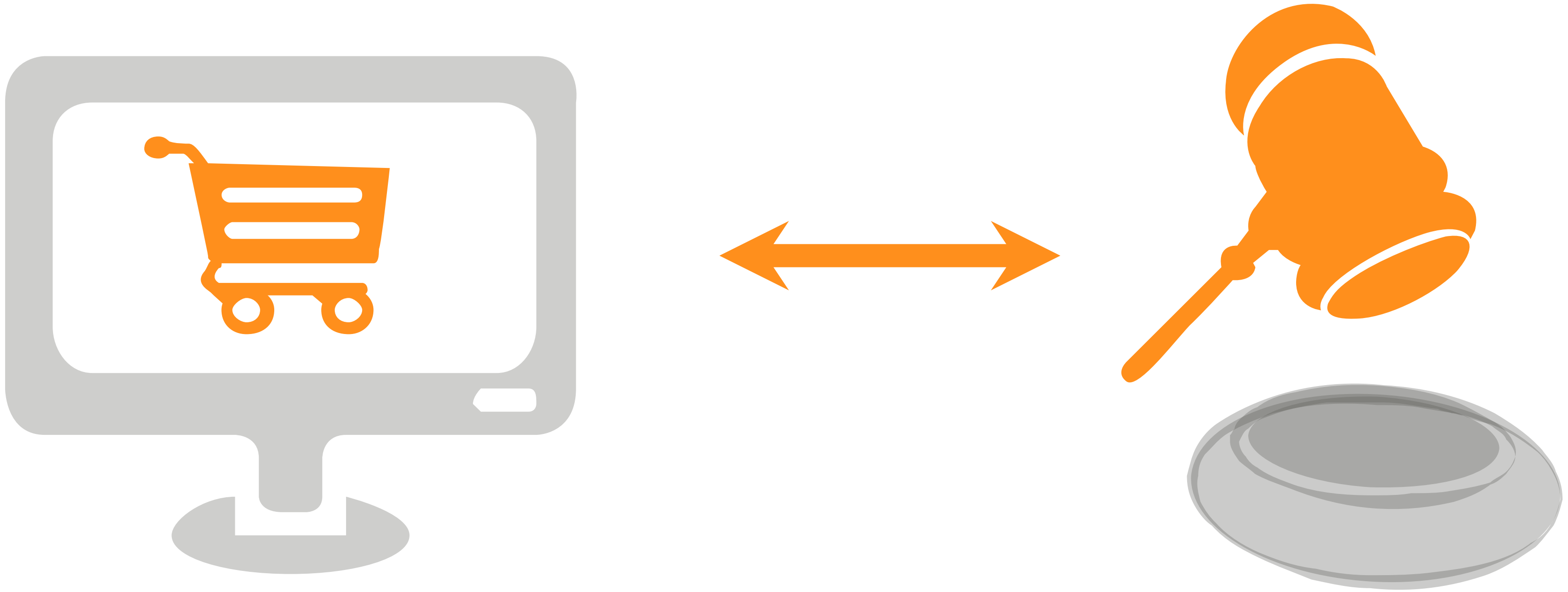
* now @Facebook

2017 AdKDD & TargetAd, Halifax, Canada

Outline

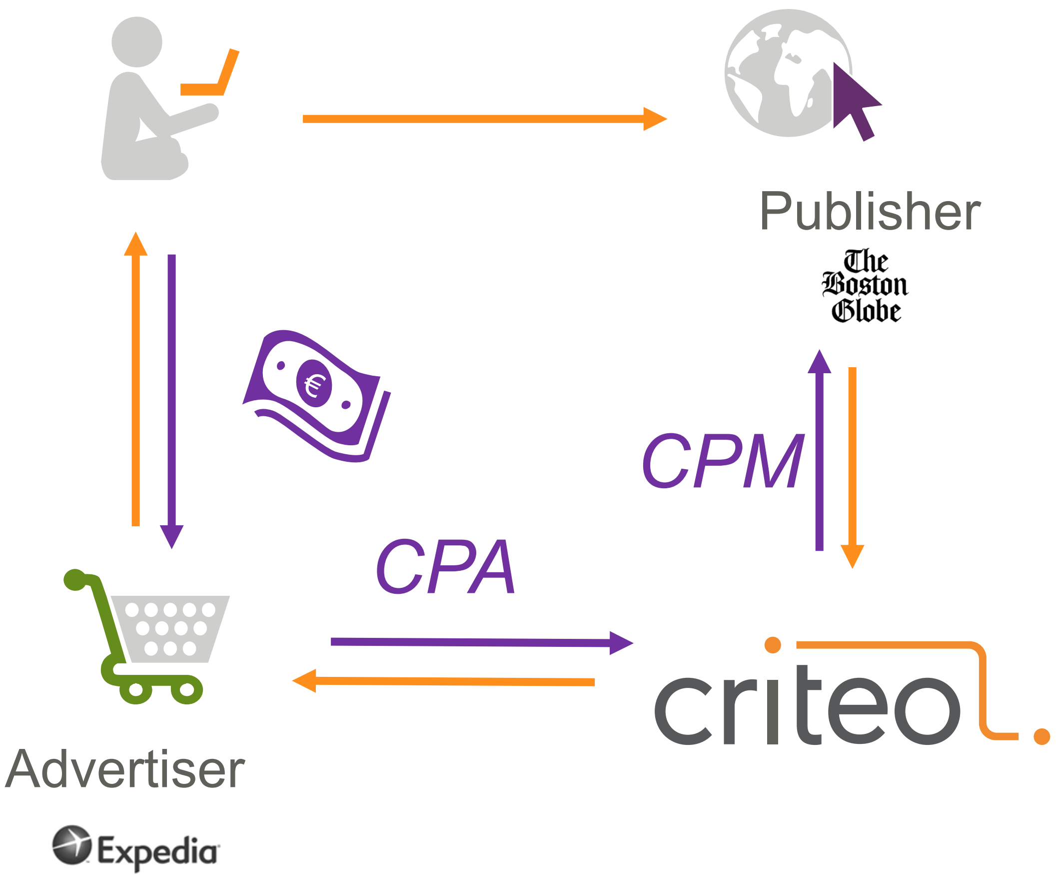
- The problem: bidding in display advertising
- Model:
  - Attribution model
  - Attribution aware bidder
- Impact on offline evaluation metrics
- Experience & results

Real-time bidding for display advertising



criteo.

Performance Display Advertising



1. User visits a publisher webpage
2. Criteo recieves real-time auction
3. Displays ad on behalf of advertiser
4. User converts on advertiser website (click / sale / lead)

Baseline Bidding Policy

- Under 2nd price auction hypothesis, dominant strategy is to bid expected value

$$bid^* = CPA \times pSale$$

« Value of a conversion »

« Probability of post-click attributed conversion »

« Post-click attributed conversions »?

- *Last-click* is the de facto attribution model…

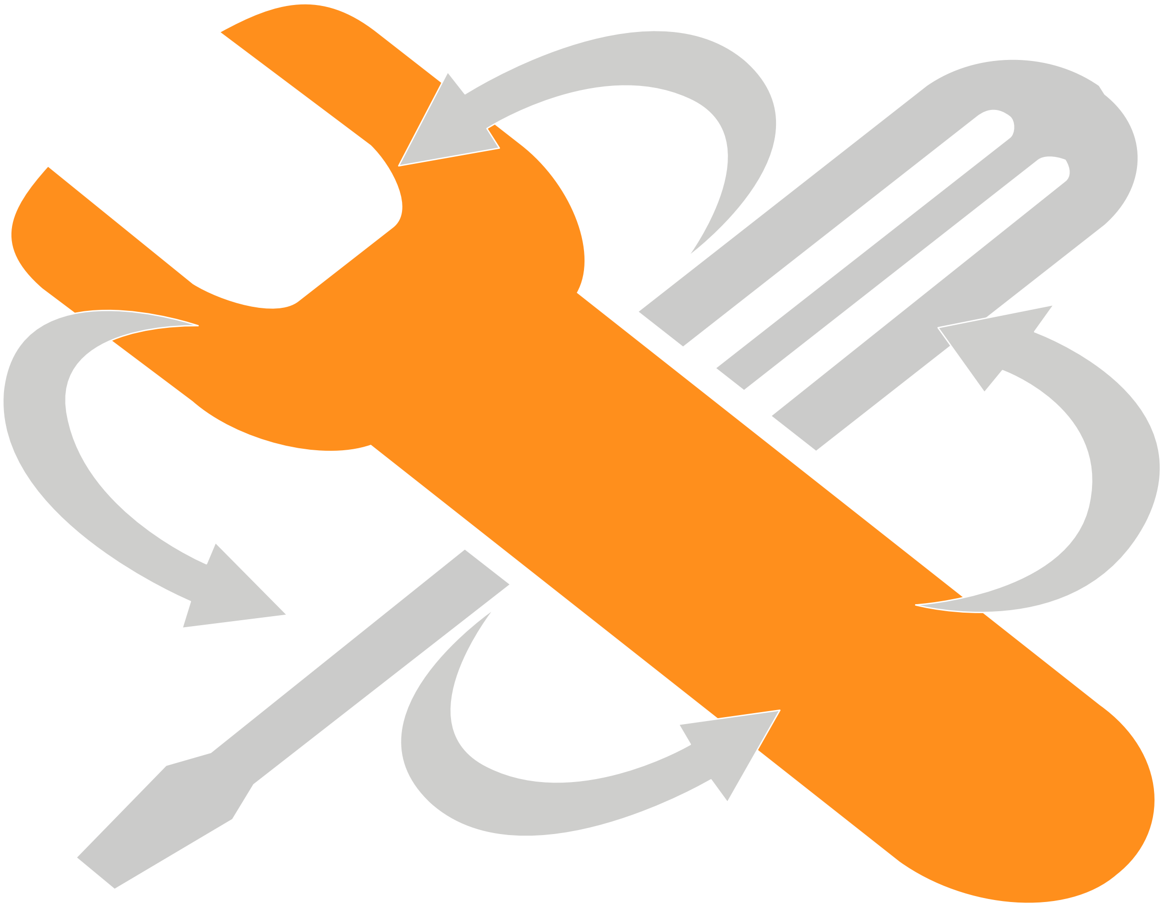
  … but advertisers are moving towards "better" attribution models:

  - Rule-based, uniform, linear, etc..
  - Data driven: regression, shapley value, etc..

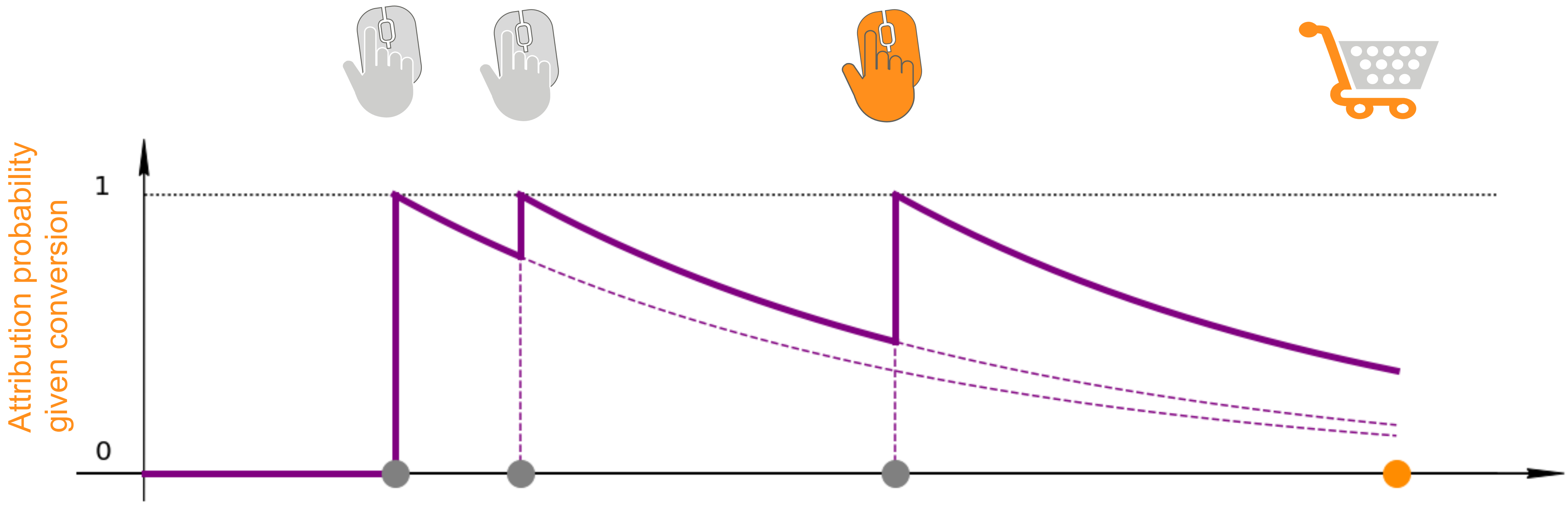
- But what is the impact from a bidder's perspective?
  - What is the optimal bidding strategy right after a click?

criteo

Attribution-aware bidder



Attribution Probability Through Time Matters



Attribution Model

- How can we model probability of getting the attribution given there will be a conversion?

$$Pr(A = 1|S = 1, X = x, \Delta = \delta) = e^{-\lambda(x)\delta},$$
$$\lambda(x) \geq 0$$

- S: Post click conversion
- A: Attributed conversion
- X: Contextual features
- Δ: Delay click/conversion

Conversion Modeling

$$bid^* = CPA \times pSale$$

« Probability of post-click attributed conversion »

- Baseline solution:
  - $0/1$ prediction problem $\Longrightarrow$ Logistic Regression
  - Large scale / latency constraint $\Longrightarrow$ Hashing trick

But what are positives / negatives?

From Attribution Model to an Attribution Aware Bidder

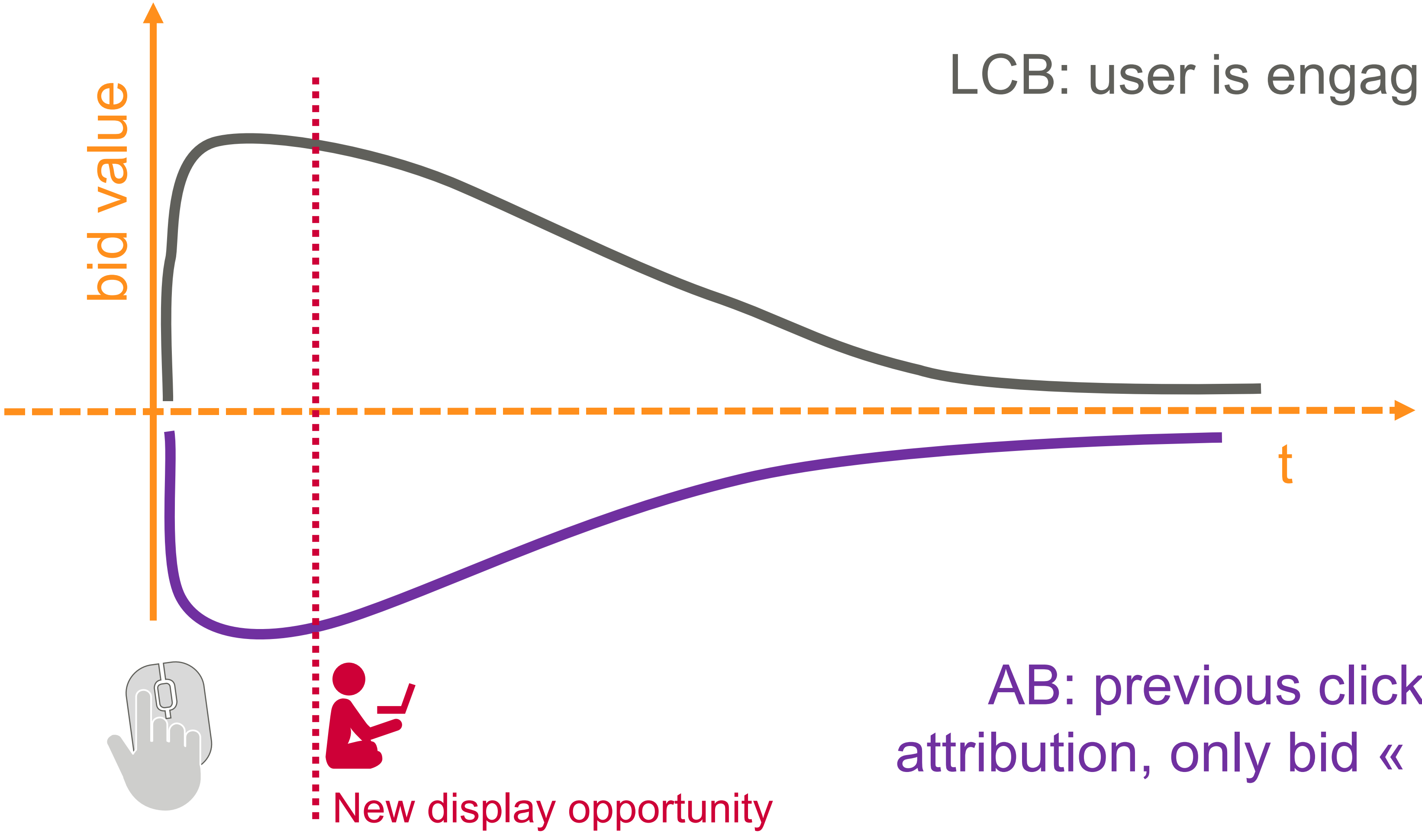
A_{LC}	0	0	1
A_{UN}	1/3	1/3	1/3
A_{FC}	1	0	0
A_{ALL}	1	1	1
A_{AM}	0.6	0.1	0.3

Cast the problem as an internal attribution problem

criteo

Attribution Aware Bidder: An Intuitive View

LCB: user is engaged, go for last-click



AB: previous click gives us the attribution, only bid « marginal value »

Attribution Aware Bidder

- Baseline Last-click Bidder (LCB)

$$bid = CPA \times Pr(A_{LC} = 1 | X = x)$$

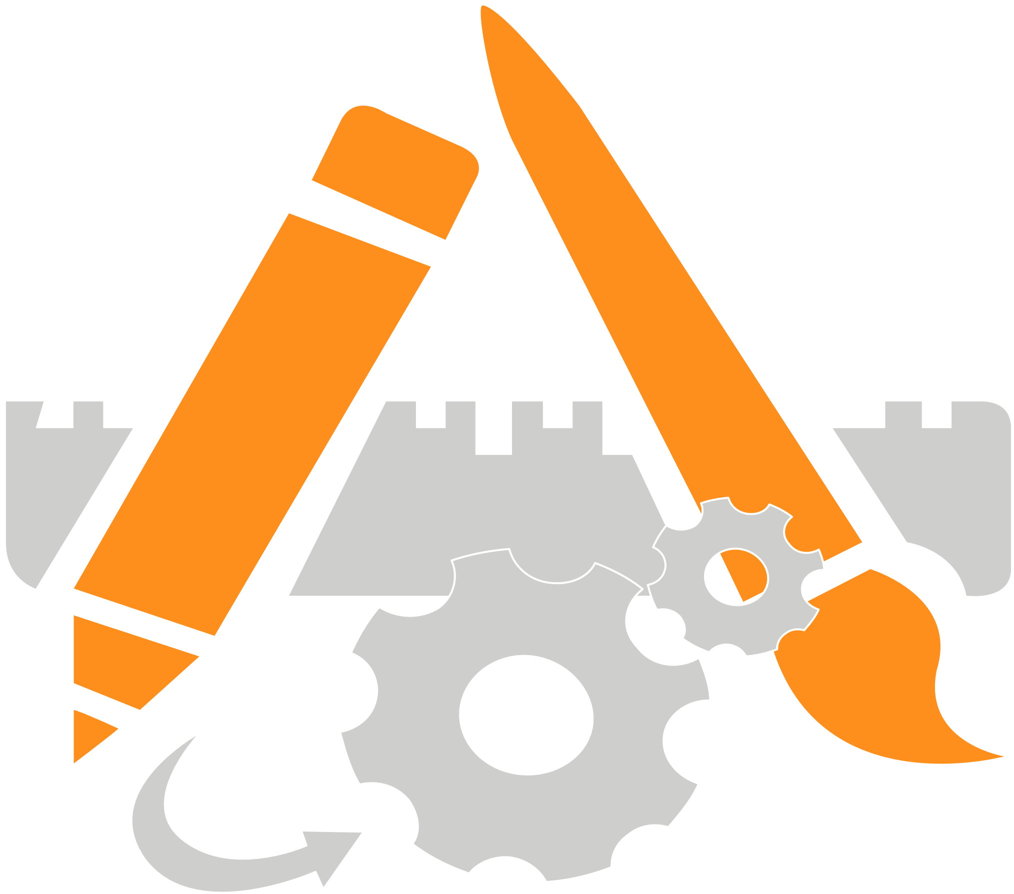
- Attribution-aware Bidder (AB):

$$bid = CPA \times Pr(A_{ALL} = 1 | X = x)\left(1 - e^{-\lambda(x)\delta_c}\right),$$

$$\delta_c : time\ elapsed\ since\ lastc\ click$$

Bid proportionally to the marginal contribution of the display

Impact on the offline evaluation metrics



criteo

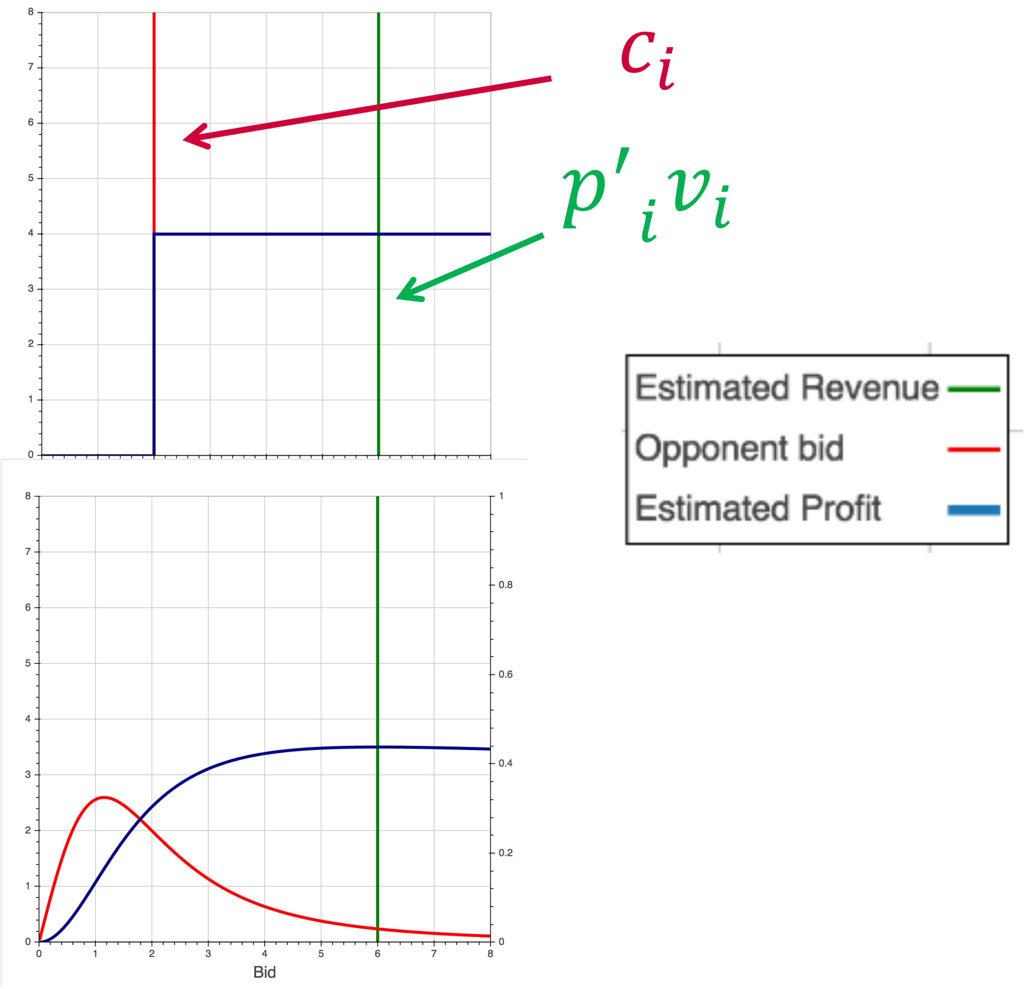
Offline Evaluation of Bidders

- *Utility* metric on logged feedbacks:

$$U(p') = \sum_{i \in D} (a_i v_i - c_i)\mathbb{I}(p'_i v_i > c_i)$$

- *Expected Utility*: add uncertainty on the cost distribution:

$$c \sim \Gamma(\alpha = \beta c_i + 1, \beta)$$



Attribution Aware Expected Utility*

- Inject attribution function in the Utility:

$$AU(p', a) = \sum_{i \in D} (a(x_i) v_i - c_i) \mathbb{I}(p'_i v_i > c_i)$$

Internal attribution function:

- can be last-click, first click, etc..
- can be the proposed attribution model

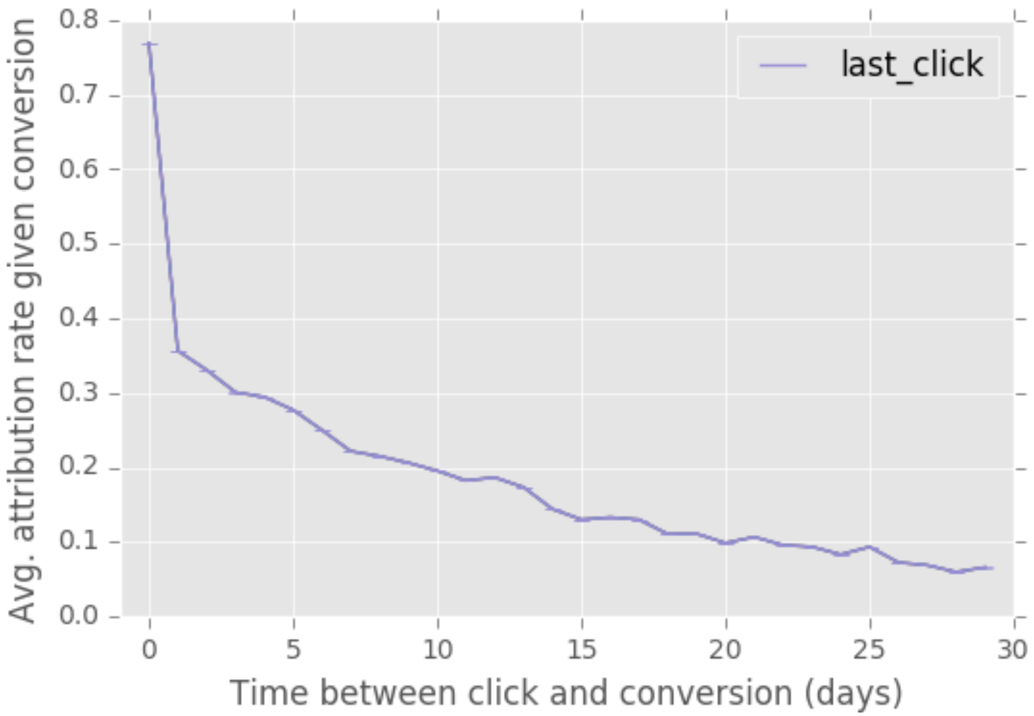
* Evaluation of the proposed metric would require a proper offline / online correlation analysis

Experiments & Results

PLACEHOLDER IMAGE

Attribution Rates vs Time

Decay of attribution rate after a click



> 40% of conversions have more than one click in the preceding 30d

Offline Evaluation - Dataset

Log sampled from 30 days of Criteo traffic

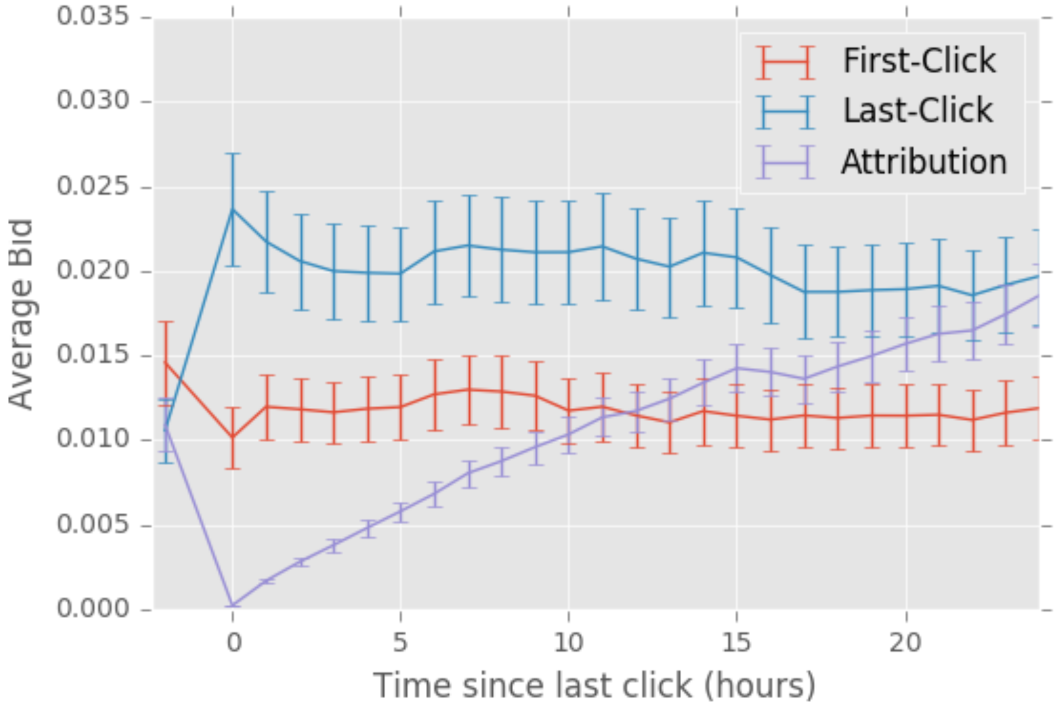
- Anonymized
- Each line is an impression with:
  - *Timestamp*
  - *Price paid*
  - *Contextual features (user, advertiser, publisher)*
  - *Click*, click position*, click number**
  - *Conversion*, conversion value**
  - *Attribution label (conversion was attributed to Criteo)*
- 16M displays, 5M clicks, 800k conversions

Will be available at http://research.criteo.com/ soon

Offline Evaluation – Impact on Bid Profiles

Post-click bid profiles for 3 bidders:

- Last-Click Bidder (LCB)
- First-Click Bidder (FCB)
- Attribution Bidder (AB)



All models are learn using regularized logistic regression + hashing trick

Offline Evaluation – Bidders Comparision

Results for 3 bidders on the Attribution Aware Expected Utility

	LCB	FCB	AB
Win Rate	0.94	0.90	0.89
$U_A{}^*, \beta = 1000$	2852 ± 43	2888 ± 43	$\mathbf{3396 \pm 53}$

- We limit user over exposure after a click
- We get closer to lift-based bidding
- We can reinvest budget on more profitable campaigns / more incremental ads

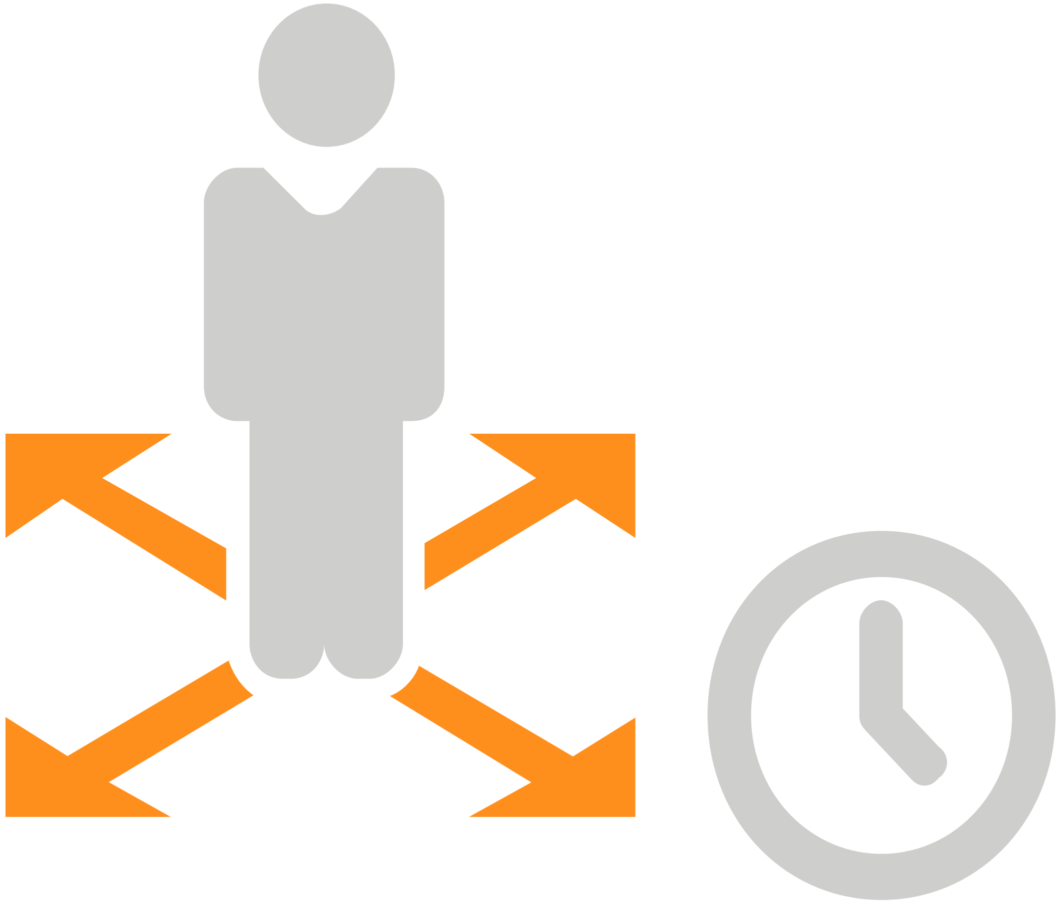
Online result

We tested online a ***simple*** modification of baseline through A/B testing:

$$bid_{test} = bid_{ref} \times A\left(1 - Be^{-\lambda\delta_c}\right)$$

ΔOEC (long term)	Revenue (short term)	Advertiser ROI	User ad exposure
$+5.5\%$ world wide	negative	positive	lower

Future Research Directions



Work in progress & Next steps

- Better attribution modeling
  - Exponential decay is naive: build a better model (e.g travel partners have different attribution schemes)
- Model both conversion lift and attribution lift
  - Delayed feedback in both cases
- Derive a robust (counterfactual) offline metric

Questions?

PLACEHOLDER IMAGE