The Multiple Attribution Problem in Pay-Per-Conversion Advertising

Patrick Jordan[1], Mohammad Mahdian[1], Sergei Vassilvitskii[1], and Erik Vee[1]

Yahoo! Research, Santa Clara, CA 95054.
{prjordan, mahdian, sergei, erikvee}@yahoo-inc.com

Abstract. In recent years the online advertising industry has witnessed a shift from the more traditional pay-per-impression model to the pay-per-click and more recently to the pay-per-conversion model. Such models require the ad allocation engine to translate the advertiser's value per click/conversion to value per impression. This is often done through simple models that assume that each impression of the ad stochastically leads to a click/conversion independent of other impressions of the same ad, and therefore any click/conversion can be attributed to the last impression of the ad. However, this assumption is unrealistic, especially in the context of pay-per-conversion advertising, where it is well known in the marketing literature that the consumer often goes through a *purchasing funnel* before they make a purchase. Decisions to buy are rarely spontaneous, and therefore are not likely to be triggered by just the last ad impression. In this paper, we observe how the current method of attribution leads to inefficiency in the allocation mechanism. We develop a fairly general model to capture how a sequence of impressions can lead to a conversion, and solve the optimal ad allocation problem in this model. We will show that this allocation can be supplemented with a payment scheme to obtain a mechanism that is incentive compatible for the advertiser and fair for the publishers.

1 Introduction

In 2009 Internet ad revenues totaled \$22.7B, of which sponsored search and display advertising accounted for 47% and 22%, respectively [14]. Although still a relatively nascent industry, the mechanism for advertising on the internet has evolved considerably over the past two decades. Initially, ads were sold on a purely CPM (cost-per-mille) basis, and it was the number of impressions that determined the payment made by the advertiser. As the marketplace matured, publishers allowed advertisers to pay per click (CPC basis), and, more recently per action [12] or conversions (CPA basis).

Auction mechanisms play a critical role in both of these formats [8] and the celebrated Generalized Second Price (GSP) mechanism has been extensively studied and analyzed [7, 1]. A crucial assumption behind the analyses is a simplistic model of user behavior, namely that the probability of the user clicking on the ad is *independent* of the number of times the user has previously viewed the ad. This is equivalent to assuming that showing the ad does not result in changes to future user behavior. The flaw in this reasoning is best illustrated by the fact that an ad loses its effectiveness over time, and the click probabilities are not going to be identical for the first and the thousandth view of the same ad by the same user [11]. If the user has not reacted (via a click or conversion) to an ad after the first 999 impressions, it is highly unlikely that the thousandth one is going to change her mind. Conversely, the first few impressions may result in a superlinear increase in conversion probability (much like having a second friend in a group increases greatly increases the probability of the user joining [2]).

This notion is known as a purchase funnel [3], and has been at the core of the marketing literature for almost a century [15]. In online advertising various, this is recognized as a major issue (see, for example, [6, 10, 5]), and analytics tools and ad hoc methods have been developed to reflect the consequences of this type of user behavior. For example, the practice of frequency capping [4, 9], whereby an advertiser limits the number of exposures of his ad to any user, is a crude way to optimize for the fact that an ad loses its effectiveness after a certain number of views. On the other hand, to the best of our knowledge, current mechanisms do not reward the publisher for displaying an ad that may not result in a click until its second or third view.

As a concrete example, consider an ad that never results in a click on the first impression, but always results in a click on the second impression. (We go through a more elaborate example in the next Section.) In this case, even with perfect click probability estimation the ad will never be shown, since every publisher does a myopic optimization, and the ad in question is guaranteed to have a zero payoff on its first view. In order to create proper incentives to the publishers, the mechanism designer must recognize that a given click or conversion is not simply result of the actions of the last publisher (as it is attributed today), but rather a result of the *aggregate* actions of all of the previous publishers. Therefore, to ensure maximum efficiency, one must attribute the conversion (and the payoffs that go with it) to all of the publishers along the chain.

In this work we mathematically formulate the *multiple attribution problem* and explore the proper method for transforming a bid per conversion to an effective bid per impression to ensure maximum efficiency. We remark that the multiple attribution problem is not only relevant to web advertising scenarios. For example, consider the problem faced by a website designer facing an increase in user traffic. Is that increase due to the last change made on the site, or is it due to the continuous work and the multitude of changes done over the past year. Similarly, suppose a brick and mortar retailer is losing clients to an online merchant. How much of that loss should be attributed to the recent history, and how much to an effect accumulated over a longer time horizon.

In addition to the optimal allocation problem in a multiple attribution setting, we explore the associated pricing problem. This problem is complicated by two constraints: a pay-per-conversion advertiser must pay only when a conversion occurs; and different impressions might be served on different publishers, and therefore it also matters how the payment of the advertiser is split between these publishers. While the first constraint can be satisfied easily, we can only prove that we can simultaneously satisfy both constraints in a special case where the opportunity cost is a constant. Our proof uses the max-flow min-cut theorem.

The rest of this paper is organized as follows: in the next section, we show how attributing a conversion to the last impression can lead to inefficiencies in the market. This motivates a model (defined in Section 3) that assumes that the user follows a Markovian process. The optimal allocation problem for this model is formulated in Section 4 as a Markov Decision Process. We solve the Bellman equations for this process in Section 5, getting a closed-form solution in a special case and a method to compute the values and the effective bid-per-impression in general. In Section 6, we prove that the allocation mechanism admits a pay-per-conversion payment scheme that is incentive compatible for the advertiser and (in the case that the opportunity cost is a constant) fair for the publishers. We conclude in Section 7 with a discussion of how our results can be generalized and applied in practice.

2 Inefficiency of the last- impression attribution scheme

The model that attributes each conversion to the last impression of the ad is built on the assumption that upon each impression, the user stochastically decides whether or not to purchase the product, independent of the number of times she has previously seen the ad. However, this is not an accurate assumption in practice, and when this assumption is violated, the last-impression attribution scheme can be inefficient. Here we explain this with a simple scenario: focus on one pay-per-conversion advertiser that has a value of $1 per conversion. Assume a user sees the ad of this advertiser four times on average. The probability of converting after viewing the ad for the first time is 0.02, and after the second viewing this probability increases to 0.1. The third and the fourth viewing of the ad will not lead to any conversions. Also, assume that this ad always competes with a pay-per-impression ad with a bid of 4 cents per impression.

First, consider a system that simply computes the average conversion rate of the ad and allocates based on that. This method would estimate the conversion rate of the ad at $(0.02 + 0.1 + 0 + 0)/4 = 0.03$. Therefore, the ad's effective bid per impression is 3 cents and the ad will always lose to the competitor. This is inefficient, since showing the ad twice gives an average expected value of 6 cents per impression, which is more than the competitor.

If we employ frequency capping and restrict the ad to be shown at most twice to each user, the above problem would be resolved, but another problem arises. In this case, the average conversion rate will be $(0.02+0.1)/2 = 0.06$, and the ad will win both impressions. This is indeed the efficient outcome, but let us look at this outcome from the perspective of the publishers. If the two impressions are on different publishers, the first publisher only gets 2 cents per impression in expectation, less than what the competitor pays. This is an unfair outcome, and means that this publisher would have an incentive not to accept this ad, thereby creating inefficiency.

Finally, note that even if the conversion rate is estimated accurately for each impression, still the usual mechanism of allocating based on expected value per impression is inefficient, since it will estimate the expected value per impression at 2 cents for the first impression. This will lose to the 4 cent competitor, and never gives the ad a chance to secure the second, more valuable, impression.

3 The Model

In this section we formalize a model that captures the fact that the user goes through a purchase funnel before buying a product, and therefore the conversion probability of an ad depends on the number of times the ad is shown. We model the user's behavior from the perspective of one pay-per-conversion advertiser A. We have T opportunities to show an ad to the user, where T is a random variable. For simplicity, we assume that T is exponentially distributed. This means that there is a fixed drop-out probability $q \in (0, 1)$, and every time a user visits a page on which an ad can be shown, there is a probability q that she will drop out after that and will not come back to another such page. Every time we have an opportunity to show an ad to this user, we must decide whether to show A's ad or the competitor's ad. Assume the value per impression of the competitor's ad is R. In other words, R is the opportunity cost of showing A's ad. We assume that R is a random variable, and is independently and identically distributed each time. We will present some of our results in the special case that R is a constant (corresponding to the case that A always faces the same competitor with a fixed value), since this case simplifies the math and allows for closed-form solutions.

We assume that the probability that the user converts (buys a product from A) is an arbitrary function of the number of times she has seen A's ad. We denote this probability by λ_j, where j is the number of times the user has seen A's ad. Typically, λ_j is unimodal, i.e., it increases at the beginning to reach a peak, and then decreases, although we will not make any such assumption.

Advertiser A's value per conversion is denoted by v. In the next section, we will discuss the problem of optimal allocation of ad space (to A or to the competitor). This can be viewed as the auctioneer's problem when trying to choose between A and its competitor to maximize social welfare, or A's problem when designing a bidding agent to submit a per-impression bid on its behalf each time. As it turns out, these views are equivalent.

The optimal allocation problem is one side of the multiple attribution problem. The other side is the problem of distributing A's payment among the publisher on which A's ad is displayed. This is an important problem when each of these pages is owned by a possibly different publisher, which is a common case in marketplaces like Google's DoubleClick Ad Exchange or Yahoo!'s Right Media Exchange [13]. We will discuss publisher fairness criteria in Section 6.

4 The Ad Allocation Problem

Given the values of the parameters of the model defined in the previous section (i.e., q, λ_i's, and the distribution of R), the goal of the ad allocation problem is to decide when to show A's ad to maximize the expected social welfare. Here the social welfare is the sum of the values that A and its competitor derive. Another way to look at this problem is to assume that at its core, the ad space is allocated through a second-price pay-per-impression auction[1], and conversion-seeking advertisers like A need to participate in the auction through a bidding

[1] This is the case in marketplaces such as Yahoo!'s Right Media Exchange.

agent that bids a per-impression value for each auction. The objective of such a bidding agent is the value to A minus its cost, which is equal to R if A wins. The difference between this objective and the objective of social-welfare maximizing auctioneer is an additive term equal to expectation of the sum of the R values. Therefore, the two optimizations are the same. In this section and the next, we solve this optimization problem by modeling it as a Markov Decision Process (MDP) and solving the corresponding Bellman equations [16]. We will also derive the value that A's bidding agent should bid to achieve the optimal outcome.

MDP formulation. We can define an MDP as follows: for each j, where $j-1$ represents the number of ad views so far, we have three states a_j, b_j, and c_j. The state a_j represents the probabilistic state right before the next time the user views a page on which an ad can be displayed. This state has a transition with probability q to the *quit* state (which is a terminal state), and another with probability $1-q$ to b_j. At b_j, the value of R is realized and we need to make a decision between not showing the ad, which would give a reward of R and takes us back to the state a_j, or to show the ad, which would take us to the state c_j. This is a probabilistic state with probability of transition of $(1-\lambda_j)$ to a_{j+1} (corresponding to the non-conversion event) and probability of transition of λ_j to a terminal *convert* state. The reward of this transition is v (the value of conversion) plus the value of the infinite sequence of alternative ads starting from this point. Since the number of page visits follows an exponential distribution, this value is $v + (1-q)E[R]/q$. The state a_1 is the starting state. Figure 1 illustrates the process.

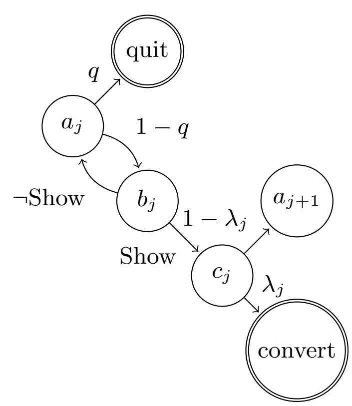

Fig. 1. Multiple Attribution MDP.

The Bellman Equation. We denote the total social welfare we obtain from this user starting from the state b_j by V_j. At this state, we need to choose between showing the competitor's ad or showing A's ad. In the former case, we immediately get a value of R and with prob. $1-q$ will be taken back to the state b_j. Therefore, the expected value in this case is $R+(1-q)V_j$. In the latter case, with probability λ_j a conversion happens, which results in a value of v for the conversion plus $(1-q)E[R]/q$ for the sequence of competitor ads we can show afterward. With probability $1 - \lambda_j$, we get no conversion and will be taken to the state b_{j+1} with probability $1 - q$. Therefore, the expected value in this case is $\lambda_j(v + (1 - q)r/q) + (1 - \lambda_j)(1 - q)V_{j+1}$, where $r = E[R]$. To summarize:

Proposition 1. *The values V_j of the expected total value starting from the state b_j satisfy the following equation:*

$$V_j = E_R[\max(R + (1 - q)V_j, \lambda_j(v + \frac{(1 - q)r}{q}) + (1 - \lambda_j)(1 - q)V_{j+1})], \quad (1)$$

where $r = E[R]$. The value V_1 indicates the maximum expected social welfare in our model.

5 Computing the values

In this section, we show how (1) can be simplified to a recurrence relation that can be used to compute V_j's. This recurrence has a simple form, but involves a function that is, in general, non-linear (depending on the distribution of R), and therefore its solution cannot be written in closed form. However, we can do this in the case that R is a constant. Also, we derive the values that a bidding agent that participates in a pay-per-impression auction on behalf of A should bid for each impression.

5.1 The general recurrence

We start with the Bellman equation (1) and simplify it in each step, eventually writing it in terms of a particular function that captures the effect of q and the distribution of R. First, we rewrite the equation in terms of new variables $W_j := (1 - q)(V_j - r/q)$. Intuitively, W_j is the maximum value starting from the state a_j, minus the value starting from this state without the presence of advertiser A. By replacing V_j's by W_j's in (1) we obtain

$$\frac{W_j}{1 - q} = \mathrm{E}[\max(R + W_j, \lambda_j v + (1 - \lambda_j)W_{j+1})] - r. \quad (2)$$

Before simplifying this equation further, notice that this means that in the optimal allocation, the advertiser A wins if and only if $R + W_j \leq \lambda_j v + (1 - \lambda_j)W_{j+1}$. Thus,

Proposition 2. *In the optimal allocation, at a point where the user has already seen A's ad $j - 1$ times, the next impression will be allocated to A if and only if the cost of this impression (R) is at most $\lambda_j v + (1 - \lambda_j)W_{j+1} - W_j$.*

To write (2) in a simpler form, we define $h(x) := \mathrm{E}[max(R, x)]$. Clearly, $h(.)$ is a function that only depends on the distribution of R. After subtracting W_j from both sides of (2), we can write this equation as

$$\frac{qW_j}{1 - q} = h(\lambda_j v + (1 - \lambda_j)W_{j+1} - W_j) - r. \quad (3)$$

Note that h is by definition a continuous non-decreasing function. For a value $\beta \geq 0$, consider the following equation in terms of the variable x: $qx/(1-q) = h(\beta - x) - r$. At $x = 0$, the right-hand side of this equation is $h(\beta) - r = h(\beta) - h(0) \geq 0$ and the left-hand side is zero. At $x = \beta$, the right-hand side is $h(0) - r = 0$ and the left-hand side is non-negative. Therefore, since the right-hand side of the equation is non-increasing in x, the left-hand side is strictly increasing, and both sides are continuous functions of x, this equation has a unique solution in $[0, \beta]$. We denote the value of this solution by $u(\beta)$.

Proposition 3. *For any value of q and distribution of R, the function $u(.)$ is well-defined, non-decreasing, and continuous, and satisfies $\forall\ \beta:\ u(\beta) \in [0, \beta]$.*

Note that $u(.)$ is defined purely in terms of the distribution of R and the value of q, and in fact, it captures all the information about these parameters that is relevant for the allocation problem. Using this function, (3) can be rewritten as:

$$W_j = u(\lambda_j v + (1 - \lambda_j) W_{j+1}). \tag{4}$$

Obtaining an explicit formula for V_j is only possible if $u(.)$ has a simple form. Unfortunately, this function is often complex and non-linear.[2] However, the above equation gives a straightforward way to compute W_j's numerically: start with a large enough j^* so that $W_{j^*} = 0$, and then move backward to compute W_j for $j = j^* - 1, \ldots, 1$. Such a value of j^* exists in most realistic scenarios; for example, any j^* such that for all $j > j^*$, $\lambda_j v$ is less than the minimum of R (say, the value of the reserve price) suffices. To summarize,

Theorem 1. *Let W_j's be the values computed using (4). Then the optimal allocation can be obtained by submitting a per-impression bid of $bid_j := \lambda_j v + (1 - \lambda_j) W_{j+1} - W_j$ on behalf of A in a state where the user has already seen the ad $j - 1$ times. The social welfare achieved by this mechanism is $r/q + W_1/(1-q)$.*

5.2 Closed-form solution for constant R

In the case that R is a constant r, we can significantly simplify the recurrence (4). First, note that by definition, $h(x) = \max(r, x)$. Therefore, $u(\beta)$ is the solution of the equation $qx/(1-q) = \max(\beta - x - r, 0)$. It is easy to see that when $\beta \geq r$, the solution of the above equation is $(1-q)(\beta - r)$, and when $\beta < r$, this solution is zero. Therefore, $u(\beta) = (1-q)\max(\beta - r, 0)$. This gives

$$W_j = (1-q)\max(\lambda_j v - r + (1 - \lambda_j) W_{j+1}, 0). \tag{5}$$

To solve this recurrence, we can expand W_{j+1} in the above expression, and iterative. This results in the following explicit expression, which can be easily verified by induction using the above recurrence (5):

$$W_j = (1-q) \max_{l \geq j-1} \left\{ \sum_{s=j}^{l} (\lambda_s v - r) \psi_s / \psi_j \right\}, \tag{6}$$

[2] For the uniform distribution, $u(.)$ is the solution of a quadratic equation; for the exponential distribution $u(.)$ cannot be written in closed form.

where $\psi_i := \prod_{t=1}^{i-1}(1-q)(1-\lambda_t)$ is the probability that the user visits at least i times and each time (except possibly the last time) does not convert on A's ad.

In the above expression an empty sum is defined as zero and an empty product is defined as one. So the final solution can be written as follows:

$$V_1 = \frac{r}{q} + \max_{l \geq 0} \left\{ \sum_{s=1}^{l} (\lambda_s v - r)\psi_s \right\}. \tag{7}$$

To summarize:

Theorem 2. *Let l^* be the value of l that achieves the maximum in (7). Then in the optimal allocation, A's ad is shown until the user converts or she sees the ad l^* times. After a conversion happens or this number of ad views is reached, the competitors ad is shown.*

6 Pricing and Publisher Fairness

In the last section, we showed how we can design a bidding agent that translates the advertiser A's values into an effective bid per impression every time there is an advertising opportunity. If this advertiser could pay per impression (we will call this the *pay-per-impression scenario*), this would have been the end of the story: on each auction, we would use the bidding agent to bid, and if A wins based on this bid, she will pay the value of the competitor's bid R. This value would be disbursed to the publisher responsible for that impression. It is not hard to see that this scheme is equivalent to the VCG mechanism from A's perspective (i.e., it allocates the good optimally and charges A the externality she imposes on others), and therefore A has incentive to truthfully report her value per conversion v. Also, the mechanism seems intuitively "fair" for publishers.

However, some advertisers are strict pay-per-conversion advertisers. For these advertisers the payment scheme should satisfy the following property:

Ex-Post Individual Rationality (Ex-Post IR): At any outcome where a conversion has not happened, A does not pay anything. At an outcome where a conversion has happened, A pays at most her value per conversion v.

In addition to the above, we require *Efficiency* (getting the optimal allocation characterized in the last section) and *Incentive Compatibiltiy (IC)*. Note that these two properties imply that in expectation, the amount the advertiser must be charged is the externality it imposes on the others. This is equal to the sum of R on impressions where A's ad is shown. In other words, in expectation, the mechanism should charge the same amount as in the pay-per-impression scenario. The challenge is to implement this while respecting Ex-Post IR.

As we will show in the next subsection, this can be achieved with a simple uniform pricing. This method is simple and works well when there is only one publisher (so there is no issue of fairness). In Section 6.2, we define and study a natural notion of fairness when there are multiple publishers. We will show that there are instances where the uniform pricing method *cannot* result in a fair distribution of payments to publishers. On the positive side, in the case of constant R, we will show that the problem can be formulated as a network flow problem, and will use the maximum-flow minimum-cut theorem to prove that a fair, ex-post IR, and incentive compatible payment rule always exists. As this is a special case of the max-flow min-cut problem, we will also be able to give a simpler and faster algorithm for computing the payments.

6.1 The uniform pricing method

The idea of the uniform pricing method is to charge the same amount for all conversions, regardless of how many ad impressions A gets prior to the conversion. This uniform cost is set at a level to get the advertiser to pay the right amount in expectation. Using the optimality of the allocation, we can show that this scheme satisfies Ex-Post IR. We first illustrate this in the ase of constant R.

First, note that $W_1 \geq 0$. This can be seen directly from the defintion of W_1 and V_1 as the optimal solution of the MDP, or from Equation (6). Let ℓ be the value that maximizes (7). Thus we have $\sum_{s=1}^{\ell}(\lambda_s v - r)\psi_s \geq 0$, or, equivalently:

$$v \geq r \cdot \frac{\sum_{s=1}^{\ell} \psi_s}{\sum_{s=1}^{\ell} \lambda_s \psi_s}. \tag{8}$$

Now consider the expected externality imposed by the advertiser on others. The probability that the ad is shown exactly ℓ times is ψ_ℓ. For some $s < \ell$ the probability that it is shown exactly s times is $\psi_s - \psi_{s+1}$. Therefore, the total expected externality imposed on others by the advertiser is $r\sum_{s=1}^{\ell} s\psi_s - r\sum_{s=1}^{\ell-1} s\psi_{s+1} = r\sum_{s=1}^{\ell}\psi_s$. On the other hand, the probability that the user converts after the i-th view is $\lambda_i\psi_i$. Thus the total probability of conversion is $\sum_{s=1}^{\ell}\lambda_s\psi_s$. Therefore if for each conversion, we charge the advertiser $r \cdot \frac{\sum_{s=1}^{\ell}\psi_s}{\sum_{s=1}^{\ell}\lambda_s\psi_s}$, the expected payment of the advertiser will be equal to the externality it imposes on others (i.e., the IC payment). Also, by Equation (8), the payment per conversion is at most v, and hence Ex-Post IR is also satisfied.

This method can be applied in the general case (when R is not a constant): On any conversion, independent of the history of impressions that lead to this conversion, we charge the advertiser an amount equal to

$$\mathit{price} := \frac{E}{P_{conv}}, \tag{9}$$

where E is the expected total externality that A imposes on the competitors, and P_{conv} is the overall probability of conversion for A. By definition, with this charging scheme in expectation A pays $\mathit{price} \times P_{conv} = E$, which is the incentive compatible payment. To show that the above price satisfies Ex-Post IR, we compare this scenario with the pay-per-impression scenario defined at the beginning of Section 6. It is easy to see that the outcome in both cases is the same and A's payment is also the same in both scenarios in expectation. Therefore, since A's utility in the pay-per-impression scenario is non-negative, it is non-negative here too, implying that $\mathit{price} \leq v$.

6.2 Publisher Fairness

There are two main motivations for studying the multiple attribution problem: the first is to ensure the efficiency of the market outcome, and the second is to ensure that each ad publisher who has contributed in the purchase funnel that has lead to a conversion gets a fair share of the conversion price. So far, we have been concerned with the first aspect: efficiency. In this section we turn to the second aspect: fairness among publishers.

We first need to define the notion of fairness for publishers. Our definition is motivated by the hypothetical pay-per-impression scenario defined at the beginning of Section 6. In this scenario, each publisher who displays A's ad, receives a payment equal to the opportunity cost of this impression. We define fairness in our setting by requiring the same payments *in expectation*:

Publisher Fairness. For each i, the expected value the i'th publisher receives from A is equal to the expected opportunity cost (R) of this publisher conditioned on A winning.

Note that this is a natural property to require, since it is natural for the publisher to request to be paid an amount at least equal to the opportunity cost of the impressions it provides (if this is not satisfied, the publisher could refuse to accept pay-per-conversion advertisers), and since the advertiser's payment is the total externality it imposes on the competitors, no publisher cannot hope to get more than its expected opportunity cost without hurting another publisher.

As we will show below, Publisher Fairness imposes a non-trivial constraint on the payments. In fact, for some payment rules like the uniform scheme defined in Section 6.1, it is not possible to distribute the payment among the publishers in a way that satisfies Publisher Fairness. To illustrate this and prepare for the result of the next section (showing that for constant R, there is a payment rule satisfying Publisher Fairness), we focus on the case of constant R, and introduce some notations.

We number the publishers in the order the user visits ad-bearing pages. Let x_{ij} be the payout to publisher j if the conversion occurs after precisely i views. This quantity is only defined for $i \geq j$, since for $i < j$, the user will either never visit publisher j, or visit this publisher after she is already converted. Also, we only define the variables x_{ij} for $i, j \leq \ell$ where ℓ is the index that maximizes the value in Equation (7), since after this index, A's ad will not be shown.

We can write our desired properties in terms of the $x_{i,j}$ variables. First, we formulate the Publisher Fairness property. For every publisher $j = 1, \ldots, \ell$, conditioned the user visiting j, the probability that it visits exactly i publishers ($i \geq j$) and then it converts is precisely $\psi_i \lambda_i / \psi_j$. Thus, the total expected payment to j, conditioned on the user visiting j can be written as $\sum_{i \geq j} x_{ij} \psi_i \lambda_i / \psi_j$. Therefore the Publisher Fairness property can be written as follows:

$$\forall j : \quad \sum_{i \geq j} x_{ij} \psi_i \lambda_i / \psi_j = r. \tag{10}$$

This property also implies that the payments are incentive compatible: since for each publisher the total payment of A is equal to the externality A imposes on its competitors on this publisher, the total expected payment of A is also equal to the total expected externality it imposes on the competitors. Therefore, all that remains is to formulate the Ex-Post IR property. The total payment of A in case a conversion happens after precisely i impressions is $\sum_{j \le i} x_{ij}$. Therefore, Ex-Post IR is equivalent to the following.

$$\forall i : \quad \sum_{j \le i} x_{ij} \le v. \tag{11}$$

We leave the proof of the following theorem to the full version of the paper.

Theorem 3. *Consider the optimal allocation with the uniform pricing rule defined in the last section. There are instances in this mechanism where there is no way to distribute the advertiser's payment among the publishers in a way that satisfies Publisher Fairness.*

6.3 Fair payments via max-flow min-cut

The main result of this section is the existence of a fair payment rule when R is constant. The proof (omitted due to lack of space) is based on formulating the constraints as flow constrains and using the max-flow min-cut theorem.

Theorem 4. *When R is a constant, the optimal allocation rule can be supplemented with a payment scheme that satisfies Incentive Compatibility, Ex-Post Individual Rationality, and Publisher Fairness.*

7 Conclusion

In this work we showed how myopic optimization by the publishers can lead to inefficient allocations in the case when displaying an impression for an advertiser changes the user's conversion probability on subsequent visits. We formulated the optimal allocation problem in this setting as a Markov Decision Process and derived the optimal allocation and a way to translate the advertiser's per-conversion value to bids for each impression. We then studied how the advertiser should be charged in the case of a conversion, and how this charge should be split between publishers in order to achieve incentive compatibility and individula rationality for the advertiser and fairness for the publishers.

Our model is fairly general, yet simple enough to be practical. Perhaps the most important assumption in the model, which is sometimes inaccurate, is that we assumed that the conversion probability depends *only* on the number of views, and *not* on the identity of the publishers that display the ad to the user. One can imagine generalizing this notion, in a manner similar to the separable click-through rate model of sponsored search – that the probability of conversion is a separable function of the number of user visits and the identity of the publisher. Another way to relax this assumption is to assume each publisher has a weight, and the conversion probability of the user at each point is a function of the total weight of the publishers that have shown the ad to the user. When all weights are 1, this model reduces to our identical publisher model. We leave this as an interesting open problem.

References

1. Gagan Aggarwal, Ashish Goel, and Rajeev Motwani. Truthful auctions for pricing search keywords. In *Proceedings of the 7th ACM conference on Electronic commerce*, EC '06, pages 1–7, New York, NY, USA, 2006. ACM.
2. Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. Group formation in large social networks: membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 44–54. ACM Press, 2006.
3. Thomas E. Barry. The development of the hierarchy of effects: An historical perspective. *Current Issues and Research in Advertising*, 10:251–295, 1987.
4. Peter J. Danaher. Advertising models. In Berend Wierenga, editor, *Handbook of Marketing Decision Models*, volume 121 of *International Series in Operations Research and Management Science*, pages 81–106. Springer US, 2008.
5. Josh Dreller. In the trenches SEM pre-click & post-click double feature: Conversion attribution and Q&A with analytics guru Eric Peterson. Search Engine Land, `http://searchengineland.com/qa-with-eric-t-peterson-on-web-analytics-for-sem-and-a-look-into-conversion-attribution-15544`, November 2008.
6. Josh Dreller. Research brief: Conversion attribution. Fuor Digital, `http://fuor.com/resources/leadership/briefs/conversion-attribution-research-brief`, December 2008.
7. Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97(1):242–259, March 2007.
8. Daniel C. Fain and Jan O. Pedersen. Sponsored search: A brief history. *Bulletin of the American Society for Information Science and Technology*, 32(2):12–13, December/January 2006.
9. Ayman Farahat. Privacy preserving frequency capping in internet banner advertising. In *Proceedings of the 18th international conference on World wide web*, WWW '09, pages 1147–1148, New York, NY, USA, 2009. ACM.
10. ClearSaleing Inc. Attribution management. `http://www.attributionmanagement.com/`.
11. R. Lewis. Where's the 'wear-out'? Working paper. Yahoo! Research, 2011.
12. Mohammad Mahdian and Kerem Tomak. Pay-per-action model for online advertising. In *Proceedings of the 3rd international conference on Internet and network economics*, WINE'07, pages 549–557, Berlin, Heidelberg, 2007. Springer-Verlag.
13. S. Muthukrishnan. Ad exchanges: Research issues. In *Proceedings of the 5th International Workshop on Internet and Network Economics*, volume 5929 of *Lecture Notes in Computer Science*, pages 1–12. Springer-Verlag, December 2009.
14. PricewaterhouseCoopers. IAB internet advertising revenue report: 2009 full-year results. `http://www.iab.net/media/file/IAB-Ad-Revenue-Full-Year-2009.pdf`, 2010.
15. Edward K. Strong. *The Psychology of Selling Advertising.* McGraw-Hill, 1925.
16. D. J. White. *Markov Decision Processes.* Wiley, 1993.