

Signaling Games and Ad Exchanges

David Kasofsky

September 1, 2015

Contents

1	Introduction	4
1.1	Motivation	4
2	Ad Exchanges	5
2.1	Publishers and Advertisers	5
2.2	Role of the Ad Exchange	5
2.3	Third Party Data	7
2.4	Attribution	8
2.5	Summary	9
3	Signaling Games	11
3.1	Basic Model	11
3.2	Equilibria	13
3.3	Costling Signaling	13
3.4	Iterated Signaling Games	13
3.5	Recommenders and Verifiers	13
4	Signaling Game Ad Exchange	14
4.1	Consumers	14
4.2	Advertisers	14
4.3	Auction Game	14
4.4	Ad Game	14
4.5	Third Party Data as Recommendation	14
4.6	Attribution as Verification	14
5	Ad Exchange Simulation	15
5.1	Agent-based Model	15
5.2	Simulation Details	15
5.3	Implementation	15

6	Simulation Results	16
6.1	Summary	16
6.2	Naive Attribution	16
6.3	Improved Attribution	16
7	References	16

1 Introduction

1.1 Motivation

2 Ad Exchanges

2.1 Publishers and Advertisers

In online advertising, publishers are those with ad space to sell. Usually this is because they control websites which create and publish content, much of which is freely accessible, and wish to monetize the site traffic the content generates. Without ad exchanges, publishers could only sell their ad space through deals negotiated directly with individual advertisers.

A publisher may represent a single site or multiple sites with different types of content. For our purposes, we will also consider Search Engines such as Google and Bing to be publishers though they are treated somewhat differently in practice. The challenge for publishers is to sell their ad space at the highest price.

Advertisers purchase ad space in order to interact with consumers via advertisements. They are usually businesses or advertising agencies acting on behalf of businesses. Advertisers allocate a portion of their total marketing budget to be spent online. The challenge for advertisers is to find cost effective ad space on which to spend their online budget.

Common criteria to evaluate the potential of a publisher's ad space include the reputation of the publisher, the consumer demographics that commonly visit the publisher's sites, and the alignment of the site's content with the advertiser's products. Once ad space has been purchased its performance may be evaluated using a process known as Attribution, which will be discussed later in great detail.

2.2 Role of the Ad Exchange

An Ad Exchange is both an aggregator of publisher inventory and a marketplace for that inventory. This is good for publishers because listing their inventory on the exchange makes it easily available to numerous advertisers. Publishers provide additional metadata about their listings, such as the type of content the ads will appear with or the demographics who are likely to visit the publishers' sites. This in turn makes it easier for advertisers to find appealing ad space.

In order to list their inventory with an exchange, a publisher signs up with and receives code from the exchange (e.g. some JavaScript) that will allow the exchange to dynamically insert ads

on the publisher's sites. For example if the publisher's site layout includes a banner ad along the side of the page, the publisher will put the exchange's code in the spot the banner ad would go.

Advertisers buy the listed ad space in auctions. Whenever a user visits a publisher's page with the exchange code on it, the exchange is signaled that there is an opportunity to show an ad. Showing an ad to a user is called an "impression" and so we say an impression is available. The advertisers on the exchange then evaluate the impression and places bids according to their evaluation. The winning advertiser's ad is shown on the page. This whole process occurs very quickly so that the ad is loaded along with the rest of the content.

Because publishers provide some metadata about their ad space, advertisers can include it in their evaluation. This highlights another important feature of ad exchanges: the ability to set bidding strategies according to impression metadata. As an impression auction must be resolved quickly, any logic to determine bid amounts is set ahead of time and is known as a bidding strategy. The following are examples of potential bidding strategies for an advertiser:

- Bid \$0.50 for impressions on sites with many male visitors
- Bid \$1 for impressions on sites with many male visitors between the ages of 20 and 40
- Bid \$1 for impressions on Sports sites
- Bid \$2 for impressions on Baseball sites but not between 1AM and 6AM

These illustrate how publisher metadata can influence advertisers' bidding strategies as well as publisher-independent factors such as time of day. Thus advertisers may have significantly differing evaluations of individual impressions. Such variation is the primary motivation for the use of auctions.

The use of auctions highlights two major difference between buying on the exchange and buying directly from publishers. First, ads are usually bought from publishers using a constant rate, such as a cost per click (CPC) or a cost per thousand impressions (CPM). Second, some number of impressions are usually guaranteed. Thus an advertiser may pay a \$3 CPM for 1 million impressions per month, or exactly \$3,000 per month. This is a common type of directly negotiated deal. Since ad exchanges use auctions to sell impressions as they become available, there is no guaranteed number of available impressions and the price may vary from impression to

impression. Therefore online advertising campaigns that spend heavily on ad exchanges usually require significantly more oversight and management than those of mostly direct deals.

2.3 Third Party Data

Third party data is data about consumers that has been collected and aggregated by third parties, i.e. neither publishers nor advertisers. In digital advertising the businesses that do so are known as Data Management Platforms (DMPs). A brief description is available at [2]. Examples of such companies are Lotame and BlueKai. Generally these companies gather data about consumers by partnering with various large e-commerce sites and monitoring activity on the site. DMPs then make their consumer data available for purchase by advertisers who use it to tailor marketing strategies, e.g. by targeting consumers with specific features.

Whereas impression metadata only describes the site the ad will appear on, DMP data describes the consumer visiting the site. Consider the following bidding strategies:

- Bid \$1 on consumers on Sports sites
- Bid \$1 on consumers with interest in Sports
- Bid \$2 on consumers with interest in Sports on Sports sites

The first is a bidding strategies based on impression metadata only and the second based only on DMP data. The third combines both impression metadata and DMP data and shows more information about an impression can make it more valuable.

A major use of third party data is to supplement impression metadata in exchange auctions. Ad exchanges integrate with DMPs so that when an impression become available, the DMP's data can also be made available to the advertisers in their evaluations. Advertisers pay a premium in order to use data from a DMP. This means that

Much third party data is collected and organized using cookies. First, a consumer is given a uniquely identifiable cookie the first time they visit a site which is monitored by a DMP. When that consumer visits certain pages or perform certain activities on the site or others monitored by the DMP, the DMP stores information related to the cookie id. For example, if a consumer visits a sporting goods site then the DMP may associate that consumer with the "sports" demographic.

Similar tracking is also used by ad servers. The first time a consumer receives an ad from a particular ad server, they are also given a uniquely identifiable cookie. This can be used to track which ads a consumer has been served. This will be seen to be critical in our discussion of Attribution.

Additionally, advertisers can use the ad server's tracking to associate data to cookie ids. Such information is known as first party data. A common tactic known as retargeting effectively flags consumers which have visited the advertiser's website and then pursues these consumers aggressively as they continue to browse the web. Many users have had creepy experiences where they encounter ads for products or brands they had recently viewed- this is retargeting.

It is important to note that cookie-based tracking has many issues as cookies cannot uniquely identify a single consumer across multiple devices or browsers. Other tracking methods are emerging to address this. For example, the Google Chrome browser encourages users to sign in to the browser with their Google account. If a consumer does this, then Google achieves very reliable cross-device tracking of that consumer. A full discussion of such methods is out of the scope of this paper but further reading can be found in [FIND SOME PAPERS FOR THIS](#)

2.4 Attribution

Before executing various advertising strategies, advertisers must decide on the events and activities they wish to cause by showing ads. When such an event occurs it is called a conversion and the associated consumer is said to have converted. The best example of conversions are purchases. Advertisers may record additional information such as the revenue generated by the conversion.

It is common for advertisers to define multiple types of conversions. Merely visiting the advertiser's homepage may be considered a type of conversion. This could be thought of as a filtering conversion type that identifies consumers that have already demonstrated an interest in the advertiser by visiting their site. Another common scheme is to have multiple conversion types which act as checkpoints within a longer process, e.g. for businesses that require customers to submit applications for approval.

Since causing specific types of conversions is an advertiser's goal, the advertiser must be able

to associate conversions with ads in order to evaluate advertising strategies. In its simplest form, this is accomplished using the same tracking which associates ads to consumers. By assigning each consumer a unique identifier, usually a cookie id, the advertiser records which ads have been shown to that consumer as well as which, if any, conversions that consumer has performed.

However, it is not enough to simply associate ad interactions and conversions to consumers. Advertisers not only wish to know which ads a consumer has interacted with but also the effect each ad had. Consumers frequently interact with numerous ads and it is not clear which if any of the ads significantly impacted the consumer's decision to convert. If advertisers could accurately divide the credit for causing a conversion among the preceding ad interactions they could also do so with the associated revenue and thereby calculate the ROI for each ad.

The process of allocating credit for conversions to the preceding ad interactions is called attribution. A particular mechanism for doing so is called an attribution model. The industry standard attribution model is Last Click. It attributes all credit for a conversion to the last-clicked ad or the last-viewed ad in lieu of a click. Plainly this does not capture the full detail of ad interaction but it has some motivating factors. It is very simple to understand as well as to implement. Variations include few parameters such as the maximum time between ad interaction and conversion for which the ad remains eligible. For instance, a click may become ineligible for credit after ten days but a view may become ineligible after five.

More sophisticated models are available to advertisers, usually from third parties. Example of such companies are Visual IQ and Adometry. These models are said to be data-driven or algorithmic as compared to Last Click which is rule-based. Such models incorporate a statistical analysis to determine the chance of conversion given interactions with certain ads or combinations of ads and allocate credit for conversions accordingly. However many advertisers do not employ more sophisticated methods because of the additional cost, implementation difficulty, and lack of transparency when using a third party.

2.5 Summary

Publishers, Advertisers, DMPs, Attribution, and Ad Exchanges are the five pillars of digital marketing. Publishers represent supply while Advertisers represent demand. DMPs provide

information about consumers while Attribution provides evaluations of advertising strategies. Ad exchanges are fluid, high-tech marketplaces which facilitate inclusion of third party data and allow timely responses to ad performance. Above all, ad exchanges make it easier for publishers and advertisers to meet their goals and to try new things.

3 Signaling Games

3.1 Basic Model

The crux of the signaling game is information asymmetry. One player has some information which is hidden from the other and wishes to convince the other to do something. Thus the informed player, called the Sender, must choose how to make his case while the second player, called the Receiver, must decide how to act based on the message from the Sender. As in typical game theory, both players wish to maximize their utility in this exchange. One can succinctly think of a signaling game definition requiring a sender, a receiver, some hidden information, the sender's possible messages, and the receiver's possible actions.

Such games are found to resemble interactions in nature, perhaps most famously described by Grafen [1] in his formalization of ideas from Zahavi [5]. Here we see an example of signaling games used to describe mate selection. Consider a species in which the females choose which males to mate with and naturally want to select the best males. It is common for the male to signal his fitness by demonstrating a handicap, e.g. the cumbersome antlers of male deer. Since it is easier for healthier bucks to carry heavy antlers, antlers size generally indicates the buck's true fitness. In this case the sender is the buck, the receiver is the doe, the hidden information is the fitness of the buck, the message is the size of antlers, and the action is to mate or not.

Spence [4] gives an example of how such games can model the hiring process. Here the sender is an applicant, the receiver is an employer, the hidden information is the competency of the applicant, the message is the credentials of the applicant, and the action is the wage offered to the applicant. Applicants want to be hired with the highest wage and employers want to pay applicants appropriately. Similarly to the antlers of deer, higher credentials are easier to earn for more competent applicants and so the employer may base his wages on the applicant's credentials.

A final simple example is the call-fold game. The sender has either a winning hand or a losing hand and can send one of two messages: bet, or check. The type of hand is hidden from the receiver. The receiver must then either call or fold. Here are the payoff matrices:

Sender has Winning Hand

	Fold	Call
Bet	1, 0	2, -1
Check	1, 0	1, 0

Sender has Losing Hand

	Fold	Call
Bet	0, 1	-1, 2
Check	0, 1	0, 1

This signaling game is much like what happens in the final round of betting in a limit poker game. Here is a version that is more like a pot limit poker game with pot size p :

Sender has Winning Hand

	Fold	Call
Bet Amount b	$p, 0$	$p+b, -b$

Sender has Losing Hand

	Fold	Call
Bet Amount b	$p, 0$	$-b, p+b$

The value of b can range from 0 to p .

With a few examples to build our intuition, let us now give a formal definition for basic signaling games. We say basic here because we will extend this model in later games. Our definition is inspired by Sobel [3].

A Basic Signaling Game is a Bayesian Game. A Bayesian game is generally defined as a tuple $G = \langle N, \Omega, \langle T_i, \tau_i, A_i, C_i, u_i, p_i \rangle_{i \in N} \rangle$ where:

- N is the set of players
- Ω is the set of initial conditions.
- T_i is the type of player i .
- $\tau_i : \Omega \rightarrow T_i$ is a function that assigns types to players

- A_i is the set of actions for player i .
- $C_i \subseteq A_i \times T_i$ defines the actions available to player i when player i is of type T_i .
- $u_i : \Omega \times A \rightarrow \mathbb{R}$ is the payoff function for player i .
- p_i is the probability distribution over Ω for player i .

3.2 Equilibria

3.3 Costly Signaling

3.4 Iterated Signaling Games

3.5 Recommenders and Verifiers

4 Signaling Game Ad Exchange

4.1 Consumers

4.2 Advertisers

4.3 Auction Game

4.4 Ad Game

4.5 Third Party Data as Recommendation

4.6 Attribution as Verification

5 Ad Exchange Simulation

5.1 Agent-based Model

5.2 Simulation Details

5.3 Implementation

6 Simulation Results

6.1 Summary

6.2 Naive Attribution

6.3 Improved Attribution

7 References

References

- [1] Alan Grafen. Biological signals as handicaps. 1990.
- [2] Forrester Research. Digital advertising comes of age. 2012.
- [3] Joel Sobel. Signaling games. 2007.
- [4] A. Michael Spence. Market signaling. 1974.
- [5] Amotz Zahavi. Mate selection- a selection for a handicap. 1975.