



HAL
open science

Causal models for Real Time Bidding with repeated user interactions

Martin Bompaire, Alexandre Gilotte, Benjamin Heymann

► **To cite this version:**

Martin Bompaire, Alexandre Gilotte, Benjamin Heymann. Causal models for Real Time Bidding with repeated user interactions. 2020. hal-02971865v1

HAL Id: hal-02971865

<https://hal.science/hal-02971865v1>

Preprint submitted on 19 Oct 2020 (v1), last revised 25 May 2021 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Causal models for Real Time Bidding with repeated user interactions

MARTIN BOMPAIRE*, Criteo AI Lab, France

ALEXANDRE GILOTTE*, Criteo AI Lab, France

BENJAMIN HEYMANN*, Criteo AI Lab, France

A large portion of online display advertising inventory is sold through real time auctions. The bidding algorithms need to estimate precisely the value of each display. Many bidding models estimate this value as the probability that a sale is attributed to this display, but this approach does not capture that a user may be shown a sequence of several displays. By mixing tools from causal reasoning and reinforcement learning to model this sequence of auctions, we derive a simple rule to improve this estimate. We test the change online in a production environment and the results validate the approach. We believe this methodology could be adapted to tackle the notoriously difficult problem of building an incremental bidder.

Additional Key Words and Phrases: real-time bidding, causality, incrementality, attribution

ACM Reference Format:

Martin Bompaire, Alexandre Gilotte, and Benjamin Heymann. 2020. Causal models for Real Time Bidding with repeated user interactions. In *Proceedings of* . ACM, New York, NY, USA, 18 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

1 INTRODUCTION

Display advertising is a form of online advertising in which a marketer pays a website owner (the publisher) for the right to show banners to its visitors (users) in the hope of triggering some future sales. It is a large industry: last year, it was estimated that display advertising generated more than 57 billions USD in the United States [Fisher, 2019].

When the internet user reaches the publisher page, he triggers a complex mechanism called Real Time Bidding (RTB). Real Time Bidding involves several intermediaries (DSP, Ad exchange, SSP, ...) in a chain of calls that finishes its run in the advertisers' server in the form of a bid request. In each advertiser's server lives an algorithmic bidding agent (the *bidder*) that implements the advertiser's strategy. When the bidder receives the bid request, it only has a few milliseconds to answer with a bid. Then, the highest bidder receives the right to show a banner to the user, in exchange for a payment that depends on the mechanism implementation. The bidder can make a marketing campaign succeed or fail, because it decides for whom, when, where and at which price to buy a banner. Estimation of economic value is an important component in most bidding architectures. The purpose of the present paper is to discuss how such estimation should be made.

The standard to evaluate the quality of a bidder is to use an *attribution* algorithm that assigns the credits of online sales to marketing events (such as the display of a banner on a news site). For instance, the most commonly used attribution algorithm is the fairly simple *last click rule*: when a sale occurs, all the credit goes to the last banner clicked that precedes the sale in the user history. This methodology has an obvious flaw because the last click rule does not

*The three authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2020 Association for Computing Machinery.

Manuscript submitted to ACM

18 measure how many additional sales the campaign is bringing. Such flaw motivated some critics in the literature. A
 19 famous experiment, reported in [Blake et al., 2015], showed that eBay used to attribute many credits to search advertising
 20 because people often do a search to reach eBay instead of typing the URL directly. The first link that would appear in
 21 the search results was usually an eBay ad. The experiment consisted in observing that shutting down search advertising
 22 did not decrease the overall performance, because people would then click on the first search result which would also
 23 redirect to eBay’s website anyway. This state of affair is unsatisfactory for the industry at large, and explains the trend
 24 in designing offline metrics and bidders with causal inference methods to improve the number of *incremental sales* –
 25 sales that would not have occurred without the marketing effort. Fortunately, several tracks of research provide tools to
 26 stay away from the *correlation is causation pitfall*.

27 Our aim is to show that tools from causal reasoning [Pearl, 1995, 2009, Peters et al., 2017] can indeed improve the
 28 performance of a bidder. However, what we propose is slightly different from the above-mentioned literature: we use
 29 causal methods to increase the performance of a bidder *measured with an attribution rule* instead of being measured
 30 as the total number of sales as in the eBay example. The motivation behind focusing on the received attribution is
 31 threefold: (a) the attribution is still the main metric in the industry, which is something that might persist because of
 32 the business models of digital advertising, (b) our approach can ingest causally motivated attribution as input, which is
 33 currently an active track of research [Dalessandro et al., 2012, Singal et al., 2019], (c) from a practical and methodological
 34 perspective, it is easier to work with attributed sales – for which we have methodologies, models and benchmarks –
 35 than with incremental sales, so while the models to address the two questions are similar, it is much easier, to test,
 36 measure, compare results and improve with attributed sales than it is with incremental sales.

37 Our model relies on the common approximation that a non clicked display has no more effect on the user than no
 38 display at all. We also explicitly assume that the display should impact the state of the next requests for this user. This
 39 allows us to produce counterfactual estimates. While our methodology is simple, our experimental results indicate that
 40 it is robust enough to allow for the rough theoretical assumptions and implementation choices we made.

41 In Section 2, we explain why we think the bidding problem is not fully solved in the literature. In Section 3, we
 42 challenge the common assumption that the bidding problem reverts to estimating accurately the probability that a
 43 sale would be attributed to this display. We provide a detailed description of the bidder’s problem and the hypotheses
 44 we need to turn it into an optimization procedure solved with classic statistical tools. In Section 4, we propose a new
 45 way to estimate this causal effect on sales, relying on some additional hypotheses on the causal links between displays
 46 and sales. In Section 5, we propose a new incrementality metric which can be used to run offline evaluations of our
 47 model. Section 6 presents some experimental offline experiments on a public data set from Criteo –a large Demand Side
 48 Platform that allows its clients to externalize the bidding process– and the results we have obtained online on a large
 49 scale experiment.

50 2 RELATED WORK AND CONTRIBUTION

51 Real Time Bidding (RTB) has been the standard for selling ad inventory on the Internet for almost one decade, fueling
 52 an extensive literature [Choi et al., 2020, Wang et al., 2016] on online bidding. In the early stages, the bidders were
 53 typically getting revenue from *clicks*, and much work was thus done on click prediction models [Chapelle et al., 2014,
 54 McMahan et al., 2013].

55 Clicks however are not the ultimate goal of advertising. What matters the most is to increase the number of *sales* on
 56 the advertiser’s website. To bridge the gap between clicks and sales, one may rely on heuristics to determine which
 57 clicked display, if any, would have caused a sale. On one hand, *last click* – the most common attribution rule – states

58 that a sale should be *attributed* to the last click preceding it. On the other hand, several more sophisticated attribution
59 rules rely on advanced Machine Learning models [Dalessandro et al., 2012, Ji et al., 2016, Singal et al., 2019, Zhang et al.,
60 2014]. Bidders are then incentivized on *attributed sales*, so that it is now usual to estimate the value of a display as “Cost
61 Per Action (CPA) \times Probability that the display will receive an attribution”. A game theoretical analysis of the situation
62 is provided in [Berman, 2018].

63 Over the decade, the bid formula shifted from this display value estimate for several reasons. First, the market globally
64 moved from second-price auctions to first-price auctions [Despotakis et al., 2019, Heymann, 2020]. The bidder now
65 needs to estimate the distribution of the price to beat – the highest bid of the competition – to compute its optimal
66 bid [Krishna, 2009]. Second, it is common for the bidder to have some constraints on the advertising campaign, such
67 as a maximum budget per day, or a maximum cost per click [Conitzer et al., 2018, Heymann, 2019]. While in several
68 cases a linear scaling applied to the display valuation may be enough to optimally satisfy a budget constraint, the value
69 of this scaling factor is not known in advance, and several articles propose to update the bidding strategy to better
70 take such constraint into account [Cai et al., 2017, Grislain et al., 2019, Lee et al., 2013, Yang et al., 2019]. In the present
71 work, we do not consider such constraints, but those methods might be applied on top of ours. Finally, bidding with the
72 probability that a sale is attributed to current display may not be optimal when there are several display opportunities
73 on the same user. The authors of [Diemert et al., 2017] present a simple but efficient heuristic for taking into account
74 user sequences with several clicks, explicitly lowering the bids right after a click. The idea that showing several displays
75 in a row does not serve the advertiser well is at the origin of several works on pacing and probabilistic throttling.
76 While an analytical solution of the pacing problem for display advertising is presented in [Fernandez-Tapia, 2019]. Most
77 industrial solutions rely on heuristics [Agarwal et al., 2014, Chen et al., 2011, Lee et al., 2013, Xu et al., 2015].

78 We propose in Section 3 a formulation of the bidder problem on the sequence of bid requests for one user, and
79 retrieve the intuitive result that we should account for the probability that the display caused the attributed sale. Our
80 formulation is closely inspired by Reinforcement Learning [Sutton et al., 1998], defining the sequences of requests
81 for one user as an episode. While Reinforcement learning has already been proposed to improve the bidder in [Cai
82 et al., 2017], previous approaches only encoded the remaining budget into the state, assuming i.i.d. requests. To our
83 knowledge, this is the first work to *explicitly assume that the display should impact the state of the next requests for this*
84 *user*.

85 Closely related to the sequentiality of the interactions with the user, the question of incrementality is an attractive
86 field of research [Lewis and Rao, 2015, Lewis et al., 2011]. Incremental sales are those that needed the display to happen
87 and may be measured with an A/B test by switching the bidder off on a part of the population. As argued in [Xu et al.,
88 2016], it would be in the interest of the advertiser that the bidders value a display with the *lift*, which we may understand
89 as the probability that the display cause the sale. The difficulty here is that we never directly observe incremental sales.
90 We either observe a sale on a user after a sequence of displays, or – in the case of an A/B test of incrementality – after
91 no display at all. This is a typical case of a causal inference problem, where we want to estimate the effect of a treatment
92 (here bidding to buy display ads) on an outcome (in this case the sale).

93 Several works have already proposed bidding methods optimized for incremental sales. For instance, [Diemert et al.,
94 2018] proposed to estimate the causal effect on sales of bidding on a user, by applying randomized bid factor on each
95 user. This allows learning models predicting which users are impacted by display ads. However, if those users are
96 seeing several displays, as it is typical, it does not allow retrieving which of those displays caused the sale. Similar to
97 our work [Rahier et al., 2020] leverage the way the causal effect is mediated to derive an uplift estimate. A method to
98 decrease the noise of incremental sales measurement is presented in [Johnson et al., 2017]

99 Some authors also proposed to use observational data to estimate the causal impact of a display. Observational
 100 studies typically require to assume that there are no unobserved confounders; but because of the large state space
 101 they usually also need additional assumptions to lower the variance. [Xu et al., 2016] propose to sample user states of
 102 random times, and instead of looking if a specific display is won, they rather check if there is at least one won display in
 103 the following time window. The (implicit) hypotheses here seems to be that the display only impact the sale through the
 104 “number of displays in a time window” variable. Similarly [Moriwaki et al., 2020] propose to unbias a model estimating
 105 the probability that a user is led to buy after a request by using inverse propensity weights, but also summarize the
 106 impact of winning a display by the increase of the variable “number of displays”. Such hypothesis seems imperfect, as
 107 variables such as the size of the display or the quality of the publisher are known to have a huge impact at least on clicks
 108 or attributed sales; and it seems therefore likely that they also have great impact on incremental sales. By contrast, our
 109 methodology allows taking into account all the information available on the user and the display opportunity. Finally,
 110 another issue of bidding for incremental sales is that the bidder is not paid for incremental sales.

111 This paper assumes a given attribution mechanism but does not challenge it. To better align bidders and advertisers
 112 objectives, a better attribution, based on causality, seems to be necessary.

113 3 THE BIDDING PROBLEM

114 In this section, we propose a formalization of the bidder’s problem trying to maximize its revenue. In particular, we
 115 carefully consider the fact that a bidder will have a sequence of opportunities to display ads to the same user. Our
 116 framework is general enough to apply to a bidder retributed for either attributed sales or incremental sales. The main
 117 result of this section is Theorem 3.2, which states that a display valuation should be the difference of two terms:

- 118 • Its impact on expectation of attributed sales.
- 119 • And its impact on expected cost paid by the bidder at later ¹ time steps.

120 The bidder should then find the bid which maximizes the expected difference of display value and cost on the current
 121 auction (for example, bidding with the display value when the auction is second price). While intuitively simple, this
 122 result allows us to cleanly separate the display valuation from all the issues related to the type of auction. The correct
 123 mathematical definition of those terms, however requires a few technical hypotheses, which are often implicitly used by
 124 most authors, and that we state explicitly. In addition, we carefully consider the fact that a bidder will have a sequence of
 125 opportunities to display ads to the same user. Our framework is general enough to apply either to a bidder maximizing
 126 attributed sales or incremental sales.

127 3.1 Mathematical formulation

128 When a user browses the internet, he encounters pages with ad banners to display. Each of these will trigger an auction,
 129 where advertisers may bid to buy the display opportunity. The highest bidder will then show a display to the user, and
 130 pay a price to the auctioneer. From the point of view of a bidder, the sequence of bid requests for a given user defines a
 131 stochastic process, where at each time step t :

- 132 • A bidder receives a bid request containing the *state* X_t of the user. It encodes everything relevant on this user
 133 and request, such as the user past interactions with the partner website, the timestamp, the displays made at
 134 previous time steps, the current website from which the request is received, the size of the ad on this page, etc.
- 135 • The bidder outputs a *bid* b_t for this auction.

¹not including the cost of *current* display

- 136 • It then receives a binary variable D_t telling if it did *win the display* (in which case $D_t = 1$) or not (and then
137 $D_t = 0$), and the associated *cost* $Cost_t$. (the cost is by definition 0 if $D_t = 0$).
- 138 • If the bidder wins, it can also later observe the click C_t , defined as 1 if the *display is clicked* and 0 if it is not.
139 (Clicks is not what the bidder optimizes for, so they could be ignored to define the bidder's problem; but we will
140 use clicks in the model of Section 4).

During the sequence, the user may buy some items, and at the end of the sequence, the advertiser may decide to attribute some of those sales to the bidder. The bidder thus observes the number of *attributed sales* S at the end of the sequence, and receives a *payment* $S \cdot CPA$ where CPA is a constant defining the value (in euro) of an attributed sale. The bidder processes in parallel the sequences of the thousand or sometimes millions of users in its campaign. The bidder chooses a *bidding policy* π , which is formally a mapping from the space of states to distributions on bids (so the bid at time t b_t is sampled from $\pi(x_t)$). The bidder maximizes its expected payoff:

$$\mathbb{E} \left[S \cdot CPA - \sum_{t \in \text{sequence}} Cost_t \right].$$

141 Note that this expectation depends on π , even if we do not write it explicitly to simplify the notations.

142 3.2 Assumptions

143 In order to derive a bidding formula from this model we formulate a series of assumptions.

144 ASSUMPTIONS 1 (FULLY OBSERVABLE MDP). *The state X_t is fully observable, and the sequence (X_t, b_t) of states and bids
145 on a given user forms a Markov Decision Process (MDP).*

146 ASSUMPTIONS 2 (INDEPENDENT USERS). *Each user is an independent instance of this MDP.*

147 With the first two hypotheses, the proposed framework would become an instance of a Reinforcement Learning (RL)
148 problem, where we would have one trajectory per user, the continuous actions would be the bids b_t , the reward at step
149 t would be $-Cost_t$, and the sequence ends with a final reward $S \cdot CPA$.

150 ASSUMPTIONS 3 (AUCTION STRUCTURE). *The auctions are second-price.*

151 More generally, the structure of the auctions (first price, second price, ...) is known. We will for simplicity assume
152 that those are only second price auctions, knowing that the bidding strategy could be adapted in the case of first price
153 auctions.

154 ASSUMPTIONS 4 (BID IMPACTS FUTURE ONLY THROUGH THE DISPLAY). *The next state X_{t+1} only depends on X_t and D_t ,
155 not on the bid b_t .*

156 Note that the cost $Cost_t$ and the display D_t depend on the bid; this assumptions states that both the bidder and the
157 environment forgets about the exact bid and cost at time t after observing D_t . Formally, this means that, for any policy
158 π of the bidder, the sequence of variables (X_t, D_t) forms a Markov chain.

159 REMARK 1. *Assumption 4 means that expectations on the future of the sequence after a display (or after no display),
160 such as $\mathbb{E}(S|X_t = x_t, D_t = 1)$ only depend on the policy of the bidder after time t , not on its exact bid at time t .*

161 The causal graph of Figure 1 summarizes those hypotheses.

162 REMARK 2. *Those assumptions may not be fully satisfied in practice. For example*

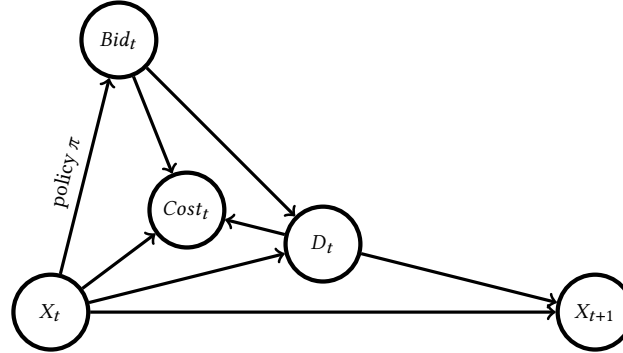


Fig. 1. Causal graph of the bidding process. The user state X_t is modeled as a Markov Chain, where the transition matrix from X_t to X_{t+1} only depends on D_t , the variable that encodes the realisation of a display at time t .

- 163 • Competitors might have some private information on the user. Since display winning rate and final sale may depend
 164 on this private information, that would make Assumption 1 wrong.
 165 • Assumption 4 would be violated if the auctioneer uses the bid at time t to compute a reserve price on a next auction
 166 for the same user at time $t + k$; or if the bidder itself decides to modulate its next bids as a function of $Cost_t$.

167 Despite those possible shortcomings, we believe those assumptions to be reasonable. We would like also to note that they are
 168 usually implicitly assumed when the display is valued with a function depending only on the state, for example with a
 169 click-through rate or conversion rate depending only on x_t , which to our knowledge includes most (if not all) prior work.

170 3.3 Greedy policy

171 As discussed before, the advertiser typically attributes the sales with some simple rule such as attributing to the last
 172 click of the user. Such a rule specifies exactly to which display the sale is attributed. This allows to *naturally* replace the
 173 final attributed sale S on the whole sequence of events by a sequence of attributed sales S_t defined at each time step: S_t
 174 is 1 if a sale is attributed to the display at time t , and 0 otherwise. In this setting, we define $S \triangleq \sum_t S_t$, and the goal is
 175 still to maximise $CPA \cdot S - \sum_t Cost_t$. Note however that the variable S_t is only observed at the end of the sequence,
 176 when the advertiser provides the attribution, and not immediately after the action (bid) at time t . In particular, S_t might
 177 be impacted by displays at later time steps.

We now define a greedy bidder: it tries to maximize, at each time step, the expected payoff at *this* time step, $\mathbb{E}(CPA \cdot S_t - Cost_t)$. It is clear that when the bidder loses the auction at time t , i.e. if $D_t = 0$, then this expected pay-off is equal to 0 (indeed the cost is 0, and no sale is attributed to a lost display). On a second price auction, it is well-known that the solution is then to bid its expected gain²:

$$b_t = CPA \cdot \mathbb{E}(S_t | X_t = x_t, D_t = 1)$$

178 On other kind of auctions, such as first price auctions, the bidder may use its knowledge (from past data) of the
 179 distribution of the clearing price to find the bid maximizing the expected gain at time t . While not optimal for the whole
 180 sequence, the greedy bidder is still widely used in practice due to its simplicity.

²Note that this expectation actually depends on the policy of the bidder after time t , because S_t may be impacted by future displays. In practice, this technicality is ignored, and the expectation is estimated with data generated with a previous version of the bidder policy.

181 **REMARK 3 (WHY THE GREEDY POLICY IS NOT OPTIMAL).** *The first reason is that the attribution is not perfect. For instance,*
 182 *supposing that a first display caused a sale that is yet to come, showing a second display will steal the attribution of the*
 183 *first display, but in total, the bidder will get the same reward (if we skipped the second display, the sale would have been*
 184 *attributed to the first). In addition, not only does the display at time t impacts the attribution, it also impacts the future*
 185 *costs. Indeed, it is quite accepted that the clicks have decreasing marginal effect on the sales probability. Hence, displays are*
 186 *likely to decrease the future costs.*

187 3.4 Optimal policy at time step τ

188 Next we derive the optimal bid at time t , assuming we know that the bidder will follow a predefined policy π in the
 189 future. We define the future cost $FCost_\tau \triangleq \sum_{t>\tau} Cost_t$, and the impact of the display on the expected reward and future
 190 costs

$$\Delta S(x_\tau) \triangleq \mathbb{E}[S|X_\tau = x_\tau, D_\tau = 1] - \mathbb{E}[S|X_\tau = x_\tau, D_\tau = 0] \quad (1)$$

191 and

$$\Delta FCost(x_\tau) \triangleq \mathbb{E}[FCost_\tau|X_\tau = x_\tau, D_\tau = 1] - \mathbb{E}[FCost_\tau|X_\tau = x_\tau, D_\tau = 0]. \quad (2)$$

192 We pinpoint that Equations (1) and (2) only refer to observational data. The following result, that relies on using Pearl's
 193 *do operator* [Pearl, 1995, 2009, Peters et al., 2017], tells us that we can nonetheless interpret them as interventional
 194 quantities.

PROPOSITION 3.1.

$$\Delta S(x_\tau) = \mathbb{E}[S|X_\tau = x_\tau, do(D_\tau = 1)] - \mathbb{E}[S|X_\tau = x_\tau, do(D_\tau = 0)]$$

and

$$\Delta FCost(x_\tau) = \mathbb{E}[FCost_\tau|X_\tau = x_\tau, do(D_\tau = 1)] - \mathbb{E}[FCost_\tau|X_\tau = x_\tau, do(D_\tau = 0)]$$

195 **PROOF.** The variable S and D_τ are d-separated in a graph were the outgoing edges from D_τ are removed. We can
 196 thus apply Pearl's Action/observation exchange [Pearl, 1995]. A similar argument applies for $Fcost_\tau$ and D_τ . \square

197 As noted previously, those quantities depend on π but do not depend on the current bid b_τ . We can now state the
 198 main result:

THEOREM 3.2 (OPTIMAL BID AT TIME STEP τ). *At time τ , the optimal bid is $b_\tau^* \triangleq \operatorname{argmax}_b \mathbb{E}[D_\tau \cdot (V(x_\tau) - Cost_\tau) | Bid_t = b]$
 with the display valuation defined as*

$$V(x_\tau) \triangleq CPA \cdot \Delta S(x_\tau) - \Delta FCost(x_\tau).$$

199 Note that the expectation in the formula of 3.2 is only on the randomness of the auction at time t . Only D_t and $Cost_t$
 200 are random variables in this formula. In particular, if the auction is second price, the optimal bid is $b^* = V(x_\tau)$. On
 201 other types of auctions, the bidder may adapt its bid accordingly, using exactly the same methods as in the case of the
 202 greedy policy.

203 **PROOF.** Noting that the cost paid at previous steps, $\sum_{t<\tau} Cost_t$, can now be viewed as a constant and thus removed
 204 from the optimisation problem, the bidder wants to maximize the expected payoff after a bid b_τ :

$$\mathbb{E}[CPA \cdot S - FCost_\tau - Cost_\tau | X_\tau = x_\tau, b_\tau]$$

205 Since the bid only impacts future through the display, when the bidder loses this auction ($D_\tau = 0$), its payoff is:

$$CPA \cdot \mathbb{E}[S|X_\tau = x_\tau, D_\tau = 0] - \mathbb{E}[FCost_\tau|X_\tau = x_\tau, D_\tau = 0]. \quad (3)$$

206 While if he wins ($D_\tau = 1$), he would receive

$$CPA \cdot \mathbb{E}[S|X_\tau = x_\tau, D_\tau = 1] - \mathbb{E}[FCost_\tau|X_\tau = x_\tau, D_\tau = 1] \quad (4)$$

207 and pay the cost $Cost_\tau$. As noted in Remark 1, the quantities in Equations (3) and (4) do not depend on the bid. Up to a
208 constant term, the payoff of the bidder is thus $D_\tau \cdot (V(x_\tau) - Cost_\tau)$.

209

□

210 3.5 Policy iteration and convergence

211 While we are not strictly in a Reinforcement Learning setup, we may define a policy iteration, which converges to the
212 optimal policy [Sutton et al., 1998] under mild technical assumptions³.

213 To be more specific: we have only derived the optimal bid at some time step τ , when the policy π at the next time
214 steps is frozen. From this, we then build a new policy: at every time step, answer with the bid which would be optimal
215 if π was used in the future. A recursion argument shows that this policy is an improvement (i.e. has a higher expected
216 return) on π at every state. Theoretically, we can then define a sequence of policies by iterating this policy improvement
217 process, converging to an optimal policy –which can be shown with a recursion on the length of the sequence when
218 those sequences are of bounded length– exactly as in the usual case of policy iteration in Reinforcement Learning.

219 4 MODELING THE CAUSAL EFFECT OF CLICKS

220 According to Theorem 3.2, we need to estimate the impact of the display on the attributed sales and on future cost,
221 $\Delta S(X_\tau)$ and $\Delta FCost(X_\tau)$. In this paper, we propose a new method to estimate $\Delta S(X_t)$ by leveraging the randomness of
222 clicks. We leave methods to estimate the impact of the display D_t on future cost $FCost_t$ to future work, and approximate
223 this impact by 0 in the online experiments.

224 Let us notice that for a given sample, we observe either what happens after $D_t = 1$ or after $D_t = 0$, but never both:
225 we need here to estimate the causal effect of the display D_t on the reward in Equation (1). Although we assume that
226 all variables are observed, this may still be challenging due to the large dimension of the state space, and the high
227 probability that $S = 0$.

In this section, we sometimes drop the index t for simplicity, as we now only care about variables at time of the
bid. Moreover, we assimilate binary random variables such as D and C to the events $D = 1$, $C = 1$. Thus we wrote
 $P(C|D, X=x)$ for $P(C=1|D=1, X=x)$; and we note \bar{D} , \bar{C} the events $D = 0$ or $C = 0$. Finally, since in practice having
several sales is very rare, we assume that S is binary, and thus note S to the event $S > 0$. This also allows us to fit S with
a logistic regression. Hence Equation (1) becomes

$$\Delta S(x) = \mathbb{P}(S|X = x, D) - \mathbb{P}(S|X = x, \bar{D}).$$

228 This quantity can be understood as “the causal effect of display D on the attributed sale S ”. We are thus trying to retrieve
229 which display(s) caused the attribution. Also, because a user is typically exposed to dozens of displays on a single day
230 and a sale is a rare event, the dilution of the sale signal prevents the use of directly fitting the model. We thus rely on a
231 few more assumptions.

³It is sufficient to assume a finite upper bound on the length of sequences.

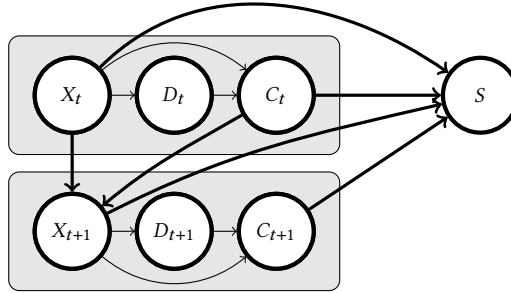


Fig. 2. Causal Graph of (X_t, D_t, C_t, S) . Comparing with Figure 1, bid and cost are omitted because we turn the focus on S . We also add the Assumption that the effect of D_t on S is mediated by C_t .

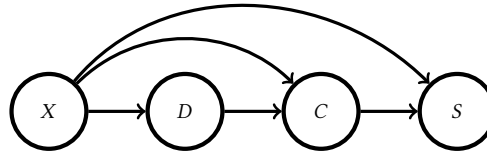


Fig. 3. Focus on the Causal graph for the Display-Click-Sale triplet (t is omitted). Only a clicked display may have an impact on the probability of seeing a sale.

232 **4.1 Assumptions**

233 A common assumption here, used implicitly by most attribution systems, is that non clicked displays have no impact
 234 on sales. This greatly reduces the number of displays one should consider when trying to find “which display might
 235 have caused the sale”.

236 ASSUMPTIONS 5 (NO POST DISPLAY EFFECT.). *The system is well represented by the causal graph in Figure 2. In particular,*
 237 *C_t blocks all directed paths starting from D_t and finishing to S .*

In other words, displays may only cause a sale through a click. This assumption is represented in the simplified graph of Figure 3. We now have:

$$\mathbb{P}(S|X = x, \bar{D}) = \mathbb{P}(S|X = x, \bar{D}, \bar{C}) = \mathbb{P}(S|X = x, D, \bar{C}).$$

238 The first equality comes from the fact that a click cannot occur without a display. The second equality can be derived
 239 using Pearl’s d-separation, after observing that $S \perp\!\!\!\perp D|(X, C)$, where $\perp\!\!\!\perp$ is the d-separation symbol.

240 REMARK 4 (DISCUSSION ON THE ASSUMPTIONS). • *The no post display effect assumption is quite commonly as-*
 241 *sumed in the industry when attributing the sales. In our case, where we are optimizing for attributed sales, this is*
 242 *even more likely to be true. For example, in the case of a “last click” attribution, the attribution will be zero if there is*
 243 *no click. While it might still be possible that a non-clicked display would cause the sale and that then another clicked*
 244 *display would make the sale to be attributed to the bidder, we believe it is reasonable to assume that attributed sales*
 245 *are caused only by clicked displays.*

- 246 • Note that in this model, the effect on displays D_t to the next state X_{t+1} is fully mediated by the click C_t . For example,
 247 this implies that the number of past displays cannot be part of the state. Since this is a feature often used by bidders,
 248 we recognize that it indicates a potential statistical relevance not taken into account in our framework.
- 249 • Finally Assumption 1 is stronger this new setting than in 3.2, because we include the clicks in the graph. Hence, there
 250 should be no unobserved variable causing both the click and the sale.

251 4.2 The incremental bidder

Since the click C is binary, we can always decompose the probability of having a sale after a display $\mathbb{P}(S|D, X=x)$ into

$$\begin{aligned} \mathbb{P}(S|D, X=x) &= \mathbb{P}(S|C, D, X=x) \cdot \mathbb{P}(C|D, X=x) + \mathbb{P}(S|\bar{C}, D, X=x) \cdot \mathbb{P}(\bar{C}|D, X=x) \\ &= \mathbb{P}(C|D, X=x) \cdot \left(\mathbb{P}(S|C, D, X=x) - \mathbb{P}(S|\bar{C}, D, X=x) \right) + \mathbb{P}(S|\bar{C}, D, X=x). \end{aligned}$$

This leads to

$$\Delta S(x) = \mathbb{P}(C|D, X=x) \cdot \left(\mathbb{P}(S|C, D, X=x) - \mathbb{P}(S|\bar{C}, D, X=x) \right) + \left(\mathbb{P}(S|\bar{C}, D, X=x) - \mathbb{P}(S|\bar{D}, X=x) \right).$$

252 Following our previous assumptions, the last term $\mathbb{P}(S|\bar{C}, D, X=x) - \mathbb{P}(S|\bar{D}, X=x)$, which we may interpret as the “pure
 253 post display effect of D on the sale”, is assumed to be 0. Hence

$$\Delta S(x) = \mathbb{P}(C|D, X=x) \cdot \left(\mathbb{P}(S|C, D, X=x) - \mathbb{P}(S|\bar{C}, D, X=x) \right). \quad (5)$$

254 All the terms in this formula may be estimated directly from the data, using standard supervised learning methods such
 255 as logistic regression. With a simple rewriting, Equation (5) becomes

$$\Delta S(x) = \mathbb{P}(C|D, X=x) \cdot \mathbb{P}(S|C, D, X=x) \cdot \left(1 - \frac{\mathbb{P}(S|\bar{C}, D, X=x)}{\mathbb{P}(S|C, D, X=x)} \right), \quad (6)$$

256 where we call the last term the *incrementality factor*. It is usually between 0 and 1 and we might think about it as “the
 257 probability that the click caused the sale”. It equals 1 if the sale is fully incremental, namely, it could not have occurred
 258 without the click: $\mathbb{P}(S|\bar{C}, D, X) = 0$. On the contrary it is valued 0 if there is no incrementality, namely if the click did not
 259 increase the sales probability: $\mathbb{P}(S|\bar{C}, D, X) = \mathbb{P}(S|C, D, X)$. While it could in theory be negative (if the click caused the
 260 user not to buy) it remains rare and in such a case we may simply submit a bid of 0, which is enough to ensure we loose
 261 the display. We clarify that the bidder is *incremental* in the sense that it bids the lift of attributed sales – or whatever is
 262 measured by the variable S : if S stands for the total number of sales, the bidder optimizes for the incremental sales.

263 **REMARK 5.** *With a greedy bidder, a display is typically valued $CPA \cdot \mathbb{P}(C_t|D_t, X=x) \cdot \mathbb{P}(S_t|C_t, D_t, X=x)$ The proposed*
 264 *model makes two changes compared to this greedy bidder:*

- 265 (1) *We multiply by the incrementally factor from Equation (6).*
 266 (2) *We replace S_t , the sale attributed to display D_t by S (sales attributed to the whole sequence), in the $\mathbb{P}(S|C, D, X=x)$*
 267 *part of the model. We pinpoint that our model does not depend on the method used to attribute sales to displays*
 268 *(typically last click) as long as it is attributed to the sequence.*

Table 1. Our generative model assumes that, given the display type ($T = a, n$ or y) and whether it is clicked C or not \bar{C} , we can determine if it leads to a sale S or not \bar{S} .

\cap	$T = a$	$T = n$	$T = y$
\bar{C}	S	\bar{S}	\bar{S}
C	S	\bar{S}	S

5 AN INCREMENTALITY METRIC

We are interested in valuing more the displays that could lead a user to a sale in the future while he would have not bought without. Namely, we would like to evaluate how good our model is at predicting $\Delta S(x)$. Given Equation (5), $\Delta S(x)$ can be split into two parts, first $\mathbb{P}(C|D, X=x)$ that is a quite mature model in the industry and a second model $\mathbb{P}(S|C, D, X=x) - \mathbb{P}(S|\bar{C}, D, X=x)$ that we have introduced in Section 4.2. Since we cannot observe both the clicked display and its counterfactual, assessing the quality of such a model offline is a complex task. In this Section, we build a random variable Y such that

$$\mathbb{P}(Y|C, D, X=x) = \mathbb{P}(S|C, D, X=x) - \mathbb{P}(S|\bar{C}, D, X=x)$$

that we cannot observe. However, we introduce a metric that is able to evaluate the capacity of our model to predict this variable Y while observing only the click C and sales S outcomes.

5.1 The generative model

As in Section 4.2 we consider only won displays, that are described by a context X , and have two observable effects: they can be clicked C or not \bar{C} and they can be followed by a sales S or not \bar{S} . Intuitively, we would like to divide those displays in four classes, which we call their *display type* $T \in \{a, n, y, d\}$:

- (1) The displays that **always** lead to a sale whether they are clicked or not : $T = a$.
- (2) The displays that **never** lead to a sale whether they are clicked or not : $T = n$.
- (3) The **incremental** displays that lead to a sale if they are clicked but do not if they are not clicked: $T = y$.
- (4) The **decremental** displays that do not lead to a sale if they are clicked but do if they are not clicked: $T = d$.

ASSUMPTIONS 6 (NO DECREMENTAL DISPLAYS). We assume for all x , $\mathbb{P}(S|C, X = x) \geq \mathbb{P}(S|\bar{C}, X = x)$.

This mean that there are no decremental displays (or equivalently, the incremental factor is positive). This assumption is quite natural in practice since we do not expect a clicked ad to discourage a user to buy a product. Of course, this display type is never observed directly, but for any given context x , we can always assume that the variable S is a function of C and an independent unobserved variable $T \in \{a, n, y\}$ as defined in Table 1⁴. The corresponding causal graph is depicted in Figure 4.

Since we are looking for incrementality, $T = y$ is our variable of interest, even if it cannot be observed from the data. Indeed, the displays that have been clicked and then have led to a sale ($C \cap S$) and the display that were not clicked and have not led to a sale ($\bar{C} \cap \bar{S}$) might either be incremental or not. To determine if they are incremental one should have observed the counterfactual. Similarly to \bar{C} and \bar{D} we denote by Y the random variable associated to the event $T = y$ and \bar{Y} to $T \neq y$.

⁴For example, we could define $\mathbb{P}(T = a|X = x) := \mathbb{P}(S|X = x, \bar{C})$ and $\mathbb{P}(T = n|X = x) := \mathbb{P}(\bar{S}|X = x, C)$.

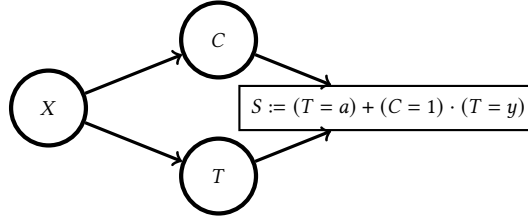


Fig. 4. Causal graph of the generative model. The display context X influences the display type $T \in \{a, n, y, d\}$ (specifically if it is incremental Y) and whether the display is clicked C . Then, T and C determine if this display will lead to a sale S following Table 1. Note that there is no causation from C to Y to illustrate the independence assumption of Proposition 5.1.

297 5.2 Reverted incremental likelihood

Y is a Bernoulli variable, hence if we have an estimator f that predicts Y given X its log-likelihood is

$$LLH_B(Y, f(X)) = Y \log f(X) + \bar{Y} \log(1 - f(X)).$$

298 Since we cannot observe Y we cannot compute this likelihood directly from the data but Proposition 5.1 gives us an
 299 unbiased estimate of this likelihood based only on the observed values C and S . This unbiased estimate *reverts the label*
 300 as in [Jaskowski and Jaroszewicz, 2012] while taking into account the partial randomness of the clicks.

PROPOSITION 5.1. *For any predictor f , the reverted incremental likelihood*

$$RLH_B(C, S, f(X)) = \frac{C}{\mathbb{P}(C|X)} LLH_B(S, f(X)) + \frac{\bar{C} \cap S}{\mathbb{P}(\bar{C}|X)} \log \frac{1 - f(X)}{f(X)}$$

301 *is an unbiased estimator of the direct incremental likelihood $LLH_B(Y, f(X))$.*

302 Provided we have a correct $\mathbb{P}(C|X)$ model, we can thus estimate the incremental likelihood from observable data.
 303 We defer the proof of Proposition 5.1 to Appendix A. The reverted incremental likelihood only relies on Assumption 6.
 304 This is interesting in practice since it allows us to evaluate offline the incremental performance of our models to
 305 perform model selection. In the case of the incremental bidder, it is useful to assess the quality of the incremental
 306 factor as a whole instead of evaluating $\mathbb{P}(S|\bar{C}, D, X=x)$ on one hand and $\mathbb{P}(S|C, D, X=x)$ on the other while having no
 307 guarantee on the quality of their ratio. Note that this still requires a click prediction model to estimate $\mathbb{P}(C|X)$. Relying
 308 on a model to assess the performances of another is not ideal but since the click prediction models are mature in the
 309 industry and trained on much more data, we can safely use them to evaluate sales prediction models. In Appendix B,
 310 we illustrate Proposition 5.1 on a data set that we simulate such that we can observe the hidden variable T . That allows
 311 us to compute the direct incremental likelihood and to compare it the reverted one to assess the performance of the
 312 incremental bidder.

313 6 IMPLEMENTATION AND RESULTS

314 6.1 Offline analysis

315 In order to evaluate our methodology, we run experiments on the Criteo Attribution Modeling for Bidding public data
 316 set [Diemert et al., 2017]. This data set consists in 16 million displays sent by Criteo on 6 million users over a period of
 317 30 days. To each display is associated a set of context features (that we have denoted by X_t so far), if the display was
 318 clicked (denoted by C_t so far), and if it has led to a sale that was attributed to Criteo (denoted by S_t so far). For simplicity,

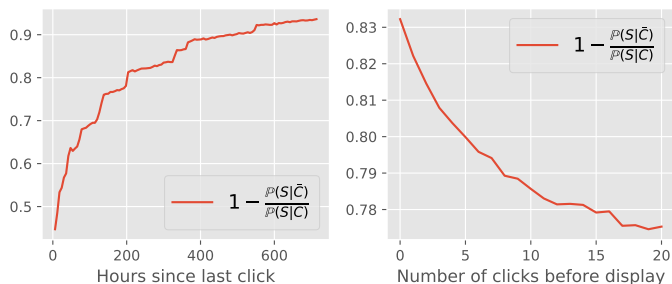


Fig. 5. Value of the incremental factor given the number of hours elapsed since the latest click. As our intuition would suggest, the most recent the click is, the less incremental the display is expected to be. Similarly, if a user has already clicked on several displays then the following clicked displays are less likely to be incremental.

319 we assume the value of an attributed sale (denoted by CPA) to equal 1. As a baseline, we train the greedy bidder that
 320 values a display with context features x by $\mathbb{P}(C_t|X=x) \cdot \mathbb{P}(S_t|C_t, X_t=x)$ (see Remark 5). Note that the context features
 321 are hashed [Weinberger et al., 2009], and each context x ends up being represented by sparse vector of dimension 2^{16} .
 322 Then, since we have access to the user identifier, we can reconstruct the user timelines. Specifically, we can determine
 323 if the sequence of displays has led to an attributed sales (denoted by S so far). With this label, we train the incremental
 324 bidder derived in Section 4.2 and compute the incremental factor from Equation (6)⁵.

325 In Figure 5, we report the average value of this factor over the displays that we have grouped in two manners, first
 326 given the time elapsed since the latest user’s click, and second, given the number of clicks already made by the user in
 327 the timeline. In both cases, we observe that the incremental factor illustrates what we could have foreseen: the more
 328 recent the click is, the less incremental the display is expected to be and, similarly, if a user has already clicked on
 329 several displays then the following clicked displays are less likely to be incremental. Note that the first relationship
 330 –between the incrementality and the time elapsed since the latest click– is at the heart of the work in [Diemert et al.
 331 2017, Figure 5] where they model it with one minus a decreasing exponential parameterized with a decay factor λ . It is
 332 interesting to see that the shape is similar to what we observe with the incremental factor but also that our approach
 333 learns a clean monotonic relationship with a non parametric model (namely we do not specify that we look for a
 334 decaying effect parameterized with a factor λ to learn). Also, our incremental bidder is not limited to the time elapsed
 335 since the latest click but can learn from all the available features (such as number of clicks before display for example).

336 On the same data set, we compare the performances of the greedy and the incremental bidders evaluated with
 337 the incremental likelihood computed thanks to Proposition 5.1. In both cases, we consider that the click prediction is
 338 evaluated separately and we isolate the part of the formula that gets multiplied by $CPA \cdot \mathbb{P}(C_t|D_t, X=x)$. That is, for the
 339 greedy bidder we use $\mathbb{P}(S_t|C_t, D_t, X=x)$ as predictor of Y and $\mathbb{P}(S|C_t, D_t, X_t=x) - \mathbb{P}(S|\bar{C}_t, D_t, X_t=x)$ for the incremental
 340 bidder. We train the models on 80% of the users and keep the remaining 20% for a test set. We report the results in
 341 Figure 6 where we varied the size of the hashing space from 2^{10} to 2^{16} . We observe that both models obtain better
 342 performances when their predictive capacities (induced by the feature space’s size) are increased. Also, we observe that
 343 the incremental bidder obtains better results than the greedy bidder. It reaches better performances than the greedy
 344 bidder with 1,000 features instead of 60,000.

⁵Note that all the experiments conducted on this data set are hosted on Google colab (url removed for anonymity reasons).

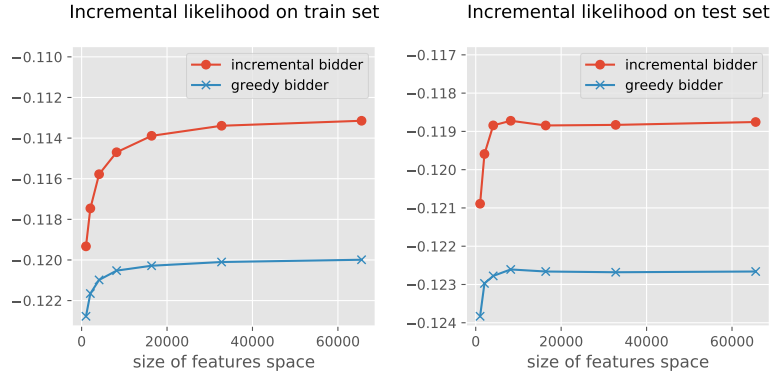


Fig. 6. Value of the incremental likelihood for our bidding formula compared to the baseline, i.e. last click model. We plot it in regards to the size of the features space that quantifies the capacity of the model.

345 One could wonder if we could train a model f to directly maximize the reverted incremental likelihood. Indeed, it
 346 sounds reasonable especially since, if $f : x \mapsto \sigma(w^T x)$ where σ is the sigmoid function as in a logistic regression, then
 347 it becomes a convex optimization problem. However, the Goodhart law stands that “when a measure becomes a target,
 348 it ceases to be a good measure”, and it might be true in our case since the variance induced by the ratio $\frac{1}{\mathbb{P}(C|X)}$ could
 349 lead to high overfitting. We thus leave it for future work.

350 6.2 Live experiment

Finally, we explain how we modify a production bidding algorithm to test our model online. The baseline is made of
 (1) a module that predicts the probability that an opportunity, if won, will be attributed a sale with the last click rule
 (the greedy bidder), and (2) a module that shades this prediction to account for the first-price auction rule. Our change
 simply consists in (a) learning $\mathbb{P}(S|\bar{C}, D, X)$ and $\mathbb{P}(S|C, D, X)$ with a logistic regression, (b) form the factor

$$\alpha(X) = \left(1 - \frac{\mathbb{P}(S|\bar{C}, D, X)}{\mathbb{P}(S|C, D, X)} \right),$$

351 (c) multiply the production module of (1) by this *incremental factor*. Many feature engineering decision were taken in
 352 step (2b), among other, we decided to learn $\alpha(X)$ with a logistic regression.

353 We tested the change on and obtained a 7.8% decrease of spend for the same campaign performance in attribution on
 354 a volume worth 7,000,000 euros on the test period, which is something remarkable in the industry.

355 7 DISCUSSION

356 We introduce a reinforcement learning formulation of the bidding problem in display advertising. We derived a strategy
 357 thanks to a causal reasoning approach. The main assumption is to suppose that only clicked displays bring additional
 358 reward. We also provide an offline metrics to sanitize the approach. While the framework could be used for incremental
 359 sales, we tested it on attributed sales for the reasons given in the introduction (it is the main metric in the industry and
 360 it is methodologically easier to test and develop on attribution).

On the experimental and implementation aspects, further work could be done, in particular to take into account the evolution of the expected future cost. We believe our framework is quite general and could be adapted to other contexts with clicks, displays and actions, in particular marketplace and recommender systems.

From a methodological perspective, we think that attribution can be useful to understand, test and improve causal models (which can later be adapted for incremental sales). From an industry wise perspective, our approach could be combined with incremental attribution framework such as [Singal et al., 2019]. Such a decomposition offers several advantage as the role of the different agents are clearly defined.

ACKNOWLEDGMENTS

We would like to thank Aloïs Bissuel, David Rohde, Eustache Diemert, Flavian Vasile, Guillaume Genthial, Ieva Grublyte, Nicolas Chrysanthos, Olivier Koch and Vincent Grosbois for their comments and suggestions.

REFERENCES

- D. Agarwal, S. Ghosh, K. Wei, and S. You. Budget pacing for targeted online advertisements at linkedin. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1613–1619, 2014.
- R. Berman. Beyond the last touch: Attribution in online advertising. *Marketing Science*, 37(5):771–792, 2018.
- T. Blake, C. Nosko, and S. Tadelis. Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica*, 83(1):155–174, 2015.
- H. Cai, K. Ren, W. Zhang, K. Malialis, J. Wang, Y. Yu, and D. Guo. Real-time bidding by reinforcement learning in display advertising. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, WSDM '17*, page 661–670, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450346757. doi: 10.1145/3018661.3018702. URL <https://doi.org/10.1145/3018661.3018702>.
- O. Chapelle, E. Manavoglu, and R. Rosales. Simple and scalable response prediction for display advertising. *ACM Transactions on Intelligent Systems and Technology*, 5(4), 2014. doi: 10.1145/2532128.
- Y. Chen, P. Berkhin, B. Anderson, and N. R. Devanur. Real-time bidding algorithms for performance-based display ad allocation. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1307–1315, 2011.
- H. Choi, C. F. Mela, S. R. Balseiro, and A. Leary. Online display advertising markets: A literature review and future directions. *Information Systems Research*, 2020.
- V. Conitzer, C. Kroer, E. Sodomka, and N. E. Stier-Moses. Multiplicative pacing equilibria in auction markets. In *Web and Internet Economics: 14th International Conference, WINE 2018, Oxford, UK, December 15–17, 2018, Proceedings*, volume 11316, page 443. Springer, 2018.
- B. Dallessandro, O. Stitelman, C. Perlich, and F. Provost. Causally motivated attribution for online advertising. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012. doi: 10.1145/2351356.2351363.
- S. Despotakis, R. Ravi, and A. Sayedi. First-price auctions in online display advertising. Available at SSRN 3485410, 2019.
- E. Diemert, J. Meynet, P. Galland, and D. Lefortier. Attribution modeling increases efficiency of bidding in display advertising. In *Proceedings of the ADKDD'17*, pages 1–6. ACM, 2017.
- E. Diemert, A. Héliou, and C. Renaudin. Off-policy learning for causal advertising, 2018.
- J. Fernandez-Tapia. An analytical solution to the budget-pacing problem in programmatic advertising. *Journal of Information and Optimization Sciences*, 40(3):603–614, 2019.
- L. Fisher. US programmatic digital display ad spending, Nov 2019. URL <https://www.emarketer.com/content/us-programmatic-digital-display-ad-spending>.
- N. Grislain, N. Perrin, and A. Thabault. Recurrent neural networks for stochastic control in real-time bidding. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019. doi: 10.1145/3292500.3330749.
- B. Heymann. Cost per action constrained auctions. In *Proceedings of the 14th Workshop on the Economics of Networks, Systems and Computation, NetEcon '19*, New York, NY, USA, 2019. Association for Computing Machinery. doi: 10.1145/3338506.3340269.
- B. Heymann. How to bid in unified second-price auctions when requests are duplicated. *Operations Research Letters*, 2020.
- M. Jaskowski and S. Jaroszewicz. Uplift modeling for clinical trial data. In *ICML Workshop on Clinical Data Analysis*, 2012.
- W. Ji, X. Wang, and D. Zhang. A probabilistic multi-touch attribution model for online advertising. *International Conference on Information and Knowledge Management, Proceedings*, 24-28-Octo:1373–1382, 2016. doi: 10.1145/2983323.2983787.
- G. A. Johnson, R. A. Lewis, and E. I. Nubbemeyer. Ghost Ads: Improving the economics of measuring online ad effectiveness. *Journal of Marketing Research*, 54(6):867–884, Dec. 2017. doi: 10.1509/jmr.15.0297.
- V. Krishna. *Auction theory*. Academic press, 2009.
- K.-C. Lee, A. Jalali, and A. Dasdan. Real Time Bid Optimization with Smooth Budget Delivery in Online Advertising. In *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising - ADKDD '13*, pages 1–9, New York, New York, USA, 2013. ACM Press. doi: 10.1145/

- 410 2501040.2501979.
- 411 R. A. Lewis and J. M. Rao. The unfavorable economics of measuring the returns to advertising. *Quarterly Journal of Economics*, 130(4):1941–1973, Nov.
- 412 2015. doi: 10.1093/qje/qjv023.
- 413 R. A. Lewis, J. M. Rao, and D. H. Reiley. Here, there, and everywhere: Correlated online behaviors can lead to overestimates of the effects of advertising. In
- 414 *Proceedings of the 20th International Conference on World Wide Web, WWW 2011*, pages 157–166, 2011. doi: 10.1145/1963405.1963431.
- 415 H. B. McMahan, G. Holt, D. Sculley, M. Young, D. Ebner, J. Grady, L. Nie, T. Phillips, E. Davydov, D. Golovin, et al. Ad click prediction: a view from the
- 416 trenches. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1222–1230, 2013.
- 417 D. Moriwaki, Y. Hayakawa, I. Munemasa, Y. Saito, and A. Matsui. Unbiased lift-based bidding system. *arXiv preprint arXiv:2007.04002*, 2020.
- 418 J. Pearl. Causal diagrams for empirical research. *Biometrika*, 82(4):669–688, 1995.
- 419 J. Pearl. *Causality*. Cambridge University Press, 2009. doi: 10.1088/1751-8113/44/8/085201.
- 420 J. Peters, D. Janzing, and B. Schölkopf. *Elements of causal inference*. The MIT Press, 2017.
- 421 T. Rahier, A. Héliou, M. Martin, C. Renaudin, and E. Diemert. Individual treatment effect estimation in a low compliance setting, 2020.
- 422 R. Singal, O. Besbes, A. Desir, V. Goyal, and G. Iyengar. Shapley meets uniform: An axiomatic framework for attribution in online advertising. *The Web*
- 423 *Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, pages 1713–1723, 2019. doi: 10.1145/3308558.3313731.
- 424 R. S. Sutton, A. G. Barto, et al. *Introduction to reinforcement learning*, volume 135. MIT Press Cambridge, 1998.
- 425 J. Wang, W. Zhang, and S. Yuan. Display advertising with real-time bidding (rtb) and behavioural targeting. *arXiv preprint arXiv:1610.03013*, 2016.
- 426 K. Weinberger, A. Dasgupta, J. Langford, A. Smola, and J. Attenberg. Feature hashing for large scale multitask learning. In *Proceedings of the 26th annual*
- 427 *international conference on machine learning*, pages 1113–1120, 2009.
- 428 J. Xu, K.-c. Lee, W. Li, H. Qi, and Q. Lu. Smart pacing for effective online ad campaign optimization. In *Proceedings of the 21th ACM SIGKDD International*
- 429 *Conference on Knowledge Discovery and Data Mining*, pages 2217–2226, 2015.
- 430 J. Xu, X. Shao, J. Ma, K. C. Lee, H. Qi, and Q. Lu. Lift-based bidding in ad selection. *30th AAAI Conference on Artificial Intelligence, AAAI 2016*, 2016.
- 431 X. Yang, Y. Li, H. Wang, D. Wu, Q. Tan, J. Xu, and K. Gai. Bid optimization by multivariable control in display advertising. *Proceedings of the ACM SIGKDD*
- 432 *International Conference on Knowledge Discovery and Data Mining*, 2019. doi: 10.1145/3292500.3330681.
- 433 Y. Zhang, Y. Wei, and J. Ren. Multi-touch Attribution in Online Advertising with Survival Theory. *Proceedings - IEEE International Conference on Data*
- 434 *Mining, ICDM, 2015-Janua(January):687–696*, 2014. doi: 10.1109/ICDM.2014.130.

435 Appendices

436 A PROOF OF PROPOSITION 5.1

With Assumption 6, \bar{Y} is the disjoint union of $T = a$ and $T = n$, hence

$$\mathbb{P}[\bar{Y}|X] = \mathbb{P}[T = a|X] + \mathbb{P}[T = n|X].$$

As the display types T are built independently of C given X , the previous equality can be written as

$$\mathbb{P}[\bar{Y}|X] = \frac{\mathbb{P}[\bar{C} \cap T = a|X]}{\mathbb{P}(\bar{C}|X)} + \frac{\mathbb{P}[C \cap T = n|X]}{\mathbb{P}(C|X)}.$$

Finally, since we can identify event $\bar{C} \cap (T = a)$ to $\bar{C} \cap S$ and $C \cap (T = n)$ to $C \cap \bar{S}$ (see Table 1), it writes

$$\mathbb{P}[\bar{Y}|X] = \frac{\mathbb{P}[\bar{C} \cap S|X]}{\mathbb{P}(\bar{C}|X)} + \frac{\mathbb{P}[C \cap \bar{S}|X]}{\mathbb{P}(C|X)},$$

437 and

$$\mathbb{P}[Y|X] = \frac{\mathbb{P}[C \cap S|X]}{\mathbb{P}(C|X)} - \frac{\mathbb{P}[\bar{C} \cap S|X]}{\mathbb{P}(\bar{C}|X)}. \quad (7)$$

Using this identity we compute the expectation of the incremental likelihood

$$\begin{aligned}
& \mathbb{E}_{X,Y}[\text{LLH}_B(Y, f(X))] \\
&= \mathbb{E}_X[\mathbb{E}_Y[Y \log f(X) + \bar{Y} \log(1 - f(X))|X]] \\
&= \mathbb{E}_X[\mathbb{P}(Y|X) \log f(X) + \mathbb{P}(\bar{Y}|X) \log(1 - f(X))] \\
&= \mathbb{E}_X \left[\left(\frac{\mathbb{P}[C \cap S|X]}{\mathbb{P}(C|X)} - \frac{\mathbb{P}[\bar{C} \cap S|X]}{\mathbb{P}(\bar{C}|X)} \right) \log f(X) \right. \\
&\quad \left. + \left(\frac{\mathbb{P}(\bar{C} \cap S|X)}{\mathbb{P}(\bar{C}|X)} + \frac{\mathbb{P}[C \cap \bar{S}|X]}{\mathbb{P}(C|X)} \right) \log(1 - f(X)) \right]
\end{aligned}$$

Since all is conditioned on X and C and S are Bernoulli variables,

$$\begin{aligned}
& \mathbb{E}_{X,Y}[\text{LLH}_B(Y, f(X))] \\
&= \mathbb{E}_{X,C,S} \left[\left(\frac{C \cap S}{\mathbb{P}(C|X)} - \frac{\bar{C} \cap S}{\mathbb{P}(\bar{C}|X)} \right) \log f(X) \right. \\
&\quad \left. + \left(\frac{\bar{C} \cap S}{\mathbb{P}(\bar{C}|X)} + \frac{C \cap \bar{S}}{\mathbb{P}(C|X)} \right) \log(1 - f(X)) \right] \\
&= \mathbb{E}_{X,S,C} \left[\frac{C}{\mathbb{P}(C|X)} \text{LLH}_B(S, f(X)) \right. \\
&\quad \left. + \frac{\bar{C} \cap S}{\mathbb{P}(\bar{C}|X)} \log \frac{1 - f(X)}{f(X)} \right].
\end{aligned}$$

438 B COMPARISON OF DIRECT AND REVERTED INCREMENTAL LIKELIHOOD ON A SIMULATED DATA 439 SET

440 In order to manifest the consistency of the incremental metric introduced in Proposition 5.1 and link it to the incremental
441 factor, we generate a toy data set where the displays are described by a 2-dimensional feature vector x , allowing us to
442 represent them on a 2d graph. Then, following our generative model, we assume that each display i , gets clicked with a
443 probability $\psi(x_i) \in (0, 1)$ and is of type $t_i = a, y$ or c with a probability that also depends on x_i . Figure 7 shows $n = 60$
444 displays represented by their feature vector $x_i \in \mathbb{R}^2$ and their click c_i , type t_i and sales s_i characteristics for $i = 1, \dots, n$
445 ⁶.

Since the data is generated, we have access to the true value of the display type t_i and more specifically to the binary variable y_i (that we identify to the event $t_i = y$) that characterizes an incremental display. Hence, for a model $f : \mathbb{R}^2 \rightarrow (0, 1)$ supposed to predict y_i , we can compute the averaged direct incremental likelihood over n displays $\frac{1}{n} \sum_{i=1}^n \text{LLH}_B(y_i, f(x_i))$. We provide in Table 2, the results for two models. First, one that directly predicts the probability $\mathbb{P}(Y|X = x_i)$. Similarly to the metric, this model is trainable only if the label y_i is observed, i.e. only on a simulated data set. The second one actually predicts the incremental part of the bidding formula (5), with two models, one trained to predict $\mathbb{P}[S|C, X = x_i]$ and a second one to predict $\mathbb{P}[S|\bar{C}, X = x_i]$. This term is actually directly linked to $\mathbb{P}(Y|X = x_i)$ since from Equation (7), we derive

$$\mathbb{P}(Y|X) = \mathbb{P}[S|C, X = x_i] - \mathbb{P}[S|\bar{C}, X = x_i].$$

⁶Note that all the experiments conducted on this simulated data set are hosted on Google colab (url removed for anonymity reasons).

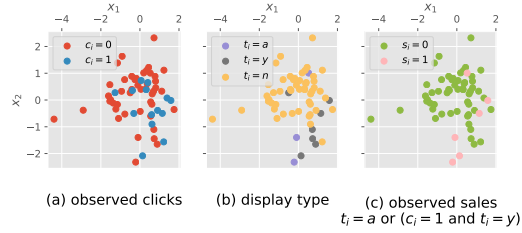


Fig. 7. Illustration of the generative model described in Table 1 where each display is depicted by its feature vectors $x \in \mathbb{R}^2$. On a real data set we observe only if a display is clicked (a) and if it leads to a sale (c), but our variable of interest is the unobservable display type (b). Though, the metric we have introduced in Proposition 5.1 allows us to evaluate the capacity of a model to predict the unobservable event $T = y$.

Table 2. Value of the *direct* incremental likelihood obtained thanks to the variable Y that we can observe on a simulated data set. We observe that without observing Y , the incremental factor achieves nearly as good performances as the theoretical maximizer $\mathbb{P}(Y|X = x)$.

	$n = 100$	$n = 1000$	$n = 10000$
$\mathbb{P}(Y X = x)$	-0.289 +/- 0.016	-0.274 +/- 0.005	-0.275 +/- 0.002
Inc. bidder	-0.293 +/- 0.017	-0.280 +/- 0.006	-0.282 +/- 0.003

Table 3. Value of the *reverted* incremental likelihood obtained thanks to Proposition 5.1. At the expense of a wider 95% confidence interval, it converges as expected to the same value as the direct likelihood (see Table 2) without observing Y .

	$n = 100$	$n = 1000$	$n = 10000$
$\mathbb{P}(Y X = x)$	-0.296 +/- 0.090	-0.274 +/- 0.025	-0.277 +/- 0.007
Inc. bidder	-0.355 +/- 0.147	-0.294 +/- 0.034	-0.280 +/- 0.010

446 *This observation makes our bidding formula a natural candidate to optimize for this incrementality metric* and we indeed
 447 observe in Table 2 that its performances are really close to the one obtained with the theoretical model $\mathbb{P}(Y|X = x_i)$.

448 Finally, we present in Table 3 the value we obtain by using the reverted incremental likelihood from Proposition 5.1:
 449 $\frac{1}{n} \sum_{i=1}^n \text{RLLH}_B(c_i, s_i, f(x_i))$ given the predictions $\mathbb{P}(C|X = x_i)$ for all $i = 1, \dots, n$ obtained from another model trained
 450 separately. We numerically illustrate the unbiasedness of the reverted incremental likelihood since we observe that it
 451 converges to the same values as the direct one. While it has more variance (leading to wider 95% confidence intervals),
 452 it is interesting to emphasize that this metric does not need to observe the hidden variable Y to correctly converge.