ABSTRACT
The real-time bidding (RTB), aka programmatic buying, has recently become the fastest growing area in online advertising. Instead of bulking buying and inventory-centric buying, RTB mimics stock exchanges and utilises computer algorithms to automatically buy and sell ads in real-time. It uses per impression context and targets the ads to specific people based on data about them, and hence dramatically increases the effectiveness of display advertising. In this paper, we provide an empirical analysis and measurement of a production ad exchange. Using the data sampled from both demand and supply side, we aim to provide first-hand insights into the emerging new impression selling infrastructure and its bidding behaviours, and help identifying research and design issues in such systems. From our study, we observed that periodic patterns occur in various statistics including impressions, clicks, bids, and conversion rates (both post-view and post-click), which suggest time-dependent models would be appropriate for capturing the repeated patterns in RTB. We also found that despite the claimed second price auction, the first price payment in fact is accounted for 55.4% of total cost due to the arrangement of the soft floor price. As such, we argue that the setting of soft floor price in the current RTB systems puts advertisers in a less favourable position. Furthermore, our analysis on the conversation rates shows that the current bidding strategy is far less optimal, indicating the significant needs for optimisation algorithms incorporating the facts such as the temporal behaviours, the frequency and recency of the ad displays, which have not been well considered in the past.

Categories and Subject Descriptors
H.3.5 [Information Storage and Retrieval]: Online Information Services—Commercial services

General Terms
Economics, Measurement, Performance

Keywords
computational advertising, real-time bidding, bidding behaviours, conversion optimisation

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. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

1. INTRODUCTION
Online advertising is one of the most fast growing area in IT industry. Over the past 10 years, its revenue has increased from $6.0 billion in 2002 to $36.6 billion in 2012, with a compound annual growth rate of 19.7%. The display-related advertising revenues totalled $12 billion in 2012, with a 9% increase from 2011 [1]. In display advertising the most significant concept in recent years is real-time bidding (RTB), or programmatic buying, where advertisers have the ability of making decisions for every impression (auction). In this paper we report interesting findings from field studies in a production ad exchange, and discuss related research challenges. RTB is a demand-side oriented concept. However, we use data from both advertisers and publishers for a better understanding of the bidding behaviour.

We illustrate briefly the structure and history of display online advertising in Figure 1. Throughout the paper we used advertiser/bidder/buyer/demand side interchangeably to comply with the industrial tradition. Since an advertiser rarely complete against himself, we consider a campaign is equivalent to an advertiser in our study. The case is similar for publisher where we used seller/supply side/website interchangeably.

Before the emergence of RTB in 2009 (i.e., announcement of support by major ad exchanges [16]), the display advertising market was mainly divided by premium contracts (since 1994), which take more than 40% of impressions, and ad networks (since 1996), which take the rest of impressions that were usually referred to as remnant. With premium contracts, publishers negotiate and make deals with advertisers directly, c.f. Block I in Figure 1. Advertisers usually propose to buy certain amount of impressions from the given placements, regardless of the identities of users, when and how many times they have seen the ad, and so on. The purchase is totally about impressions and relies on the reputation of the publisher and reported audience profiles [11, 6, 2]. At the other end, the publisher needs to guarantee the delivery of impressions that have been agreed upon otherwise a penalty fee would be incurred [25, 27]. The pricing model used in contracts are mostly cost-per-mille (CPM). Since advertisers have no control over the inventories or users, it is more difficult to deploy goal-driven campaigns (e.g booking a ticket) than branding ones (e.g announcing a new product). These contracts are sometimes called guaranteed delivery display advertising [5].

With ad networks, publishers register placements and offer impressions from them for sale. Impressions are largely sold using the second price auction in ad networks. Adver-
tisers (or their delegates) also register with ad networks to participate in auctions. C.f. Block II in Figure 1. However, the impressions in ad networks are non-guaranteed, as opposed to premium contracts.

It is the ad network’s responsibility to understand the webpage and the user, and to select advertisers based on their pre-defined targeting rules. The understanding of webpage is usually referred to as contextual advertising, where ad networks crawl, parse, and extract keywords which summarise the target. Advertisers bid on these keywords which is very similar to sponsored search [24, 33, 3, 29]. A more advanced approach is to learn a model including various features of webpages, which could then be used to compute a relevance score of advertisers’ targeting criteria [17, 7, 8]. The understanding of user is usually referred to as behaviour targeting, where ad networks utilise the browsing history of a specific user to infer his interests, as well as geographical location, local time, etc for target matching [32, 23, 30].

In ad networks advertisers largely adopt the cost-per-click (CPC) or cost-per-acquisition (CPA) pricing models where they only pay when a certain goal is achieved. These choices reduce their risks thus are good for goal-driven campaigns. But then it is ad networks’ responsibility to optimise to maximise clicks or conversions. In order to take the measurement of performance into account, ad networks usually employ the generalised second price auction (GSP) [13] which allows them to apply bid biases (e.g., the quality score) that usually weight the historical clickthrough rate (CTR) or conversion rate (CVR) heavily.

For the publisher side, an important research topic is to allocate impressions to premium contracts and ad networks. If chooses contracts, the publisher also needs to decide which specific contract to fulfil when multiple contracts have overlapping targeting rules. It is possible that ad networks bring in more revenue, but if a publisher sends too many impressions to this revenue channel and fails to fulfil a contract, he needs to pay the goodwill penalty. In [27] the authors used a two-phase models to sample, compute a compact allocation plan, and assign impressions online to advertisers who submit contracts with overlapping targeting rules. This balancing problem was discussed in [25] by using a certainty equivalent control heuristic to show the necessity of using both channels to reach the global maximum revenue. In [15] it was generalised as an online stochastic optimisation problem where given a set of resources, demands for resource arrive online with associated properties. Given a general priori about the demands, one has to decide whether and how to satisfy a demand when it arrives. The goal is to find a valid assignment (strategy) with the maximum total payoff.

When there were more and more ad networks, a problem led to the birth of ad exchanges: the excessive impressions in some ad networks. It is preferable to have more demand than supply because intense competition leads to higher revenue of both ad networks and publishers. However, when there are plenty of impressions unsold, ad networks try hard to find buyers. Besides, a common practice for advertisers was to register with multiple ad networks to find cheap inventories, or at least to find enough impressions within their budget constraints. They only found that managing numerous channels difficult and inefficient (e.g., how to split the budget). Ad exchanges, like Google AdX, Yahoo! Right Media, Microsoft ad exchange, were created to address this problem by connecting hundreds of ad networks together, c.f. Block III in Figure 1. Advertisers now have a higher chance to locate enough impressions with preferred targeting rules; publishers may receive higher profit, too, because of more bidders potentially.

There are new research problems introduced by ad exchanges. In the pioneer work of [20] the author discussed several issues including the truthfulness of auctions [19], call-out optimisation [9], arbitrage bidding and risk analysis, etc. The most significant feature introduced by ad exchanges is real-time bidding, which queries bidders for a bid for every impression. Along with the query, ad exchanges send meta data of the webpage and the user. The exposure of this meta data enables advertisers to switch from the inventory-centric to user-centric optimisation. There are also noteworthy attempts to introduce concepts from finance market [28]. These attempts will make the ad exchange more mature and attractive.

When advertisers want to take advantage of RTB, they work with ad exchanges through 3rd party platform that are usually referred to as demand-side platform (DSP). DSPs are delegates of advertisers that answer bidding requests and optimise campaigns at the impression level. Comparing with ad networks, the advantages of using DSPs are: 1) advertisers do not need to manage their registration with many ad networks; 2) they can optimise at a finer granularity and a higher frequency because of local impression logs instead of aggregated reports from ad networks; 3) DSPs are also more customisable to better suit advertisers’ goals. For example, advertisers traditionally set a frequency cap on their campaigns (the maximum times to display the ad to the same user).
user). Now the cap can be applied to user groups, or even a single user for optimal efficiency.

At the other end, supply-side platforms (SSP) were created to serve publishers. Similarly, SSPs provide a central management console with various tools for publishers’ ultimate goal: the yield optimisation. For example, SSPs allow publishers to set a reserve price for a specific placement, or even against a specific advertiser. Some SSPs also allow publishers to have preference over bidders (bid bias).

In Figure 1 arrows with number describe the process of RTB:

1. An impression is created on publisher’s website;
2. The bidding request is sent to ad exchanges through ad network or SSP;
3. The ad exchanges query DSPs for advertisers’ bids;
4. The DSP could contact data exchanges for 3rd party user data;
5. The bid is generated and submitted; winner will be selected at ad exchanges, then at SSP. Following the reversed path the winner’s ad will be displayed on publisher’s website to the specific user.

Note that ad exchanges normally only query DSPs for bids, not ad networks. Also note that there are two types of links missing from the current marketplace. We find these links interesting and potentially useful. Firstly, SSPs normally do not contact data exchanges for 3rd party user data. However, if publishers could understand and model specific users, they will be in a better position of revenue optimisation (e.g. setting up an optimal reserve price for this auction based on the forecast of bidding activity with user data), although such query would incur cost. Secondly, ad exchanges normally do not send impressions to each other even if they remain unsold. This is largely due to business considerations and it leads to the fact that DSPs and SSPs register with multiple ad exchanges. A unified and interconnected marketplace is what we find attractive and it is certainly good for the whole eco-system.

In the remaining of this paper we represent our empirical study on a production ad exchange. We analyse the various datasets from both demand and supply side and report research challenges introduced by ad exchanges and RTB.

2. THE EMPIRICAL STUDY OF RTB

We first report the statistics of dataset used in the empirical study and represent some general analysis on bidders and bids. Then we discuss two specific aspects (bidding behaviour, frequency and recency factor) and associated problems (floor price detection, daily pacing, and selective bidding).

2.1 Dataset and the System’s Properties

We conducted our study in a production ad exchange based in the UK. Datasets were created both for demand and supply side. For advertisers, we sampled impression, click, and conversion logs from Feb to May 2013. In total we sampled 52,850,635 impressions, 72,958 clicks, and 37,978 conversions. Note that we sampled the conversions only and back traced associated clicks and impressions to make the dataset self-contained (i.e. any impressions and clicks associating with a sampled conversion would be in the dataset). Since these datasets were sampled, we also used various reports (e.g. analytical performance, site domain, attributed conversions, frequency/recency distribution, and so on) to understand the overall performance of advertisers.

We also recorded auction logs for placements from registered publishers of the ad exchange. In total we sampled 12,965,119 auctions from 50 placements, ranging from 14 Dec 2012 to 24 May 2013. These placements are from 16 websites of different categories, including news, finance, pc & console games, gadgets, sports, entertainment, and so on. Since auction logs were sampled, we also used various reports to understand the overall performance of publishers.

In the following of this section we report some statistics that we found interesting in the study.

2.1.1 Periodic Patterns

Due to the common daily routine of human there are always periodic patterns of our activity. In Figure 2 and Figure 3 we can observe clear daily patterns of both impressions and clicks from a single website. For number of impressions, there was also a weak weekly pattern that during weekends less visitors come to the website. There was also a daily pattern for conversions: during daytime the numbers of conversions were larger and the CVRs were higher compared to sleeping hours, c.f. Figure 5 and Figure 4.

Note there are two types of conversions: post-click that the user saw an impression, clicked, and finally converted (while remaining on the advertiser’s website and without any interruption); post-view that the user saw an impression, but did not click, but still converts (by visiting the website later on directly). These two types were plotted separately.

2.1.2 The Impact of the Soft Floor Price

In modern ad exchanges publishers are usually given the option to set a soft floor price $p_s$ along with the hard floor price (the traditional reserve price) $p_h$. The auction pro-
2.2 Bidding Behaviours

We now study the advertisers bidding behaviours influenced by the design of the RTB systems.

2.2.1 Bids' Distribution

A widely adopted assumption in optimal auction design research is that the private values of advertisers follow log-normal distribution [21, 14, 31, 22]. Bids are then generated from the private values. However, in the dataset we found that the 1st highest bid did not show strong log-normal distribution properties, nor did the 1st and 2nd highest bids, or all bids.

We split auctions by hour-of-day since bids vary greatly throughout a day, c.f. Figure 7. From 50 placements and 160 days we obtained 192k <placement, hour> tuples. For each tuple we fit highest bids, highest + second highest bids, and all bids.

We now study the advertisers bidding behaviours influenced by the design of the RTB systems.

2.2.2 The Daily Pacing

The daily pacing refers to the way that advertisers spend their budget in a single day. Usually an advertiser submits a daily budget for his campaign, and choose from spending it evenly throughout a day (uniform pacing) and as fast as possible (no pacing). The no pacing setup may lead to premature stop easily, which means the budget depletes too quickly, so advertisers cannot capture traffic later in the day, that may have high quality impressions. An instance of premature stop is illustrated in Figure 9 (day 1). The uniform pacing also suffers from the traffic problem: if high quality impressions appear in early part of the day, the pacing setup would not be able to capture all of them; if there is not enough traffic in the late part of the day, the pacing setup would not be able to spend all the budget (usually called under-delivery). In sum they are not good daily pacing strategies although being widely used in DSPs.

In Figure 8 we plot the hourly mean of number of bidders or impressions in that hour. Over all bidders or impressions the number of bidders and impressions are well fitted to Poisson distribution. The mean of the Poisson distribution is thus the daily budget for his campaign, and choose from spend-
their maximum in a day. Generally speaking, the number of bidders peaks in the morning but the number of impressions peaks afterwards. This lag indicates the unbalance of supply and demand of the market in certain hours. For example, there are more bidders competing over limited impressions in the morning, resulting in high winning bids as shown in Figure 7, which in turn costs more of advertisers. However, this higher cost does not necessarily lead to higher performance. From Figure 4 we can see that both post-view and post-clicks CVR peak in the evening, which argues that intense bidding activities in the early hours are not reasonable.

This distribution of number of bidders throughout a day may be due to the mixture of hour-of-day targeting and no daily pacing setup. Most of advertisers wish to skip the last night hours because of the low CVR. Some of them may use the no daily pacing setup to avoid the risk of under-delivery as we discussed before. When these bidders start to bid in the morning, they win lots of impressions with high bids, and they quit as their budgets depletes quickly. We leave the validation of this hypothesis to the future works.

A reasonable alternative is the dynamic pacing against the performance. It is intuitively correct to spend more budget in hours that generate more clicks or conversions, and less in low performing hours.

Note that this problem is not the same as a typical budgeted multi-arm bandit problem that has been discussed extensively in the literature [18, 12, 10]: 1) In this allo-

Figure 7: The time series snippet of winning bids and its hourly average of a single placement. The hourly average series peaks around 6-8am every day when there are less impressions but more bidders.

2.3 Conversion Rates and Selective Bidding

Considering the maintenance cost (e.g. servers and bandwidth) it is not wise for an advertiser to submit a bid every time he receives a bidding request. Among various factors that he uses to decide to bid or not, the frequency and recency factor are important ones that have not been given enough attention to.

2.3.1 The Frequency Factor

The frequency factor (or frequency cap, FC) defines how many times ads (a.k.a. creatives) would be displayed to a single user. It can be applied to campaign groups, campaigns, and creatives separately. For example, given a campaign with 3 ads, an advertiser could set FC(campaign) = 6, FC(ad1) = 2, FC(ad2) = 3, and FC(ad3) = 4 and these settings would work together. The same user would at most see ad1 twice, ad2 three times, ad3 four times, and all ads displayed to him together would not exceed six times.

The frequency factor is normally set based on historical data and needs to be adjusted constantly during the flight time of the campaign, because different campaigns ask for very different FCs, as illustrated in Figure 10. The campaign 1 from the left plot received the highest CVR with 6-10 impressions, which is also true for the cumulative CVR metrics. However the campaign 2 from the right plot received the highest CVR with 2-5 impressions. If the campaign 1 sets a frequency cap of 2-5 impressions, most of conversions would not be achieved. If the campaign 2 sets a frequency cap of 6-10, nearly half of impressions could be wasted. Therefore, a good FC setting is crucial to the efficiency of advertising.

Before setting an optimal FC, we need to find the right metrics to measure the efficiency of different FCs. See the example in Table 1 where we compare popular metrics. Assume there are 100 users; the CPM is fixed at 10; the advertiser’s conversion goal is worth 500. From the table we can see using different metrics could lead to very different
The long term impact: FC = 3 gives more conversions at low cost, and once these users are attracted they have a higher chance of converting again in future.

### 2.3.2 The Recency Factor

The recency factor (or recency cap, RC) helps to decide to bid or not based on how recently the ad was displayed to the same user. It also works at different levels including campaign groups, campaigns, and creatives. For example, an advertiser can set RC(campaign) = (1 hour) so that all ads from this campaign would be displayed to the same user only once in every hour. Similar to the frequency factor, RCs are useful to achieve the best advertising efficiency. For example, displaying ads intensively incurs high cost but little effect (or even getting people annoyed) for some campaigns (e.g. financial services). Users need to think, compare, then make decisions. A better strategy is to display the same ad right after the "thinking time" to remind users. However, for some campaigns (e.g. booking flight) users would convert very quickly or not convert at all, which requires relatively intense advertising.

Figure 11 plots the RC against CVR of two different campaigns in the dataset. Both campaigns show the highest

### Table 1: The comparison of different metrics against frequency caps (FC).

<table>
<thead>
<tr>
<th>fc</th>
<th>cvr</th>
<th>(cumulative) cvr</th>
<th>convs</th>
<th>cpa</th>
<th>roi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0</td>
<td>-</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>0.0150</td>
<td>0.0150</td>
<td>3</td>
<td>667</td>
<td>1.50</td>
</tr>
<tr>
<td>3</td>
<td>0.0067</td>
<td><strong>0.0167</strong></td>
<td>2</td>
<td>600</td>
<td>1.25</td>
</tr>
<tr>
<td>4</td>
<td>0.0025</td>
<td>0.0150</td>
<td>1</td>
<td>667</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>0.0020</td>
<td>0.0140</td>
<td>1</td>
<td>714</td>
<td>0.88</td>
</tr>
<tr>
<td>6</td>
<td>0.0000</td>
<td>0.0117</td>
<td>0</td>
<td>857</td>
<td>0.70</td>
</tr>
<tr>
<td>7</td>
<td>0.0000</td>
<td>0.0100</td>
<td>0</td>
<td>1000</td>
<td>0.58</td>
</tr>
</tbody>
</table>
Figure 11: The recency factor against CVR plot from two different campaigns. Both campaigns show the highest CVR at the 1-5 minutes level. However for the campaign 2 the CVR is still not negligible after a long time (14-30 days). If the advertiser uses a more strict RC setting (e.g. do not display ads to users who were first exposed 14 days ago) he could lose potential conversions. On the other hand, using a loose RC setting for the campaign 1 will only waste budget since the CVR is very low after 14 days.

CVR at the 1-5 minutes level. However for the campaign 2 the CVR is still not negligible after a long time (14-30 days). If the advertiser uses a more strict RC setting (e.g. do not display ads to users who were first exposed 14 days ago) he could lose potential conversions. On the other hand, using a loose RC setting for the campaign 1 will only waste budget since the CVR is very low after 14 days.

Another interesting observation is given in Figure 12. Similarly to the frequency factor, the analysis of efficiency of RCs requires metrics based on the understanding of advertising goal, which we do not repeat here.

The above analysis shows the importance of setting up proper FCs and RCs. At present these settings are at most at the creative level. With real-time bidding, advertisers can push them to a finer granularity: setting FCs and RCs for individual users. The individual FCs and RCs can then be used to decide to bid or not on a specific impression.

3. CONCLUSION

In this paper we introduced the history of real-time bidding and discussed research issues related to the demand side. In fact, RTB, ad exchanges, and DSP are conceptual ideas; what we really want to explore are the behaviour of advertisers and their delegates in the market, and the challenges brought by the impression-level bidding and user-centric bidding, distinguished from bulk buying and inventory-centric buying. Through analysis of datasets acquired from a production ad exchange, we discovered that floor price detection, daily pacing, and frequency/recency setting are problems not addressed. We explained their importance to advertisers however leave the development and evaluation of algorithms to the future works.

4. REFERENCES


